Biochemical and Biophysical Research Communications 526 (2020) 751–755

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Machine-learning-based quality control of contractility of cultured human-induced pluripotent stem-cell-derived cardiomyocytes

Ken Orita [a], Kohei Sawada [a], Nobuyoshi Matsumoto [a], Yuji Ikegaya [a, b, *]

[a] Graduate School of Pharmaceutical Sciences, The University of Tokyo, Tokyo, 113-0033, Japan
[b] Center for Information and Neural Networks, National Institute of Information and Communications Technology, Suita City, Osaka, 565-0871, Japan

ARTICLE INFO

Article history:
Received 1 March 2020
Accepted 25 March 2020
Available online 4 April 2020

Keywords:
Machine learning
SVM
UMAP
Heart
iPSC

ABSTRACT

The precise and early assessment of cardiotoxicity is fundamental to bring forward novel drug candidates to the pharmaceutical market and to avoid their withdrawal from the market. Recent preclinical studies have attempted to use human-induced pluripotent stem-cell-derived cardiomyocytes (hiPSC-CMs) to predict clinical cardiotoxicity, but the heterogeneity and inconsistency in the functional qualities of the spontaneous contractility of hiPSC-CMs across cell culture wells and product lots still matter. To rapidly assess the functional qualities of hiPSC-CMs without histological labeling, we optically detected the contractility of confluently cultured hiPSC-CMs using bright-field microscopy. Using a method that consisted of data preprocessing, data augmentation, dimensionality reduction, and supervised learning, we succeeded in precisely discriminating between functionally normal and abnormal contractions of hiPSC-CMs.

1. Introduction

Off-target cardiotoxicity often causes discontinuation of drug development or drug withdrawal from the pharmaceutical market [1]. One of the factors that makes it difficult to predict cardiotoxicity in preclinical studies is species differences in cardiomyocytes. To overcome this problem of species differences, recent studies have attempted to use human-induced pluripotent stem-cell-derived cardiomyocytes (hiPSC-CMs) to assess their contractility, action potentials, and viability [2–4], and classical animal models for cardiac safety assays are expected to be substituted with hiPSC-CMs [5,6]. However, at present, hiPSC-CMs are experimentally vulnerable, and their morphology and physiology vary depending on the product lots and experimental conditions, even when the experimental protocols and experimenters are identical. We have previously established a convolutional-neural-network-based method for the automated assessment of the cell morphology of cultured hiPSC-CMs [7]. In contrast, there are few methods that can assess the quality of the contractile functions of hiPSC-CMs [8]. In this study, we optically quantify the contractility of hiPSC-CMs using bright-field microscopic videos, instead of measuring isometric or isotonic contractions, and extract the contraction waves directly from time-lapse video images. Using dimensionality reduction, data augmentation, data preprocessing, and binary classification [9,10], we divide the contraction waves into 'normal' contraction (experimentally useable) and 'abnormal' contraction (experimentally unusable) waves. More specifically, we utilize i) a sliding window method for data augmentation, ii) fast Fourier transform for data preprocessing, and iii) uniform manifold approximation and projection (UMAP) for nonlinear dimensionality reduction. We then train a support vector machine (SVM) to evaluate the qualities of individual contraction waves [11].

2. Materials and methods

2.1. hiPSC-CM culture

iCell Cardiomyocytes[2] (Cellular Dynamics International, Inc.; Madison, WI, USA), commercially available hiPSC-CMs, were cultured as previously reported [7]. In short, hiPSC-CMs were seeded at 40,000 cells/well in 96-well plates and formed a 2D monolayer structure. After 5–7 d of incubation, they were treated

Abbreviations: hiPSC-CM, human-induced pluripotent stem-cell-derived cardiomyocyte; UMAP, uniform manifold approximation and projection; SVM, support vector machine; FFT, fast Fourier transform.

* Corresponding author. Laboratory of Chemical Pharmacology, Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan.
E-mail address: yuji@ikegaya.jp (Y. Ikegaya).

https://doi.org/10.1016/j.bbrc.2020.03.141
0006-291X/

with drugs to increase the variability in the qualities of the hiPSC-CM cultures.

