

Differentiating Engineered Tissue Images and Experimental Factors to Classify Cardiomyocyte Content

Samira Mohammadi^{1,2,#}, Mohammadjafar Hashemi^{1,3,#}, Ferdous Finklea^{1,4}, Elizabeth
Lipke^{1,5}, and Selen Cremaschi^{1,6,*}

¹212 Ross Hall, Department of Chemical Engineering, Auburn University, Auburn, AL
36849, U.S.A.

²+1(334)844-4970, szm0108@auburn.edu

³+1(334)844-2003, mzh0111@auburn.edu

⁴+1(334)844-2003, fbt0001@auburn.edu

⁵+1(334)844-2003, elipke@auburn.edu

⁶+1(334)844-4970, selen-cremaschi@auburn.edu

*Corresponding author

#These authors have contributed equally to this work

Abstract

In this study, we employed machine learning (ML) to classify the cardiomyocyte (CM) content on day 10 of the differentiation of human-induced pluripotent stem cell (hiPSC)-laden microspheroids using easily acquirable non-destructive phase-contrast images taken in the middle of differentiation and tunable experimental parameters. Scale-up suspension culture, use of engineered tissues to support stem cell differentiation, and CM production for improved control over cellular microenvironment in the suspension system need non-destructive methods to track engineered tissue development. The ability to couple images that capture experimenter perceived “good” or “bad” batches based on visualization at early differentiation time points with actual experimental outcomes in an unbiased way is a step towards building these methods. In recent

years, ML techniques have been successfully applied to identify critical process parameters and employ this information to build models that describe process outcomes in cell production and hiPSC differentiation. Building upon these successes, here, we utilize convolutional neural networks (CNNs) to build a binary classifier model for CM content on differentiation day 10 (dd10) for hiPSC-CMs. We consider two separate data sets as potential input features for the classification models. The first set includes phase-contrast images of microspheroid tissues taken on days 3 and 5 of the differentiation batches at different experimental conditions. The second set supplements the images with tunable experimental differentiation parameters, such as cell concentration and microspheroids' size. The CM content classes were *sufficient* and *insufficient*. The accuracy of the CNN classifier using images only was 63%. The addition of experimental features increased the accuracy to 85%, indicating the importance of tunable parameters in predicting CM content.

Keywords: cardiac differentiation, machine learning, convolutional neural networks, image processing, human pluripotent stem cell, engineered heart tissue

Impact Statement

Machine learning approaches were employed to predict the final cardiomyocyte content class (sufficient versus insufficient) of engineered cardiac tissue microspheroids produced through suspension-based cardiac differentiation of human-induced pluripotent stem cell-laden engineered tissue microspheroids. The models used specified experimental features and data collected using non-destructive, inexpensive methods, specifically phase-contrast images taken during the initial days of differentiation as inputs. The best model was a convolutional neural network trained using experimental features and differentiation day 5 images. It classified the cardiomyocyte content

with 85% accuracy and replicated and formalized experimenter's visual intuition about differentiation outcomes by incorporating images from early time points.

Introduction

Machine learning (ML) techniques are powerful tools to predict experimental outcomes using multi-factor experimental and/or visual data in bioprocesses for cell production and biomanufacturing. The information extraction from multi-factor data and images for these processes may not be possible without using ML techniques. Tracking and predicting final engineered tissue properties using non-destructive approaches and without the use of reporter genes is challenging. Particularly when starting with stem cells, the process of engineered tissue formation can take multiple weeks or longer; the ability to predict outcomes earlier and identify and tune key experimental parameters is essential to establishing more efficient approaches for commercial translation. Machine learning techniques have been applied to analyze and integrate information from high-resolution, three-dimensional images of developing tissues (1), for a range of applications in medical image computing, including histopathological image analysis (2) and for guiding bioprinting of engineered tissue scaffolds (3, 4). Employing machine learning techniques to detect characteristics of engineered tissues, extract information and hidden features and integrate image and non-image information has the potential to provide valuable guidance for identifying optimal initial experimental input parameters, *e.g.*, matrix composition, tissue structure, and cell seeding density, for engineered tissue formation and tracking and predicting outcomes, *e.g.*, endpoint cellular composition and potentially even tissue maturation and function.

Recently, there has been an increased focus on producing cardiac tissue from human-induced pluripotent stem cells to address cardiovascular diseases (CVDs), the number one cause

of mortality in recent years and the cause of one death every 36 seconds in the United States (5). Due to the limited regenerative capacity of human cardiomyocytes (CMs) and the loss of over one billion CMs post-heart attack (6), there is a considerable need for cardiomyocytes and cardiac tissue to replace and regenerate damaged heart tissue. Human-induced pluripotent stem cells (hiPSCs) have been demonstrated to be a promising cell source for hiPSC-derived CM (hiPSC-CM) production and may contribute to developing and testing therapeutics for CVD (7). Large-scale production of hiPSC-CMs in suspension culture is needed for use in cell therapy, drug discovery, and disease modeling (8).

Although suspension differentiation platforms hold considerable potential for scaling-up hiPSC-CM production, obtaining consistent and robust cardiac differentiation outcomes has been challenging (9) due to the complex and multifactorial process of 3-dimensional (3D) differentiation platforms (10). Accurately predicting cardiac differentiation outcomes, such as CM content on a certain differentiation day, for 3D platforms, especially using data collected using non-destructive and inexpensive methods, is an important enabling step for production scale-up. Through visual inspection, experienced experimenters have the opinion that it is possible to distinguish “good” hiPSC-CM production batches, although the potential perceived differences are difficult to describe. Extracting the underlying information from images and coupling this visual information with actual experimental outcomes in an unbiased way is challenging but could be a useful tool in accurately predicting outcomes. Frequently, day 10 flow cytometry data is used to characterize bioreactor-produced cardiomyocytes and determine if batches can move forward for downstream applications (11-13). The ability to non-destructively predict cardiac differentiation outcomes at an earlier time point would significantly reduce the time and expense

involved for hiPSC-CM and engineered cardiac tissue production by avoiding the continuation of the differentiation process for low-yield batches.

In recent years, machine learning (ML) techniques have been successfully used to represent expensive and complex systems with reduced-order models. These techniques map the information from process inputs to the outcomes (outputs). The input data can consist of different types, such as process parameters, categorical variables, and images. Different ML algorithms with various types of data have also been implemented in bioprocesses (e.g. (14, 15)). Increasing computational power and the progress in ML algorithms have enabled using the information in images for building predictive models. For example, CM contraction, measured based on single cardiomyocyte deformation dynamics, was quantified using image correlation analysis (16), and plant diseases were detected using neural networks trained based on images of the plants (17).

