

An unsupervised domain adaptation approach to classification of stem cell-derived cardiomyocytes

Carolina Pacheco and René Vidal

Center for Imaging Science, Mathematical Institute for Data Science,
Department of Biomedical Engineering, Johns Hopkins University.

Abstract. The use of human embryonic stem cell-derived cardiomyocytes (hESC-CMs) in applications such as cardiac regenerative medicine requires understanding them in the context of adult CMs. Their classification in terms of the major adult CM phenotypes is a crucial step to build this understanding. However, this is a challenging problem due to the lack of labels for hESC-CMs. Adult CM phenotypes are easily distinguishable based on the shape of their action potentials (APs), but it is still unclear how these phenotypes are expressed in the APs of hESC-CM populations. Recently, a metamorphosis distance was proposed to measure similarities between hESC-CM APs and adult CM APs, which led to state-of-the-art performance when used in a 1 nearest neighbor scheme. However, its computation is prohibitively expensive for large datasets. A recurrent neural network (RNN) classifier was recently shown to be computationally more efficient than the metamorphosis-based method, but at the expense of accuracy. In this paper we argue that the APs of adult CMs and hESC-CMs intrinsically belong to different domains, and propose an unsupervised domain adaptation approach to train the RNN classifier. The idea is to capture the domain shift between hESC-CMs and adult CMs by adding a term to the loss function that penalizes their maximum mean discrepancy (MMD) in feature space. Experimental results in an unlabeled 6940 hESC-CM dataset show that our approach outperforms the state of the art in terms of both clustering quality and computational efficiency. Moreover, it achieves state-of-the-art classification accuracy in a completely different dataset without retraining, which demonstrates the generalization capacity of the proposed method.

Keywords: Domain adaptation, LSTM, embryonic cardiomyocytes

1 Introduction

The insufficient supply of oxygen-rich blood to the heart, known as Ischaemic Heart Disease (IHD), has remained the global leading cause of death for more than 15 years, taking the lives of almost 18 million people in 2016 [1]. Along with prevention, there is a need for innovative approaches to treat IHD. In particular, cardiomyocyte (CM) transplantation has shown favorable results of remuscularization in animals [2], which is promising for post-myocardial infarction patients.

Human embryonic stem cell-derived cardiomyocytes (hESC-CMs) successfully resemble embryonic CMs in terms of structure and function [3]. Thus, they are an important source of CMs not only for regenerative medicine, but also for other applications such as drug screening. However, any application requires to understand hESC-CM characteristics relative to the ones of adults CMs, which remain unclear. A first step in that direction is to study the presence of the major adult CM phenotypes (atrial, ventricular, etc) in hESC-CM populations.

Initial approaches to classification of hESC-CMs like [4] or [5] were based on handcrafted action potential (AP) features with ad-hoc thresholds in small datasets, which made them subjective and difficult to transfer to other datasets. New electrophysiological recording techniques increased the size of the datasets from dozens and hundreds to several thousand samples [6], which required automatic methods for their classification. However, the lack of ground truth labels for hESC-CMs makes this problem challenging. Gorospe et al. [7, 8] proposed to leverage the existence of electrophysiological models of adult CMs to build a 1 nearest neighbor classifier based on a minimum deformation distance called metamorphosis. Their method, 1NN-M, achieved good performance in hESC-CM populations but it is computationally expensive because it requires to solve 20 optimization problems for each new sample to be classified. To overcome this drawback, a recurrent neural network (RNN) based classifier was proposed in [9]. They trained the RNN in a semi-supervised way using labeled adult CM APs and unlabeled hESC-CM APs, showing significant computational advantages with respect to 1NN-M and reaching similar, but not better, clustering quality. Although its computational advantages are undeniable, the approach presented in [9] lacks in a fundamental aspect: the classifier is not aware of the existence of two different domains. Classical machine learning algorithms rely on the assumption that training and testing data are sampled from the same distribution. Unfortunately, this assumption does not hold in the case of adult CMs and hESC-CMs, and therefore we argue that a domain adaptation approach is needed to appropriately train the RNN.