2.2. Data acquisition

Bright-field videos of cultured hiPSC-CMs (1,280 × 1,080 pixels, 16-bit intensity, 20 fps, 30 s) were obtained before and 1–3 d after the drug treatments using a 20 × objective and CQ1 confocal quantitative image cytometer (Yokogawa Electric, Tokyo, Japan) and saved as a TIFF stack.

A total of 624 videos were obtained from four independent experiments. All videos were carefully inspected by four well-trained experimenters [7]. Of the 624 videos, 68 were easily found to exhibit extremely bad conditions and thus were excluded from the following analyses. The remaining 556 videos were categorized as either 'normal' ($n = 366$ videos) or 'abnormal' ($n = 190$ videos) based on the contractile qualities of the cultured hiPSC-CMs. This categorization was consistent across the well-trained experimenters, although it may be difficult for beginners to achieve the same result.

2.3. Data analysis

Data analyses were performed using Python 3, and the summarized data were reported as the means ± standard error of the means (SEMs) unless otherwise specified. $P < 0.05$ was considered statistically significant.

2.3.1. Quantification of contraction properties of hiPSC-CMs

The TIFF stacks of 600 frames were imported to Fiji (a distribution of ImageJ) [12]. To investigate the longitudinal changes in the background pixel intensity, we used the 200-frame simple moving average (SMA) around the frame of interest [13], which was calculated using the average intensity projection mode. The periodic contraction of hiPSC-CMs was quantified by calculating the absolute value changes between the t-th frame of interest (I_t) and the 200-frame SMA of the t-th frame (SMA_t). For every frame, each pixel in SMA_t was subtracted from the corresponding pixel in I_t, and the difference was used as an absolute value change. All the absolute value changes were averaged for every frame. The changes in the pixel intensities (hereafter called 'contraction waves') were obtained for the middle 400 frames out of a total of 600 frames. In each video, the amplitudes of the contraction wave were normalized between 0 and 1.

2.3.2. Dataset split (10-fold cross-validation)

The dataset ($n = 556$ contraction waves) was randomly split into 10 subsets ($S_1, S_2, \ldots, S_{10}$). Then one subset (*e.g.*, S_1) and the rest of the subsets (*e.g.*, $S_2, S_3, \ldots, S_{10}$) were assigned as a testing subset and a training subset, respectively. This procedure was repeated for all S_i (Fig. 1A) [14]. Each subset randomly contained 'normal' videos and 'abnormal' videos with a nearly constant normal/abnormal ratio of approximately 1.9.

2.3.3. Data augmentation (sliding window method)

Each contraction wave ($n = 556$ waves in total) was snipped off by the sliding window method [15], where the window size and the stride were 200 frames and 40 frames, respectively. Then, the total number of contraction waves was increased 6-fold.

2.3.4. Data preprocessing (fast Fourier transform)

We used the fast Fourier transform (FFT) algorithm (Python-implemented) to convert contraction waves ($n = 556$ waves before data augmentation or $n = 3{,}336$ waves after data augmentation) into the frequency domain data [16].

2.3.5. Dimensionality reduction (uniform manifold approximation and projection)

Uniform manifold approximation and projection (UMAP) is a nonlinear dimensionality reduction algorithm, a machine learning method based on a combination of Riemannian geometry and algebraic topology [17,18]. A total of 400 (before the data augmentation) or 200 (after the data augmentation) dimensions of the contraction waves (before or after the FFT) were reduced to two dimensions (Python-implemented with the following default parameters: *n_neighbors* = 15, *min_dist* = 0.1, and *metric* = 'euclidean').