Convolutional neural networks (CNN) (18) are commonly used for image processing. For instance, Vardhana, Arunkumar (19) discuss the implementation of the CNN for biomedical image segmentation and disease detection by classification. Kusumoto and Yuasa (20) developed CNN models using phase-contrast images to identify morphology-based cell types for differentiated cells from induced pluripotent stem cells. Orita, Sawada (21) trained a CNN model using bright-field images of human-induced pluripotent stem cell-derived cardiomyocytes to classify cell culture quality into normal and abnormal groups. The automatic feature extraction capability of CNN makes them a robust toolset for image processing to generate and select features that are the most predictive of the output. CNN models can handle more than one type of data, making them an excellent candidate method for building models for biological systems with experimental and other types of data, such as images, present.

In our previous experimental studies, we established the ability to produce engineered cardiac tissue through hiPSC hydrogel encapsulation and direct differentiation within the 3D engineered tissue microenvironment (22-24), eliminating the need to pre-differentiate, dissociate, and re-assemble CMs. The cardiac differentiation in these hydrogel encapsulation platforms has been confirmed by displaying an expression of appropriate cardiac markers and genes, well-defined and aligned sarcomeres and T-tubules formation, and appropriate response of cardiac tissues to outside pacing and drug treatment (22-25). We then transitioned this approach to suspension-based cardiac differentiation by using a microfluidic system to form spheroidal hiPSC-laden engineered microtissues (25, 26). Using this system, multiple tunable parameters can be adjusted during the differentiation protocol timeline, phase-contrast images of the microspheroids (including microsphere and microrod (27)) are taken at regular time intervals and the CM content is evaluated on differentiation day 10 (dd10) using flow cytometry for all differentiation batches.

In this study, we investigated the ability to classify CM content on day 10 of hiPSC-laden microspheroid differentiation using easily acquired non-destructive images taken on days 3 and 5 (middle of the differentiation period, after small molecules addition) and tunable experimental parameters from days -3 (three days before the start of differentiation) through day 0 of differentiation (Figure 1). The percentage of cells that are cardiomyocytes on the specific differentiation day is defined as CM content. It is our hypothesis that ML can capture the information contained in the combination of the tunable experimental parameters from the early days of differentiation with the phase-contrast images of the microspheroids taken during differentiation (Second row in Figure 1 and Figure 2) to distinguish batches that will have sufficient CM content on day 10 from those that will have insufficient CM content. Based on our recent experience in hiPSC-laden microspheroid cardiac differentiation (25) and other studies

which have defined a range of target hiPSC-CM percentages for successful CM differentiation (65-92%) (28-31), output was labeled as *sufficient* for CM content from flow cytometry above 70% and *insufficient* for otherwise. Convolutional neural networks (CNNs) were implemented to predict the CM content class. The best classifier had an accuracy of 85%. A brief description of the experiments, data collection, and construction of the classification model are included in the Methods. The model training details are discussed in Computational Experiments, followed by the classification model statistics and their analysis in Results and Discussion. The takeaway message is summarized in Conclusions.

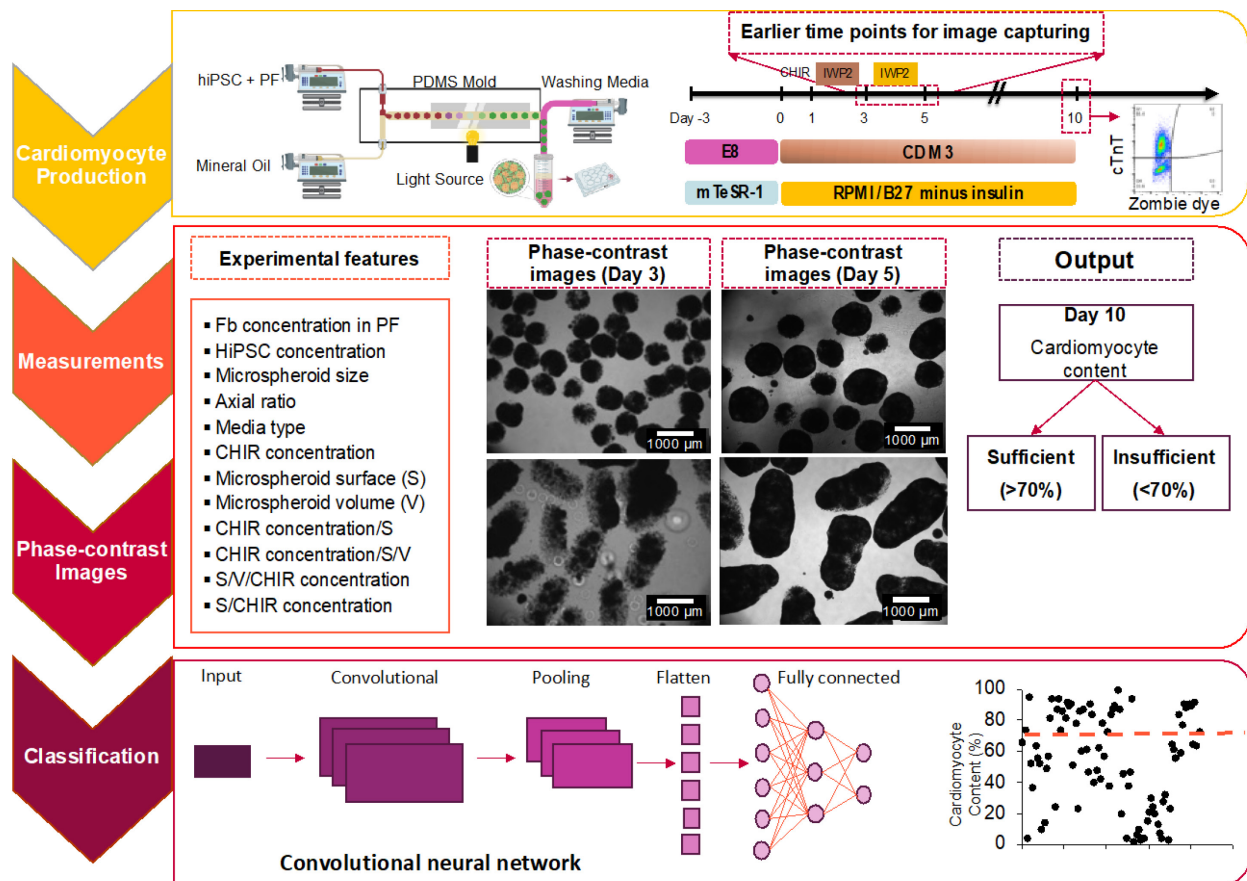


Figure 1. Schematic of the process to collect experimental data for building a classification model of CM content from cell encapsulation experiments and classification model workflow. Human induced pluripotent stem cells (hiPSCs) were resuspended in PF precursor solution, including the photo initiator at a concentration of 30, 40, 50, and 60 million cells mL⁻¹ of PF. The precursor solution and mineral oil were infused into the top and bottom inlets of the PDMS mold, respectively. With breaking the surface tension of the precursor solution, the microspheroids of hydrogel were created and crosslinked in the outlet by