Domain adaptation addresses the problem of optimizing the performance in one domain (called target domain), given training data in a different domain (called source domain). We propose to use the RNN architecture presented in [9], but train it in a different way. We consider the output of its hidden layer as a feature space shared by adult CMs and hESC-CMs. The domain shift between their distributions in the feature space is then reduced by adding their maximum mean discrepancy (MMD) to the loss function. The RNN classifier is trained using a subset of 1600 samples from an unlabeled 6940 hESC-CM APs dataset and 1600 adult CM APs from electrophysiological models. Experimental results confirm that the addition of a domain adaptation term to the loss function improves with respect to the state of the art in terms of clustering quality, and at the same time keeps the computational advantages of previous RNN-based approaches. Moreover, it also reaches state-of-the-art classification accuracy in a completely different dataset without retraining (outperforming previous RNN-based methods), which further demonstrate the advantages of our approach.

2 Methods

2.1 Problem formulation

Let $\Omega_e = \{\mathbf{x}_j^e\}_{j=1}^{N_e}$ be an *unlabeled* hESC-CM APs dataset from the target domain, where the sequence $\mathbf{x}_j^e = \{x_j^e(k) \in \mathbb{R}\}_{k=1}^K$ represents the j th hESC-CM AP and K is the total number of samples in one cycle length. Hereafter we will refer to this dataset as *embryonic* because of its resemblance of embryonic CM APs. Let $\Omega_a = \{(\mathbf{x}_i^a, y_i^a)\}_{i=1}^{N_a}$ be a *labeled* adult dataset from the source domain, where $\mathbf{x}_i^a = \{x_i^a(k) \in \mathbb{R}\}_{k=1}^K$ is the i th adult AP and $y_i^a \in \{0, 1\}$ is its ground truth label ($y_i^a = 0$ denotes atrial and $y_i^a = 1$ denotes ventricular). We consider the problem of assigning a label \hat{y}_j^e to each $\mathbf{x}_j^e \in \Omega_e$, where $\hat{y}_j^e = 0$ denotes atrial-like and $\hat{y}_j^e = 1$ denotes ventricular-like. Let $\delta \in \{e, a\}$ indicate the embryonic or adult domain. $\mathbb{P}\{\mathbf{x}|\delta = e\}$ denotes the probability density function of APs in embryonic domain and $\mathbb{P}\{\mathbf{x}|\delta = a\}$ denotes the probability density function of APs in the adult domain. We assume: (i) $\mathbb{P}\{\mathbf{x}|\delta = e\} \neq \mathbb{P}\{\mathbf{x}|\delta = a\}$, and (ii) $\mathbb{P}\{y|\mathbf{x}, \delta = e\} = \mathbb{P}\{y|\mathbf{x}, \delta = a\}$ (covariate shift assumption).

Classifying samples from Ω_e using training data from Ω_a corresponds to unsupervised domain adaptation, which according to [10] can be addressed via: instance weighting, self-labeling approaches, clustering-based methods, or feature representation methods. Instance weighting approaches require shared support between both distributions, which does not hold in our case because there are embryonic APs never observed in adult data. On the other hand, self-labeling approaches as well as clustering-based methods often rely on computing similarities between samples, which can be computationally expensive for APs. Thus, we use a feature representation approach in which probability distribution functions of both domains are forced to be similar in a learned feature space $\varphi(\mathbf{x})$.