2.3.6. Prediction of contractile properties of hiPSC-CMs

We used the SVM algorithm for classification [11]. It took two dimensions in the UMAP space as inputs and determined the decision boundary across two classes (*i.e.*, normal or abnormal) in the UMAP space. Because the shape of the data structure in the UMAP space was nonlinear (Fig. 2B), we used the radial basis function kernel. To correct the class imbalance problem, we set the class weight as abnormal/normal = 1.9. The SVM parameters regarding the shape of the decision boundary (*i.e.*, gamma and C) were optimized in the training subsets using the grid search algorithm with 5-fold cross-validation so that the classification accuracy could be maximized.

To evaluate the post-training SVM model, we used four indices: accuracy, precision, recall, and F1. These four indices are defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN},$$

$$\text{Precision} = \frac{TP}{TP + FP},$$

$$\text{Recall} = \frac{TP}{TP + FN},$$

$$\text{F1} = \frac{2Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FP + FN},$$

where TP, FP, FN, and TN represent the true positive, false positive, false negative, and true negative ratios of classification, respectively.

3. Results

A total of 556 videos of confluently cultured hiPSC-CMs were taken using bright-field microscopy. Four well-experienced experts watched 556 original videos and classified them into 366 normal and 190 abnormal cell cultures (Fig. 1A). The beat rates were significantly higher in the normal hiPSC-CMs than in the abnormal hiPSC-CMs (Fig. 1C; 0.44 ± 0.04 Hz (normal) and 0.31 ± 0.09 Hz (abnormal), $P = 3.3 \times 10^{-87}$, $t = 24$, $\nu \approx 227$, $n = 366$ (normal) and 190 (abnormal) waves, Welch's t-test). Each video was transformed into a contraction wave based on time changes in the pixel intensities in the video images (Fig. 1B and C). The periodic repetition in each contraction wave was reminiscent of typical *in vivo* ECGs [19] and pulsatile blood flow [20].

UMAP was used to reduce a total of 400 dimensions (*i.e.*, 400 video frames) of the contraction waves to two dimensions (Fig. 2A). Since UMAP preserves both global and local data structures, we expected that the UMAP representations of the contraction waves would reflect the fundamental properties in the contractility of hiPSC-CMs (Fig. 2A). Indeed, the 556 data points of the normal or