using visible light. Microspheroids were collected at the end of the PDMS mold and removed from the oil phase, and cultured in two different stem cell media for expansion for 3 days. Cardiac differentiation was initiated on day 0 using two different differentiation protocols, as shown in the timeline. Experimental features that were measured and intentionally changed were fibrinogen (Fb) concentration in polyethylene glycol - fibrinogen (PF) and hiPSCs concentration in PF on day -3, microspheroid size and shape measured on day -2, and CHIR concentration on day 0. The phase-contrast images of microspheroids were taken on days 3 and 5 of differentiation. The output was CM content in each batch of encapsulation which was measured by flow cytometry data on day 10 of differentiation. Convolutional neural networks were the machine learning method used to build the CM content classification model.

Methods

HiPSC Culture, Encapsulation within a Hydrogel, Expansion, and Directed Differentiation

Two hiPSC lines, IMR-90 Clone 1 and the Un-Arc 16 Facs II cell line (32), which is genetically encoded with a voltage (ArcLight) fluorescent indicator, were used in this study. The hiPSCs were cultured in stem cell media, mTeSR-1 medium (Stem Cell Technologies) or E8 medium as described (25, 33), in T-25 flasks coated with Matrigel (Corning) or Geltrex (Gibco), respectively. Briefly, hiPSCs were detached using Accutase (Innovative Cell Technologies) for about two minutes in the incubator and resuspended in stem cell media supplemented with ROCK inhibitor (5-10 μ M, RI, Y 27632, Stem Cell Technologies).

PEG-fibrinogen (PF) as a biomaterial for encapsulation was synthesized as previously described (25, 34). To characterize the PF fibrinogen concentration, the Pierce BCA assay (Thermo Scientific) was used. By changing the fibrinogen amount during the synthesis, PF with a range of different fibrinogen concentrations can be produced (8-14 mg/mL). A novel microfluidic system was used as previously described (27) to prepare hiPSC-laden microspheroids. PF at a final protein concentration from 8.6 to 12.3 mg/mL in Phosphate Buffered Saline (PBS) was used in this study. To prepare the precursor solution, hiPSCs were resuspended in PF with different concentrations (30, 40, 50, and 60 million cells mL⁻¹ PF). Then, this precursor solution was injected from the top inlet and mineral oil was infused from the bottom inlet of the PDMS mold to break the surface tension of the precursor solution; using this approach, microspheroidal hydrogels were created (Figure 1). To manipulate the size and shape of these microspheroids, experimental

parameters, including junction size, precursor solution and mineral oil flow rates, and outlet channel diameter, were adjusted as previously described (27). To photo crosslink the liquid PF-cell mixture, a 2.7W light source (Prior Lumen 200) was employed. Stem cell media supplemented with ROCK inhibitor (5-10 μ M, RI, Y 27632, Stem Cell Technologies) was used for collecting and culturing the microspheroids for 24 h (day -3). Microspheroids were then cultured for an additional 48 h in mTeSR 1 or E8 medium with daily media changes (days -2 and -1).

Two differentiation protocols were used, as briefly described here. For the first differentiation protocol, cardiac differentiation was initiated (day 0) by activation and inhibition of the Wnt signaling based on a method previously established for use in stirred tank bioreactors (12). To reduce the differentiation cost and produce the clinically relevant quantities of hiPSC-CMs, a cost-effective and xenobiotic-free (xeno-free) media, chemically defined medium, 3 components (CDM3) (35) was used. CDM3 medium was supplemented with 5 μ M CHIR99021 (Stem Cell Technologies) on day 0 of differentiation, and the well plate was placed on a shaker. After 24 hours (day 1), the medium was replaced with CDM3 medium supplemented with 5 μ M IWP2 (Stem Cell Technologies) to inhibit Wnt signaling. After 48 hours, the medium was changed to RPMI/B27 containing insulin (Invitrogen), with subsequent medium changes every other day until day 7 and every 3 days thereafter. Cardiac tissue contraction was observed on day 10 (Supplementary Video 1). In the second differentiation protocol, cardiac differentiation (36) was initiated (day 0) by switching the medium from mTeSR-1 to RPMI/B27 minus insulin (RPMI/B27-I, Thermo Fisher) supplemented with CHIR (10-12 μ M, Stem Cell Technologies). After 24 h (day 1), the medium was removed, and fresh RPMI/B27-I (4 mL) was added. After an additional 48h (day 3), 2 mL of old media was combined with 2 mL of fresh RPMI/B27-I supplemented with IWP2 (5 μ M, Stem Cell Technologies). After another 48 h (day 5), media was replaced with

RPMI/B27-I (4 mL); on day 7, RPMI/B27-I was changed to RPMI/B27 medium (Thermo Fisher). RPMI/B27 medium was replaced every three to four days following differentiation.

Phase-contrast images were taken (Ti Eclipse, Nikon equipped with an Andor Luca S camera) on differentiation days 3 (dd3) and 5 (dd5) of the hiPSC-laden microspheroids, as shown in Figure 2.