2.2 Maximum Mean Discrepancy

Maximum mean discrepancy [11] corresponds to the distance between the mean of two probability distribution functions mapped into a reproducing kernel Hilbert space (RKHS), embedding their samples via $\psi(\cdot)$. An estimation of the MMD between two datasets $\Omega_a = \{\mathbf{x}_i^a\}_{i=1}^{N_a}$ and $\Omega_e = \{\mathbf{x}_j^e\}_{j=1}^{N_e}$ is given by

$$\widehat{\mathcal{MMD}}^2(\Omega_a, \Omega_e) = \sum_{i=1}^{N_a} \sum_{i'=1}^{N_a} \frac{\mathcal{K}(\mathbf{x}_i^a, \mathbf{x}_{i'}^a)}{N_a^2} + \sum_{j=1}^{N_e} \sum_{j'=1}^{N_e} \frac{\mathcal{K}(\mathbf{x}_j^e, \mathbf{x}_{j'}^e)}{N_e^2} - \sum_{i=1}^{N_a} \sum_{j=1}^{N_e} \frac{2\mathcal{K}(\mathbf{x}_i^a, \mathbf{x}_j^e)}{N_a N_e}, \quad (1)$$

where $\mathcal{K}(x, y)$ is a positive semidefinite kernel such that $\mathcal{K}(x, y) = \psi(x)^\top \psi(y)$. The Gaussian kernel $\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma_k^2}\right)$ is commonly used.

The equation in (1) allows us to estimate how different the distributions are based on their samples. The MMD estimator has been successfully applied to learn appropriate kernels for cross-domain SVM-based classification, regression and video concept detection, among others [12, 13]. This estimator has also been

recently applied with fixed kernels as a metric to learn the parameters of generative networks [14], and the parameters of feature extraction layers for multi-task learning in multiple domains [15], which is closely related to our task.

2.3 Network Architecture

As shown in Fig. 1, we use the architecture proposed in [9]: one input layer, one hidden LSTM layer of dimension $p = 3$, and a single sigmoid unit as output layer. The LSTM layer is explicitly considered as a feature extractor, such that $\mathbf{x} \mapsto \varphi_{\mathcal{W}_{\mathcal{F}}}(\mathbf{x}) = h(\mathbf{x}, K) \in \mathbb{R}^3$. Note that $h(\mathbf{x}, K)$ emphasizes that the output cell at time K depends on the entire input sequence.

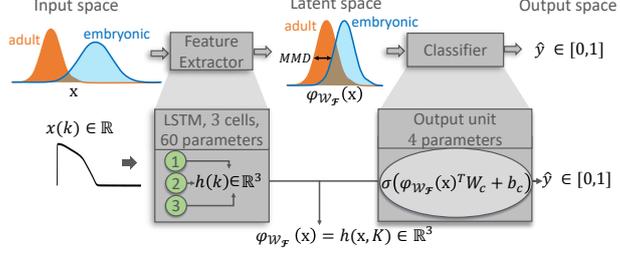


Fig. 1: Domain adapted classification approach.

The feature vector $\varphi_{\mathcal{W}_{\mathcal{F}}}$ depends on the parameters of the feature extractor $\mathcal{W}_{\mathcal{F}}$, and thus it is learnable. For a given set of parameters $\mathcal{W} = \{\mathcal{W}_{\mathcal{F}}, W_c, b_c\}$, we represent the classifier as the function $f_{\mathcal{W}}(\mathbf{x}) = \hat{y}$ that maps an action potential \mathbf{x} to a predicted label \hat{y} .

2.4 Domain Adaptation Objective Function

We aim to enforce similarity between the probability density functions of both domains in feature space, i.e. $\mathbb{P}\{\varphi_{\mathcal{W}_{\mathcal{F}}}(\mathbf{x})|\delta = a\} \approx \mathbb{P}\{\varphi_{\mathcal{W}_{\mathcal{F}}}(\mathbf{x})|\delta = e\}$, while training a classifier with source domain data. In that sense, the network learns to classify samples in a space in which embryonic and adult data “are similar”. We propose an objective function that builds on top of the semisupervised loss presented in [9] as follows