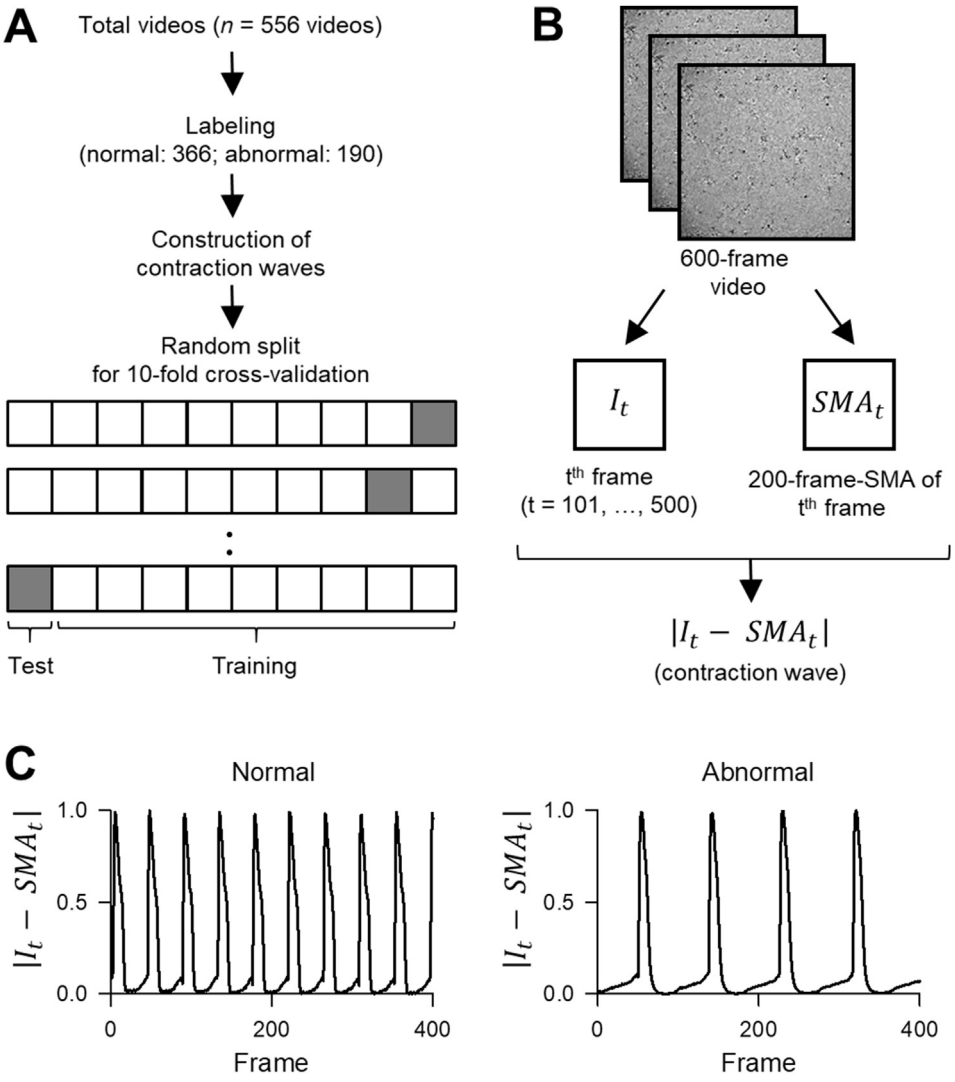

Fig. 1. Grouping of contraction waves obtained from bright-field videos of hiPSC-CMs. *A*, Bright-field videos of hiPSC-CMs were manually labeled as either 'normal' or 'abnormal' videos. The video datasets were randomly split into 10 consecutive folds. Each fold was used once as a testing set, while the remaining 9 folds were used as training sets. *B*, Construction of contraction waves from bright-field videos. First, the mean pixel intensity of a 200-frame simple moving average image (SMA_t) around the *t*-th frame was obtained as the baseline intensity. SMA_t was then subtracted from the pixel intensity of the *t*-th image (I_t). The consecutive absolute values of (I_t - SMA_t) were defined as a contraction wave (see Materials and methods). *C*, Representative contraction waves of normal (*left*) and abnormal (*right*) cultures of hiPSC-CMs.

abnormal videos appeared to be relatively separated in the UMAP space (Fig. 2B). We thus sought to classify the data points as either 'normal' or 'abnormal' using the SVM. For a 10-fold cross-validation, we split the 556 contraction waves randomly into 10 distinct subsets (Fig. 1A). To assess the performance of the post-training SVM classifier, we adopted four indices: accuracy, precision, recall, and F1 score. We found that these parameters were 0.85 ± 0.04, 0.88 ± 0.11, 0.66 ± 0.13, and 0.75 ± 0.09, respectively (Fig. 3; means ± SEMs of 10 trained models).

To further improve the performance of the SVM, we employed two strategies: data preprocessing and data augmentation (Fig. 2A). For data preprocessing, we applied FFT to the contraction waves to decompose the time changes in the beat patterns of hiPSC-CMs into the frequency component [16]. When we classified these FFT datasets using the SVM, the accuracy, precision, recall, and F1 score were 0.89 ± 0.05, 0.92 ± 0.09, 0.76 ± 0.09, and 0.83 ± 0.07, respectively (Fig. 3). For data augmentation [21], we used a sliding window method to increase the total number of datasets from 556 to 3,336 waves. The accuracy, precision, recall, and F1 score were 0.89 ± 0.02, 0.92 ± 0.04, 0.75 ± 0.05, and 0.82 ± 0.04, respectively (Fig. 3). Finally, we combined the FFT and the sliding window method, that is, the FFT was applied to the 3,336 augmented waves. The accuracy, precision, recall, and F1 score were 0.89 ± 0.02, 0.93 ± 0.04, 0.73 ± 0.05, and 0.82 ± 0.04, respectively (Fig. 3). The accuracy, recall, and F1 score were significantly higher when the FFT and/or sliding window method was used (Fig. 3).