Cardiomyocyte content of each batch of differentiated engineered cardiac tissue microspheroids was assessed on day 10 by flow cytometry as previously described (25). Following microspheroid dissociation, cells were washed with PBS, and unlabeled samples were collected. The cells were then incubated in Zombie dye (Biotium), fixed and permeabilized overnight with cold Foxp3 Fixation/Permeabilization (ThermoFisher) working solution at 4 °C. Cells were labeled with primary antibodies (cTnT 1:400, Thermo Scientific; MF20 1:200, AbCam; IgG Isotype 1:1000, ThermoFisher) for 1 hour at room temperature, washed with Permeabilization Buffer, and incubated in secondary antibody (1:300, Alexa Fluor 647, Invitrogen) for 45 minutes. Following washing, cells were analyzed with a CytoFLEX LX flow cytometer (Beckman Coulter). For visualization of produced CMs, differentiated microspheroids were dissociated and replated on Matrigel (Corning) coated coverslips for 2-4 days. Then the hiPSC-CMs were immunostained with alpha-sarcomeric actinin (Sigma Aldrich) as described previously (25). Images were taken by using a Nikon A1R laser-scanning confocal microscope. Representative images of day 30 hiPSC-CMs derived through CDM3 microspheroid differentiation and extracted morphological data using an unbiased algorithm, SarcOmere Texture Analysis (SOTA) (37) are shown in Supplementary Figure 1. To verify the relative functionality of CDM3 differentiated microspheroids from different batches, videos of microspheroid tissue contractions were taken using an Andor Luca S camera

attached to the microscope on day 10 and analyzed by using motion-tracking software, an open-source MATLAB code (38). Data are provided in Supplementary Figure 2 and Video 1.

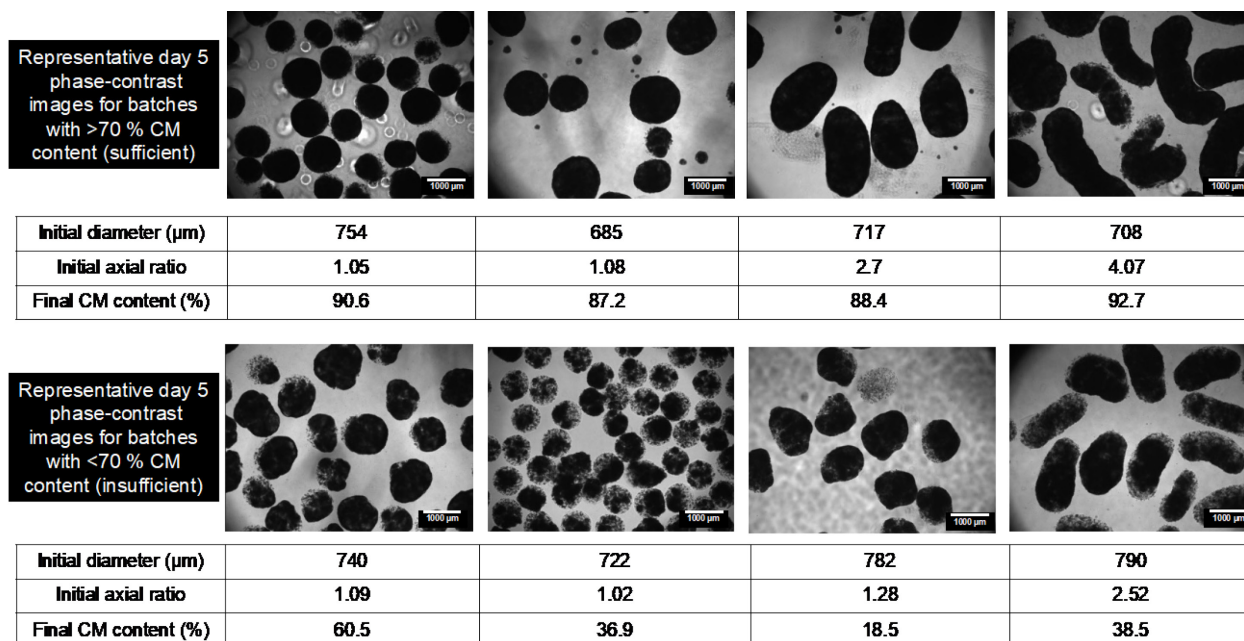


Figure 2. Representative differentiation day 5 phase-contrast images for four individual batches with CM content higher than 70% and four batches with CM content lower than 70%. The initial tunable size (diameter) and shape (axial ratio) of each batch have been provided in the table.

Data Used for Building the Classification Models

Two types of data were collected from the experiments during the differentiation timeline. Several experimental feature values were tuned and measured from day -3 to day 0 of the differentiation. These features included PF fibrinogen concentration, hiPSC concentration, microspheroid size (diameter) and axial ratio (ratio between axis), media type, and CHIR molecule concentration. These experimentally specified adjustable features were used to derive additional features. The derived features included microspheroid surface area (S), the ratio of CHIR concentration to microspheroid surface and its inverse (CHIR/S, S/CHIR), and the ratio of the CHIR concentration to microspheroid surface area to volume ratio (CHIR concentration/S/V) and its inverse (S/V/CHIR concentration).

The second data type consisted of phase-contrast images taken on days 3 and 5 of differentiation. These time points, days 3 and 5, were selected because they are early in the differentiation process. Furthermore, day 5 is the first time point after finishing incubating the samples with the small molecules. And at this time point, experienced experimenters felt that they could reliably guess whether or not the final CM content of the resulting engineered cardiac tissue microspheres would be “good” (sufficient) based on visualization of the differentiating hiPSC-laden engineered tissue microspheroids. For the purposes of this study, which aimed to predict the CM content classes of the cardiac tissues on day 10 of differentiation, the latter time points (such as day 7) were considered too late. Being able to predict the CM content at an earlier time point (days 3 or 5 versus day 7) enables earlier termination of differentiation of batches with insufficient CM content.

The output was the corresponding CM content of the resulting engineered cardiac tissue microspheroids measured using flow cytometry on day 10 of differentiation. The data included 280 images of 56 batches for day 3 and 303 images of 64 batches for day 5 of engineered tissue microsphere cardiac differentiation experiments, with their output labeled as *sufficient* and *insufficient* for CM content above and below 70%, respectively. The data was divided to test and training sets with a ratio of 20% and 80%, respectively, for both day 3 and day 5 points. The data was relatively balanced, with 53.5% *sufficient* and 46.5% *insufficient* CM content labels.

Classification Model Construction and Performance Metrics

Convolutional neural networks (CNN) (18) mimic the human brain’s visual perception mechanism (39). They are mainly used for image processing and can handle mixed data types (19, 40). Unlike most ML modeling approaches, where feature engineering and extraction are crucial

steps for building accurate models, in CNN models, the features are extracted automatically and used for predictions (39).