$$\frac{1 - \lambda}{N_a} \left(\sum_{i=1}^{N_a} \ell(y_i^a, f_{\mathcal{W}}(\mathbf{x}_i^a)) \right) + \frac{\lambda}{N_e(N_e - 1)} \left(\sum_{j=1}^{N_e} \sum_{j' \neq j} \ell_u(f_{\mathcal{W}}(\mathbf{x}_j^e), f_{\mathcal{W}}(\mathbf{x}_{j'}^e)) \right) + \gamma \widehat{\mathcal{MMD}}^2 \left(\{\varphi_{\mathcal{W}_{\mathcal{F}}}(\mathbf{x}_i^a)\}_{i=1}^{N_a}, \{\varphi_{\mathcal{W}_{\mathcal{F}}}(\mathbf{x}_j^e)\}_{j=1}^{N_e} \right), \quad (2)$$

where $\gamma \geq 0$ and $\lambda \in [0, 1]$ modulate the relative importance given to the domain adaptation term and unsupervised term, respectively. The first two terms of (2) correspond to the supervised and unsupervised losses presented in [9]. $\ell(\cdot, \cdot)$ is the crossentropy loss, and $\ell_u(\cdot, \cdot)$ is an unsupervised contrastive loss that depends on the similarity between the samples. The domain adaptation term is the square of the empirical estimator of MMD in feature space using a Gaussian kernel.

2.5 Metrics

Classification Accuracy. When ground truth labels are available, it is the ratio between correctly classified samples and the total number of samples.

Davies-Bouldin Index (DBI) [16]. The DBI is a measure of clustering quality, which is used as a proxy for classification performance when ground truth labels are not available. It corresponds to the ratio between the intra-cluster and inter-cluster dispersion and should be as small as possible.

3 Experiments

3.1 AP Data

Adult CM AP Data. Electrophysiological models of adult atrial [17] and ventricular [18] CMs were paced at 1.5 Hz to generate 800 APs of each class by randomizing their parameters as described in [9]. Fig. 2(a) shows their normalized version (maximum value 1 and resting membrane potential 0).

hESC-CM AP Data: Optical recording dataset [6]. Large *unlabeled* dataset composed of 6940 APs optically recorded from 9 cell aggregates paced at 1.5 Hz. Fig. 2(b) shows the normalized hESC-CM APs from this dataset.

hESC-CM AP Data: Single cell recording dataset [4]. Small *labeled* dataset composed of 52 APs recorded from spontaneously beating hESC-CMs. The nonlinear mapping proposed in [19] was used to adjust them to 1.5 Hz pacing rate. As shown in Fig. 2(c), 16 of them are atrial-like (blue) and 36 of them are ventricular-like (red). This dataset is only used for testing purposes.

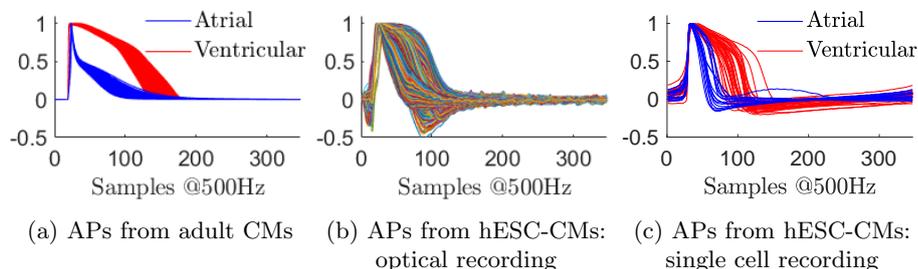


Fig. 2: Action potentials: (a) 1600 adult CMs, (b) 6940 unlabeled hESC-CMs, and (c) 52 labeled hESC-CMs.

3.2 Implementation Details

The classifier was implemented in Keras [20] with TensorFlow backend and trained using the RMSprop optimizer ($\epsilon = 0.003$). The weights were initialized from the Sup-LSTM network presented in [9]: a network trained for 100 epochs in a fully supervised way using 300 adult atrial and ventricular APs.

$N_e = 1600$ embryonic APs from the optical recording dataset and $N_a = 1600$ adult APs are used for training and validation in balanced batches of 32 samples. The data was split into 10 folds, and the average performance of 10-fold crossvalidation experiments is reported. Once the network is trained, a forward pass in the optical recording dataset is performed to compute the clustering quality (DBI) of the output. Similarly, a forward pass in the single cell recording dataset is done to compute the classification accuracy.