4. Discussion

In this study, we established a machine-learning-based method for the semiautomatic quality assessment of cultured hiPSC-CMs in terms of their contractile function. The onset times of contractions were different among videos; thus, the data structures in the contraction waves are complex and cannot be assessed by absolute times. To analyze these complex data, previous studies measured linear parameters, such as the beat rates, periodicity, and amplitude modulation. However, these so-called "biological" parameters are arbitrarily selected by humans and may not necessarily capture all

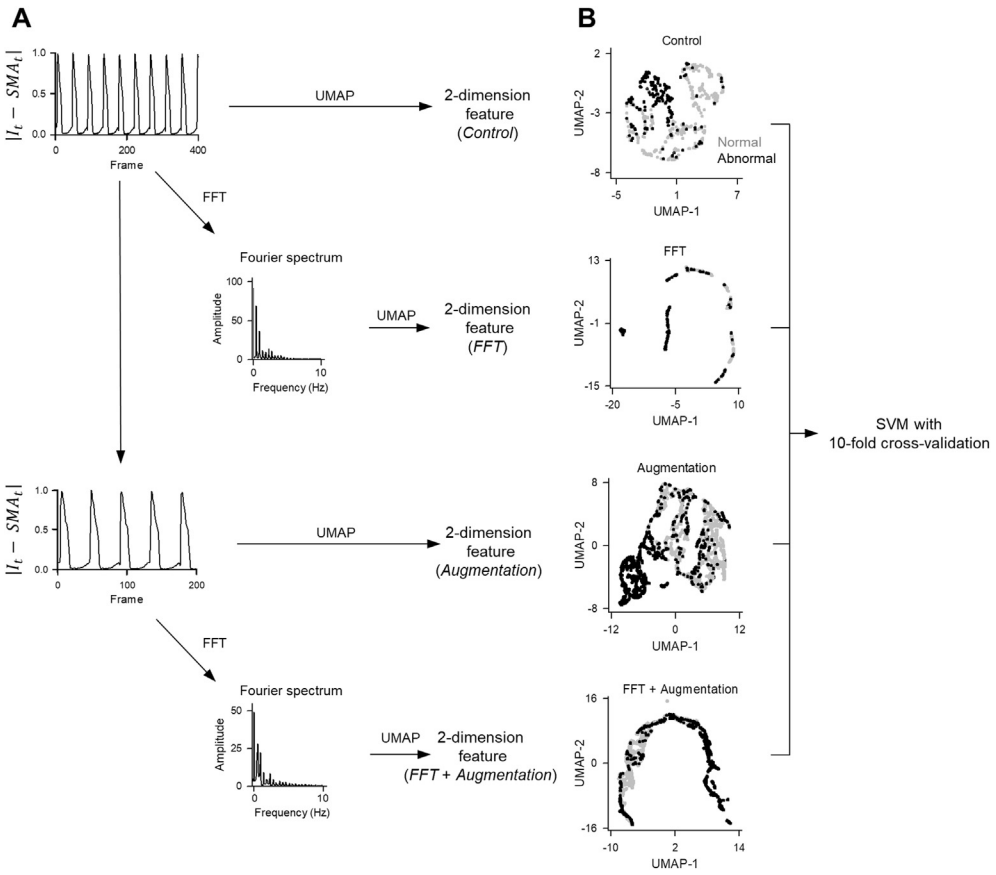

Fig. 2. Visualization of two-dimensional features of contraction waves. *A*, Four strategies for the two-dimensional characterization of the contraction waves. First, UMAP was applied directly to 400-frame contraction waves (*Control*). Second, the FFT was applied to the contraction waves, and UMAP was applied to the FFT spectra (*FFT*). Third, the contraction waves were shortened to increase the number of data points six-fold, and UMAP was applied to the augmented datasets (*Augmentation*). Fourth, the FFT was applied to the augmented datasets, and UMAP was applied to the FFT spectra (*FFT + Augmentation*) (see Materials and methods). Then, the SVM was trained using the training datasets, and its classification performance was evaluated using the testing datasets described in Fig. 1A. Training and evaluation were repeated ten times. *B*, Visualization of the two-dimensional features in the UMAP spaces constructed in *A*. Normal and abnormal hiPSC-CM cultures are indicated by the *gray* and *black* dots, respectively.