Accuracy (41), recall (42), precision (42), and Mathew's correlation coefficient (MCC) (43) were the metrics used for comparing the performance of the classifiers. All these metrics are calculated using the confusion matrix shown in Figure 3. The four cells in Figure 3 correspond to True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Our study associated positive and negative classes with *sufficient* and *insufficient* CM content classes.

		Predicted CM Content	
		Sufficient	Insufficient
Actual CM Content	Sufficient	TP	FN
	Insufficient	FP	TN

Figure 3. Confusion matrix for calculating the classification performance metrics based on the actual and predicted CM content classes. TP: True positive, TN: True negative, FN: False negative, FP: False positive.

Accuracy (41) calculates the ratio of the correct predictions of the classes (Eq. 1), and it is a value between 0 and 1, corresponding to wrong and perfect classification for all the points, respectively. Recall (42) presents the ratio of the actual positive class points predicted correctly, with a value between 0 and 1 (Eq. 2). A value of zero suggests that all the positive class points were incorrectly classified as negative, and a value of one corresponds to all correct class identification for the positive class points. Precision (42) indicates the proportion of the times the points predicted to be in the positive class corresponding to the actual positive class (Eq. 3). Finally, MCC (43) is a metric between -1 and 1, representing the correlation between the predicted

and actual classification for all the points (Eq. 4). The value of one for MCC means a perfect positive correlation between the predictions and original labels. The value of -1 shows a perfect inverse correlation suggesting that all the prediction classes were wrong for the data points. The value of 0 corresponds to the situation where the classification is no different than the random label assignment for each point.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (2)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (3)$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4)$$

Computational Experiments

Details of the CNN Classification Model Architecture, Training, and Comparison Studies

Two types of classification models based on the input data type were constructed to predict the CM content of engineered cardiac tissue microspheroids on differentiation day 10 and compared to each other. In the first model, only phase-contrast images were used as inputs to the CNN model (CNN_{image}). The second model employed mixed data types, i.e., experimental features and images (CNN_{mixed}). Monte-Carlo cross-validation (44) with 30 replications was used for tuning the parameters in CNN models with a validation set ratio of 20%. The comparison metrics were calculated for each model using the test data. Models were trained and evaluated using Keras (45) library in Python 3.7. The experimental features consisted of twelve features listed in Methods

section, and the images had a size of 496×658 pixels. The final architecture of both $\text{CNN}_{\text{image}}$ models of days 3 and 5 consisted of five convolutional layers with filter sizes of 16, 32, 64, 64, and 64, respectively. A pooling layer with (2, 2) stride was added right after the first convolutional layer. After flattening, two dense layers with 1000 and 500 fully connected neurons were employed. For $\text{CNN}_{\text{mixed}}$, parallel to the $\text{CNN}_{\text{image}}$ architecture, two dense layers with 16 fully connected neurons were implemented for the experimental features. The outputs from both parts were combined and fed to two dense layers of 256 and 128 fully connected neurons. The activation function for all the layers was Relu function, except for the output layer, where a sigmoid function was utilized. The trained models are available at the Cremaschi research group GitHub repository (https://github.com/SCremaschi-Research-Group/Image_processing).

For a comprehensive evaluation of the classification models, the two classifiers of $\text{CNN}_{\text{image}}$ and $\text{CNN}_{\text{mixed}}$ trained using either day 3 or day 5 images were compared to two models from our previous studies (46, 47). The first model was a Gaussian process classifier ($\text{GP}_{\text{Efeatures}}$) trained only using experimental features as inputs (47), and the second model used support vector machines ($\text{SVM}_{\text{image}}$) with images as inputs (46). Finally, the performance of the best classifier was tested on another cell line, IMR-90, to evaluate the model's capability to predict CM content outcomes for cell lines not included in the training set. The data set from the IMR-90 cell line included 26 data points from 7 batches.

Results and Discussion

Figure 4 shows the classification results for $\text{CNN}^{\text{d3}}_{\text{image}}$, $\text{CNN}^{\text{d3}}_{\text{mixed}}$, $\text{SVM}^{\text{d3}}_{\text{image}}$, $\text{CNN}^{\text{d5}}_{\text{image}}$, $\text{CNN}^{\text{d5}}_{\text{mixed}}$, $\text{SVM}^{\text{d5}}_{\text{image}}$ (46), and $\text{GP}_{\text{Efeatures}}$ (47) based on the four metrics of accuracy, recall, precision, and MCC calculated for the test points. The superscripts of each model represent the day of differentiation images used for training the models, where d3 and d5

correspond to days 3 and 5, respectively. According to Figure 4, $\text{CNN}^{\text{d5}}_{\text{mixed}}$ has the highest metric values, with 0.85 for accuracy, 0.82 for recall, 0.92 for precision, and 0.72 for MCC. The recall of $\text{CNN}^{\text{d5}}_{\text{mixed}}$ is the only metric lower than two other classifiers, $\text{SVM}^{\text{d3}}_{\text{image}}$ and $\text{SVM}^{\text{d5}}_{\text{image}}$, with a recall of 0.93 and 0.92, respectively. Although $\text{SVM}_{\text{image}}$ had a higher recall using both day 3 and 5 images, all other metric values were lower than those of $\text{CNN}^{\text{d5}}_{\text{mixed}}$, with a difference of around 0.2. These metric values indicate that both $\text{SVM}_{\text{image}}$ models correctly label more of the true positives; however, they do so at the expense of yielding more false negatives and positives, given their lower accuracy and precision values. The classifier with the second-highest metric values is $\text{SVM}^{\text{d5}}_{\text{image}}$, followed by $\text{SVM}^{\text{d3}}_{\text{image}}$, $\text{CNN}^{\text{d3}}_{\text{mixed}}$, $\text{GP}^{\text{Efeatures}}$, $\text{CNN}^{\text{d5}}_{\text{image}}$, and finally $\text{CNN}^{\text{d3}}_{\text{image}}$. The accuracies for different classifiers range between 0.85 to 0.54, indicating that all models have reasonably identified the correct classes; however, there is a considerable change in accuracies of the $\text{CNN}_{\text{mixed}}$ and $\text{CNN}_{\text{image}}$ for both models trained using either day 3 or day 5 images. The MCCs of the classifiers are between 0.1 and 0.72, and the model with the lowest MCC employs day 3 images only.