Four cases are studied: (i) Supervised learning $\lambda = 0$ and $\gamma = 0$ (Sup-LSTM); (ii) Semisupervised learning with metamorphosis distances $\lambda = 0.1$ and $\gamma = 0$ (Semi-M-LSTM), (iii) Supervised learning with domain adaptation $\lambda = 0$ and $\gamma = 1$ (DA-Sup-LSTM); and (iv) Semisupervised learning with metamorphosis distances and domain adaptation $\lambda = 0.1$ and $\gamma = 5$ (DA-Semi-M-LSTM). They are trained for 100 epochs, except Sup-LSTM which converges in 15 epochs.

3.3 Results

Classification results for the 9 cell aggregates of the optical recording dataset are depicted in Fig. 3 along with the mean clustering quality index (DBI). In all cases the LSTM network suggests heterogeneity in the cell clusters and generates smooth classification boundaries, which coincides with previous findings [6, 8]. As reported in [9], Sup-LSTM results are easily distinguishable from all the other approaches and lead to significantly worse clustering quality, which supports the idea that adult CMs and hESC-CMs intrinsically belong to different domains. In that sense, the simple addition of the domain adaptation term to the loss function (DA-Sup-LSTM) improves the mean clustering quality from 0.2793 to 0.2412, which makes it comparable to Semi-M-LSTM (0.2449), but with significant computational advantages since it does not require any computation of metamorphosis distances. However, the addition of the unsupervised term along with the domain adaptation (DA-Semi-M-LSTM) leads to the best performance, outperforming the state-of-the-art method (1NN-M) in terms of clustering quality. Therefore, the effects of the semi-supervised term and the domain adaptation term seem to be complementary.

Table 1 and Fig. 4 summarize the results for the studied cases in terms of the clustering quality achieved in the optical recording dataset as well as the classification accuracy obtained in the single cell recording dataset.

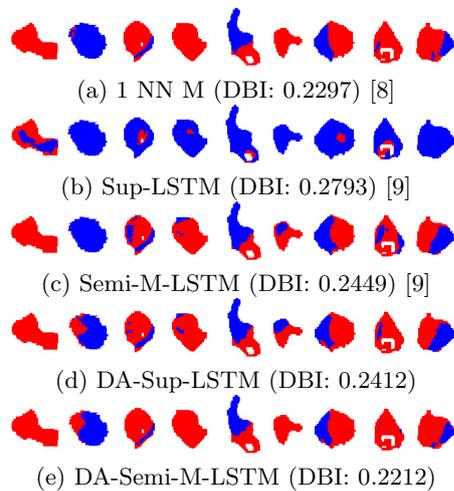


Fig. 3: LSTM classification results (each pixel corresponds to one hESC-CM AP). Blue indicates atrial-like phenotype and red indicates ventricular-like phenotype.

Table 1: Performance comparison in the 10-fold crossvalidation experiments.

	Accuracy \uparrow			DBI \downarrow		
	mean	median	(std)	mean	median	(std)
1NN M [8]	0.9615	0.9615	(0.0000)	0.2297	0.2297	(0.0000)
Sup-LSTM [9]	0.3269	0.3269	(0.0000)	0.2793	0.2795	(0.0009)
Semi-M-LSTM [9]	0.7154	0.7596	(0.1301)	0.2449	0.2420	(0.0059)
DA-Sup-LSTM (ours)	0.8385	0.9135	(0.1339)	0.2412	0.2408	(0.0047)
DA-Semi-M-LSTM (ours)	0.9673	0.9904	(0.0472)	0.2212	0.2197	(0.0072)

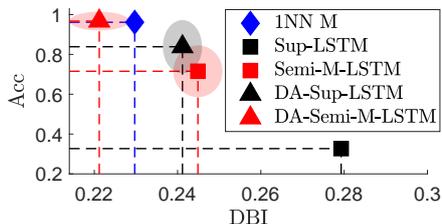


Fig. 4: Comparison of performance in 10-fold crossvalidation. Mean performance is marked by solid-colored symbols and variability is shown by translucent ellipses (whose semi-axes correspond to standard deviations).