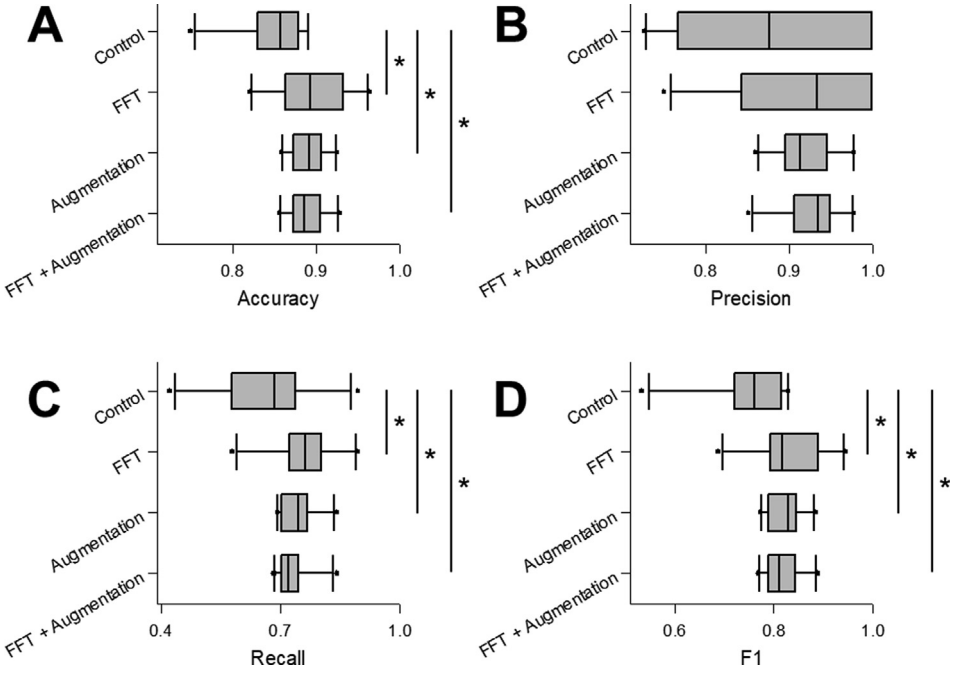

Fig. 3. Comparison of SVM classification performances. *A*, Comparisons of the accuracy in a 10-fold cross-validation among four strategies. *B–D*, The same as *A* but for the precision, recall, and F1 score, respectively. *$P < 0.05$, Steel test.

significant features hidden in the datasets. In the present study, we avoided extracting any biological parameters from the datasets, and instead, we reduced the dimensions of the entire datasets *per se*. Our method was based on a serial combination of UMAP, a nonlinear dimensionality reduction algorithm, and a nonlinear SVM. The classification performance was significantly high for the original wave data compared with the surrogate data (*data not shown*), but the classification scores were further improved by data preprocessing and data augmentation. The FFT serves to remove the phase variance of the contraction waves; note that the FFT converts a time-series signal from the time domain to the frequency domain and extracts the temporal features [19,22]. Indeed, FFT-converted datasets exhibited more simplified data structures in the UMAP space than those of the original contraction waves (Fig. 2B) and thus contributed to an improvement in the performance of the SVM.

The data augmentation was also valid. Our datasets were relatively small in size and were imbalanced between 'normal' and 'abnormal' classes. These factors may obscure the decision boundary of the SVM. Data augmentation enhances the robustness of the decision boundary and indeed succeeded in improving the overall classification performance of the SVM. This improvement is not trivial. If our data were sufficiently large, deep learning tools, such as one-dimensional convolutional neural networks, would be applicable to improve the classification performance [23]. In practice, large datasets are not always available, especially when rare and expensive materials, such as human-induced pluripotent stem cells, are used. Our approach represents one of the solutions when the amount of data is limited.