The best model, $\text{CNN}^{\text{d5}}_{\text{mixed}}$, was used to evaluate if the CM content on day 10 for the IMR-90 cell line could be classified as sufficient or insufficient. For the IMR-90 cell line data, the accuracy was 0.46, recall and precision were 0.39, and MCC was -0.08, revealing that $\text{CNN}^{\text{d5}}_{\text{mixed}}$ cannot accurately predict sufficient or insufficient CM content classes for this cell line. These results suggest that the models trained using data from only one cell line could not be directly used for predictions for a different cell line CM content. The ML models, such as the classifier models trained in this study, are data-driven, meaning they only perform accurately within the ranges of the training data. To train models that can predict CM content for other cell lines or multiple cell lines, we posit that the training data should include the necessary information, such as images,

experimental features, and CM content labels for multiple lines. Future work is needed to assess the performance of classification models trained using multiple cell line data.

The results agree with our hypothesis that the information contained in the images of differentiating engineered tissue microspheroids combined with tunable experimental features can be extracted to build a classifier model to distinguish experimental runs with *sufficient* and *insufficient* CM contents on day 10 of the differentiation. The inclusion of features that non-destructively capture the information about the progression of stem cell growth and differentiation within the engineered tissue microspheres, represented by the images in our study, improves the prediction of the CM content class in comparison to the $GP_{Efeatures}$, which only utilizes tunable experimental features from the early days of the differentiation process (day -3 to 0). On the other hand, the lower metric values of the CNN_{image} and SVM_{image} compared to CNN_{mixed} reveal the critical impact experimental features play in the accurate prediction of the output class. The difference is more significant between CNN_{image}^{d5} and CNN_{mixed}^{d5} , whose accuracy and precision values are 0.22 apart, and recall and MCC values vary by 0.24 and 0.42, respectively. The accuracy, recall, precision, and MCC of CNN_{image}^{d3} and CNN_{mixed}^{d3} are also apart by 0.18, 0.3, 0.4, and 0.37, respectively. This result may be partially due to the limited number of data available to train the CNN classifiers because CNNs, a deep learning method, require a large number of data for accurate predictions. However, the improvement in the CNN classifier predictions with the addition of tunable experimental features indicates that the limited data is not the sole reason for the lower metric values yielded by CNN_{image} .

The information contained in the images, even though this cannot be directly linked to the ongoing biological processes within the engineered tissues, accompanied by the right set of experimental features, is essential for building accurate classification models for bio-processes like

the current one. In the future, adding images from earlier time points, including the pre-differentiation expansion phase, and other features capturing information regarding the differentiation process could enhance the prediction results. Features like dissolved oxygen, pH, media composition at different time points, and their changes through the process can provide high-quality information for building predictive models, similar to our previous study on the production of CMs in a bioreactor (11). With a more extensive feature set, we will take another step in constructing predictive models and consider building regression models instead of classifiers for estimating CM content on day 10 of differentiation. It is also important to note that expert knowledge should be incorporated when building predictive ML models for the hiPSC differentiation process to train accurate models. The expert knowledge informs the construction of the features that should be considered for training the models via identifying the relevant measurements containing the pertinent information and advising in forming additional engineered features generated by modifying the measurements. Cardiomyocyte functionality and maturation in the batches with *sufficient* and *insufficient* CM content were not compared here and could be different. We plan to investigate employing ML to predict the optimal differentiation conditions for enhancing hiPSC-CM functionality and maturation using related outcomes, such as contraction velocity; cell morphology, sarcomere organization, and length by using an unbiased algorithm, SarcOmere Texture Analysis (SOTA) (37); mitochondrial location and morphology, etc., in a future study (Supplementary Figures 1 and 2). Our study demonstrated that training ML models to predict differentiation outcomes using correct features representing the process is a promising field.

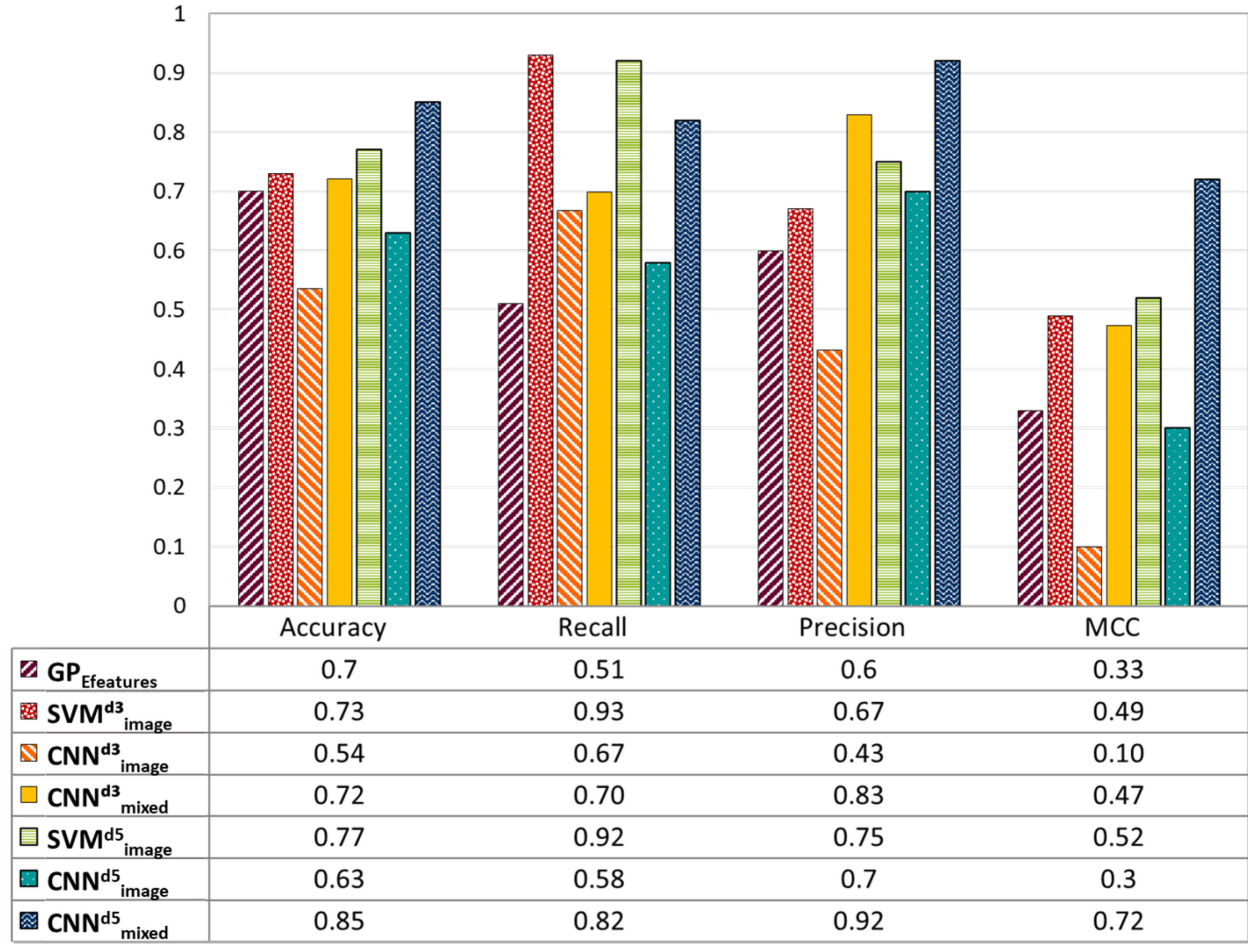


Figure 4. Comparison of performance metrics, accuracy, recall, precision, and MCC, for predicting CM content on day 10 of hiPSC differentiation using seven classification models. **GP_{Efeatures}**: Only experimental features were used as input for a Gaussian process classifier. **SVM^{d3}_{image}** and **SVM^{d5}_{image}**: Phase-contrast images taken on day 3 and day 5 of differentiation, respectively, were used to train support vector machine classifiers. **CNN^{d3}_{image}** and **CNN^{d5}_{image}**: Days 3 and 5 images were used to train convolutional neural network classifiers. **CNN^{d3}_{mixed}** and **CNN^{d5}_{mixed}**: Both experimental features and images were used as the input features of convolutional neural network classifiers, using images taken on either day 3 or day 5 of differentiation.

Conclusion

In this study, classification models for predicting the CM content were constructed using convolutional neural networks, and these classifiers were compared to two previously developed models. The first set of models (CNN_{image} and SVM_{image}) used phase-contrast images of the differentiating hiPSC-laden engineered tissue microspheroids on days 3 and 5 of differentiation. The second model type (CNN_{mixed}) used a combination of the images and experimental features as

inputs. The results demonstrated that CNN^{d5}_{mixed} had the highest accuracy, recall, precision, and MCC of 0.85, 0.82, 0.92, and 0.72, respectively. Its performance in predicting the CM content classes of the resulting engineered cardiac tissue microspheroids on day 10 of differentiation was significantly better than the other six classifiers.

Disclosure Statement

None of the authors of this article has any competing interests to declare; therefore, no competing financial interests exist for any of the authors.

Funding Information

The funding provided by the National Science Foundation (NSF) grants no 1743445 and 2135059 is acknowledged. Ferdous Finklea had an NSF Graduate Research Fellowship.

References

1. Hartmann J, Wong M, Gallo E, Gilmour D. An image-based data-driven analysis of cellular architecture in a developing tissue. *eLife*. 2020;9:e55913.
2. Komura D, Ishikawa S. Machine Learning Methods for Histopathological Image Analysis. *Comput Struct Biotechnol J*. 2018;16:34-42.
3. Conev A, Litsa EE, Perez MR, Diba M, Mikos AG, Kavraki LE. Machine Learning-Guided Three-Dimensional Printing of Tissue Engineering Scaffolds. *Tissue Engineering Part A*. 2020;26(23-24):1359-68.
4. Kim J, McKee JA, Fontenot JJ, Jung JP. Engineering Tissue Fabrication With Machine Intelligence: Generating a Blueprint for Regeneration. *Frontiers in Bioengineering and Biotechnology*. 2020;7.
5. [Available from: <https://www.cdc.gov/heartdisease/facts.htm>.
6. Kempf H, Andree B, Zweigerdt R. Large-scale production of human pluripotent stem cell derived cardiomyocytes. *Advanced Drug Delivery Reviews*. 2016;96:18-30.
7. Li J, Hua Y, Miyagawa S, Zhang J, Li L, Liu L, et al. hiPSC-derived cardiac tissue for disease modeling and drug discovery. *International journal of molecular sciences*. 2020;21(23):8893.
8. Tani H, Tohyama S. Human Engineered Heart Tissue Models for Disease Modeling and Drug Discovery. *Frontiers in Cell and Developmental Biology*. 2022;10.
9. Williams B, Löbel W, Finklea F, Halloin C, Ritzenhoff K, Manstein F, et al. Prediction of human induced pluripotent stem cell cardiac differentiation outcome by multifactorial process modeling. *Frontiers in bioengineering and biotechnology*. 2020:851.

10. Gaspari E, Franke A, Robles-Diaz D, Zweigerdt R, Roeder I, Zerjatke T, et al. Paracrine mechanisms in early differentiation of human pluripotent stem cells: insights from a mathematical model. *Stem Cell Research*. 2018;32:1-7.
11. Williams B, Löbel W, Finklea F, Halloin C, Ritzenhoff K, Manstein F, et al. Prediction of Human Induced Pluripotent Stem Cell Cardiac Differentiation Outcome by Multifactorial Process Modeling. *Frontiers in Bioengineering and Biotechnology*. 2020;8(851).
12. Halloin C, Schwanke K, Löbel W, Franke A, Szepes M, Biswanath S, et al. Continuous WNT control enables advanced hPSC cardiac processing and prognostic surface marker identification in chemically defined suspension culture. *Stem cell reports*. 2019;13(2):366-79.
13. Kahn-Krell A, Pretorius D, Ou J, Fast VG, Litovsky S, Berry J, et al. Bioreactor suspension culture: Differentiation and production of cardiomyocyte spheroids from human induced pluripotent stem cells. *Frontiers in bioengineering and biotechnology*. 2021;9:674260.
14. Wu SG, Wang Y, Jiang W, Oyetunde T, Yao R, Zhang X, et al. Rapid prediction of bacterial heterotrophic fluxomics using machine learning and constraint programming. *PLoS computational biology*. 2016;12(4):e1004838.
15. di Sciascio F, Amicarelli AN. Biomass estimation in batch biotechnological processes by Bayesian Gaussian process regression. *Computers & Chemical Engineering*. 2008;32(12):3264-73.
16. Kamgoué A, Ohayon J, Usson Y, Riou L, Tracqui P. Quantification of cardiomyocyte contraction based on image correlation analysis. *Cytometry Part A: The Journal of the International Society for Advancement of Cytometry*. 2009;75(4):298-308.
17. Vishnoi VK, Kumar K, Kumar B. Plant disease detection using computational intelligence and image processing. *Journal of Plant Diseases and Protection*. 2021;128(1):19-53.
18. LeCun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W, et al. Handwritten digit recognition with a back-propagation network. *Advances in neural information processing systems*. 1989;2.
19. Vardhana M, Arunkumar N, Lasrado S, Abdulhay E, Ramirez-Gonzalez G. Convolutional neural network for bio-medical image segmentation with hardware acceleration. *Cognitive Systems Research*. 2018;50:10-4.
20. Kusumoto D, Yuasa S. The application of convolutional neural network to stem cell biology. *Inflammation and regeneration*. 2019;39(1):1-7.
21. Orita K, Sawada K, Koyama R, Ikegaya Y. Deep learning-based quality control of cultured human-induced pluripotent stem cell-derived cardiomyocytes. *Journal of pharmacological sciences*. 2019;140(4):313-6.
22. Kerscher P, Kaczmarek JA, Head SE, Ellis ME, Seeto WJ, Kim J, et al. Direct Production of Human Cardiac Tissues by Pluripotent Stem Cell Encapsulation in Gelatin Methacryloyl. *ACS Biomaterials Science & Engineering*. 2016.
23. Kerscher P, Turnbull IC, Hodge AJ, Kim J, Seliktar D, Easley CJ, et al. Direct hydrogel encapsulation of pluripotent stem cells enables ontomimetic differentiation and growth of engineered human heart tissues. *Biomaterials*. 2016;83:383-95.
24. Chang S, Finklea F, Williams B, Hammons H, Hodge A, Scott S, et al. Emulsion-based encapsulation of pluripotent stem cells in hydrogel microspheres for cardiac differentiation. *Biotechnology Progress*. 2020;36(4):e2986.
25. Finklea FB, Tian Y, Kerscher P, Seeto WJ, Ellis ME, Lipke EA. Engineered cardiac tissue microsphere production through direct differentiation of hydrogel-encapsulated human pluripotent stem cells. *Biomaterials*. 2021;274:120818.