Finally, our method must be carried out with careful consideration because we used contraction waves to quantify the contractile properties of hiPSC-CMs. We obtained the contraction waves from the entire image in each video frame, with the local contractile properties being averaged out because we aimed to focus on global contractile properties. Thus, our method may overlook local abnormalities in the contractile properties of hiPSC-CMs. To scrutinize the local contractile properties, one has to use raw video data as an input of machine learning. At present, these machine learning techniques that can handle time-lapse video data are still premature and computationally insufficient [24].

Declaration of competing interest

The authors declare no conflict of interest associated with this manuscript.

Acknowledgments

This work was supported by JST ERATO (JPMJER1801), JSPS Grants-in-Aid for Scientific Research (18H05525), and the Human Frontier Science Program (RGP0019/2016). This work was conducted partially as a program at the International Research Center for Neurointelligence (WPI-IRCN) of The University of Tokyo Institutes for Advanced Study at The University of Tokyo.

References

[1] J.L. Stevens, T.K. Baker, The future of drug safety testing: expanding the view and narrowing the focus, Drug Discov. Today 14 (2009) 162–167, https://doi.org/10.1016/j.drudis.2008.11.009.

[2] M. Takeda, S. Miyagawa, S. Fukushima, A. Saito, E. Ito, A. Harada, R. Matsuura, H. Iseoka, N. Sougawa, N. Mochizuki-Oda, M. Matsusaki, M. Akashi, Y. Sawa, Development of in vitro drug-induced cardiotoxicity assay by using three-dimensional cardiac tissues derived from human induced pluripotent stem cells, Tissue Eng, Part C Methods 24 (2018) 56–67, https://doi.org/10.1089/ten.tec.2017.0247.

[3] A. Sharma, W.L. McKeithan, R. Serrano, T. Kitani, P.W. Burridge, J.C. del Álamo, M. Mercola, J.C. Wu, Use of human induced pluripotent stem cell–derived cardiomyocytes to assess drug cardiotoxicity, Nat. Protoc. 13 (2018) 3018–3041, https://doi.org/10.1038/s41596-018-0076-8.

[4] B.J. van Meer, A. Krotenberg, L. Sala, R.P. Davis, T. Eschenhagen, C. Denning, L.G.J. Tertoolen, C.L. Mummery, Simultaneous measurement of excitation-contraction coupling parameters identifies mechanisms underlying contractile responses of hiPSC-derived cardiomyocytes, Nat. Commun. 10 (2019) 1–9, https://doi.org/10.1038/s41467-019-12354-8.

[5] G. Gintant, P.T. Sager, N. Stockbridge, Evolution of strategies to improve preclinical cardiac safety testing, Nat. Rev. Drug Discov. 15 (2016) 457–471, https://doi.org/10.1038/nrd.2015.34.

[6] H. Wei, C. Wang, R. Guo, K. Takahashi, K. Naruse, Development of a model of ischemic heart disease using cardiomyocytes differentiated from human induced pluripotent stem cells, Biochem. Biophys. Res. Commun. 520 (2019) 600–605, https://doi.org/10.1016/j.bbrc.2019.09.119.

[7] K. Orita, K. Sawada, R. Koyama, Y. Ikegaya, Deep learning-based quality control of cultured human-induced pluripotent stem cell-derived cardiomyocytes, J. Pharmacol. Sci. 140 (2019) 313–316, https://doi.org/10.1016/j.jphs.2019.04.008.

[8] E.K. Lee, D.D. Tran, W. Keung, P. Chan, G. Wong, C.W. Chan, K.D. Costa, R.A. Li, M. Khine, Machine learning of human pluripotent stem cell-derived engineered cardiac tissue contractility for automated drug classification, Stem Cell Reports 9 (2017) 1560–1572, https://doi.org/10.1016/j.stemcr.2017.09.008.