26. Seeto WJ, Tian Y, Pradhan S, Kerscher P, Lipke EA. Rapid Production of Cell-Laden Microspheres Using a Flexible Microfluidic Encapsulation Platform. *Small*. 2019;15(47):1902058.
27. Tian Y, Lipke EA. Microfluidic Production of Cell-Laden Microspheroidal Hydrogels with Different Geometric Shapes. *ACS Biomaterials Science & Engineering*. 2020;6(11):6435-44.
28. Dahlmann J, Kensah G, Kempf H, Skvorc D, Gawol A, Elliott DA, et al. The use of agarose microwells for scalable embryoid body formation and cardiac differentiation of human and murine pluripotent stem cells. *Biomaterials*. 2013;34(10):2463-71.
29. Sung T-C, Li H-F, Higuchi A, Kumar SS, Ling Q-D, Wu Y-W, et al. Effect of cell culture biomaterials for completely xeno-free generation of human induced pluripotent stem cells. *Biomaterials*. 2020;230:119638.
30. Branco MA, Cotovio JP, Rodrigues CA, Vaz SH, Fernandes TG, Moreira LM, et al. Transcriptomic analysis of 3D cardiac differentiation of human induced pluripotent stem cells reveals faster cardiomyocyte maturation compared to 2D culture. *Scientific reports*. 2019;9(1):1-13.
31. Fonoudi H, Ansari H, Abbasalizadeh S, Larijani MR, Kiani S, Hashemizadeh S, et al. A universal and robust integrated platform for the scalable production of human cardiomyocytes from pluripotent stem cells. *Stem cells translational medicine*. 2015;4(12):1482-94.
32. Shinnawi R, Huber I, Maizels L, Shaheen N, Gepstein A, Arbel G, et al. Monitoring human-induced pluripotent stem cell-derived cardiomyocytes with genetically encoded calcium and voltage fluorescent reporters. *Stem cell reports*. 2015;5(4):582-96.
33. Kempf H, Kropp C, Olmer R, Martin U, Zweigerdt R. Cardiac differentiation of human pluripotent stem cells in scalable suspension culture. *Nature protocols*. 2015;10(9):1345.
34. Almany L, Seliktar D. Biosynthetic hydrogel scaffolds made from fibrinogen and polyethylene glycol for 3D cell cultures. *Biomaterials*. 2005;26(15):2467-77.
35. Burridge PW, Matsa E, Shukla P, Lin ZC, Churko JM, Ebert AD, et al. Chemically defined generation of human cardiomyocytes. *Nature methods*. 2014;11(8):855-60.
36. Bonow RO, Mann DL, Zipes DP, Libby P. Braunwald's heart disease e-book: A textbook of cardiovascular medicine: Elsevier Health Sciences; 2011.
37. Sutcliffe MD, Tan PM, Fernandez-Perez A, Nam Y-J, Munshi NV, Saucerman JJ. High content analysis identifies unique morphological features of reprogrammed cardiomyocytes. *Scientific reports*. 2018;8(1):1-11.
38. Huebsch N, Loskill P, Mandegar MA, Marks NC, Sheehan AS, Ma Z, et al. Automated video-based analysis of contractility and calcium flux in human-induced pluripotent stem cell-derived cardiomyocytes cultured over different spatial scales. *Tissue Engineering Part C: Methods*. 2015;21(5):467-79.
39. LeCun Y, Bengio Y, Hinton G. Deep learning. *nature*. 2015;521(7553):436-44.
40. Yuan Z, Jiang Y, Li J, Huang H. Hybrid-DNNs: Hybrid deep neural networks for mixed inputs. *arXiv preprint arXiv:200508419*. 2020.
41. Guyon I, Elisseeff A. An introduction to variable and feature selection. *Journal of machine learning research*. 2003;3(Mar):1157-82.
42. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*. 2009;45(4):427-37.
43. Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*. 1975;405(2):442-51.

44. Burman P. A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. *Biometrika*. 1989;76(3):503-14.
45. Gulli A, Pal S. Deep learning with Keras: Packt Publishing Ltd; 2017.
46. Mohammadi S, Finklea F, Hashemi ML, Elizabeth , Cremaschi S, editors. Image classification of experimental yields for cardiomyocyte cells differentiated from human induced pluripotent stem cells. International Symposium on Process Systems Engineering; 2022; Kyoto, Japan. Computer Aided Chemical Engineering: Elsevier.
47. Mohammadi S, Cremaschi S, editors. Classification of cardiomyocyte content differentiated from human induced pluripotent stem cells2021: AIChE.