[9] S.B. Kotsiantis, D. Kanellopoulos, P.E. Pintelas, Data preprocessing for supervised leaning, Int. J. Comput. Sci. 1 (2006) 111–117, https://doi.org/10.5281/zenodo.1082415.

[10] A. Le Guennec, S. Malinowski, R. Tavenard, Data augmentation for time series classification using convolutional neural networks, in: ECML/PKDD Work. Adv. Anal. Learn, Temporal Data, 2016.

[11] B.E. Boser, I.M. Guyon, V.N. Vapnik, A training algorithm for optimal margin classifiers, in: Proc. Fifth Annu. Work. Comput. Learn. Theory - COLT '92, ACM Press, New York, New York, USA, 1992, pp. 144–152, https://doi.org/10.1145/130385.130401.

[12] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis, Nat. Methods 9 (2012) 676–682, https://doi.org/10.1038/nmeth.2019.

[13] R.J. Hyndman, Moving averages, in: M. Lovric (Ed.), Int. Encycl. Stat. Sci., Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 866–869, https://doi.org/10.1007/978-3-642-04898-2_380.

[14] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, in: C.S. Mellish (Ed.), Proc. 11th Int. Jt. Conf. Artif. Intell., Morgan Kaufmann, 1995, pp. 1137–1143.

[15] M. Ali, M.W. Jones, X. Xie, M. Williams, TimeCluster: dimension reduction applied to temporal data for visual analytics, Vis. Comput. 35 (2019) 1013–1026, https://doi.org/10.1007/s00371-019-01673-y.

[16] E.O. Brigham, R.E. Morrow, The fast Fourier transform, IEEE Spectr 4 (1967) 63–70, https://doi.org/10.1109/MSPEC.1967.5217220.

[17] L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, 2018. http://arxiv.org/abs/1802.03426.

[18] E. Becht, L. McInnes, J. Healy, C.-A. Dutertre, I.W.H. Kwok, L.G. Ng, F. Ginhoux, E.W. Newell, Dimensionality reduction for visualizing single-cell data using UMAP, Nat. Biotechnol. (2018), https://doi.org/10.1038/nbt.4314.

[19] M. Sato, N. Matsumoto, A. Noguchi, T. Okonogi, T. Sasaki, Y. Ikegaya, Simultaneous monitoring of mouse respiratory and cardiac rates through a single precordial electrode, J. Pharmacol. Sci. 137 (2018) 177–186, https://doi.org/10.1016/j.jphs.2018.06.009.

[20] N. Matsumoto, Y. Takahara, N. Matsuki, Y. Ikegaya, Thoracotomy reduces intrinsic brain movement caused by heartbeat and respiration: a simple method to prevent motion artifact for in vivo experiments, Neurosci. Res. 71 (2011) 188–191, https://doi.org/10.1016/j.neures.2011.06.011.

[21] T. Dao, A. Gu, A.J. Ratner, V. Smith, C. De Sa, C. Ré, A kernel theory of modern data augmentation, Proc. Mach. Learn. Res. 97 (2019) 1528–1537. http://www.ncbi.nlm.nih.gov/pubmed/31777848.

[22] J. Aoe, R. Fukuma, T. Yanagisawa, T. Harada, M. Tanaka, M. Kobayashi, Y. Inoue, S. Yamamoto, Y. Ohnishi, H. Kishima, Automatic diagnosis of neurological diseases using MEG signals with a deep neural network, Sci. Rep. 9 (2019) 5057, https://doi.org/10.1038/s41598-019-41500-x.

[23] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, D.J. Inman, 1D Convolutional Neural Networks and Applications: A Survey, 2019. http://arxiv.org/abs/1905.03554.

[24] D. Tran, J. Ray, Z. Shou, S.-F. Chang, M. Paluri, ConvNet Architecture Search for Spatiotemporal Feature Learning, 2017. http://arxiv.org/abs/1708.05038.