(DA-Semi-M-LSTM) that succeeds in outperforming the 1NN-M also in terms of classification accuracy (0.9673). This is a powerful result because our approach not only outperforms the state-of-the-art method, but also it is significantly faster. The classification of the single cell recording dataset is reported to take approximately 12 seconds with the most efficient algorithm for 1NN-M in 2 8-core computer nodes with 8 hyperthreaded 2.3 GHz CPUs per node [8], whereas in our case it takes less than 0.4 seconds in one 2.2 GHz CPU with 2 cores, 4 threads.

The first row of Fig. 5 shows the distribution of adult samples and unlabeled hESC-CM samples in latent space in the four cases studied. Noticeably, atrial and ventricular samples are located in different areas of the feature space, and the hESC-CMs form a one dimensional path between them. The second row of Fig. 5 represents the histogram of samples along this path. The effect that the domain adaptation term has in the data distribution in the feature space is subtle for the supervised case (DA-Sup-LSTM vs Sup-LSTM): it induces a more balanced distribution of hESC-CMs samples in the location of atrial and ventricular adult CMs. In the semisupervised case however, Semi-M-LSTM generates a high concentration of embryonic samples far from adult CM data, and the addition of the domain adaptation term significantly reduces this effect (DA-Semi-M-LSTM).

Whereas Semi-M-LSTM and DA-Sup-LSTM are comparable in terms of clustering quality, the domain adaptation approach performs significantly better in terms of classification accuracy (0.8385 vs 0.7154). Note that this corresponds to a forward pass on the single cell recording dataset, so it shows that the the domain adaption method provides better generalization across datasets. However, their classification accuracy is still far from the one achieved by 1NN-M (0.9615). Again it is the complementary action of semisupervised and domain adaptation terms

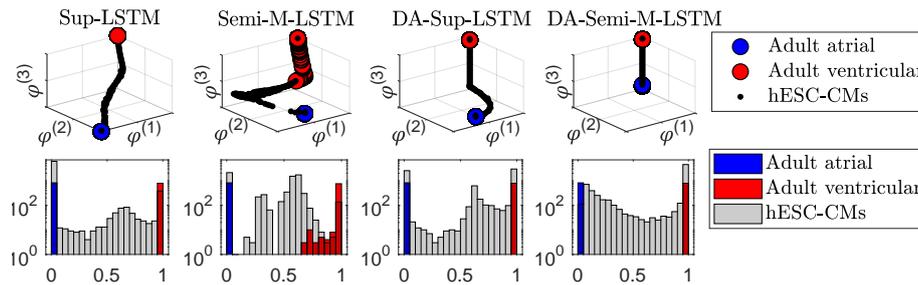


Fig. 5: Distribution of adult CMs and hESC-CMs in latent space (first row) and histogram of their projection on the one dimensional path (second row).

4 Conclusion

We have applied for the first time the concept of domain adaptation to address discrepancies related to the stem cell differentiation process. Moreover, the proposed approach has proven to be useful, since it outperforms the state-of-the-art method for classification of hESC-CM APs not only in terms of clustering quality, but also in terms of computational efficiency and inter-dataset generalization.

Acknowledgement. This work has been supported by NIH #5R01HD87133. The authors thank Dr. Gianni Gorospe for insightful discussions, and Dr. Renjun Zhu and Prof. Leslie Tung for providing the hESC-CMs dataset.

References

1. World Health Organization: World Health Statistics 2018: Monitoring health for the Sustainable Development Goals. (2018)
2. Hartman, M.E., Chong, J.J., Laflamme, M.A.: State of the art in cardiomyocyte transplantation. In: Cardiac Regeneration. Springer (2017) 177–218
3. Kehat, I., Kenyagin-Karsenti, D., Snir, M., Segev, H., Amit, M., Gepstein, A., Livne, E., Binah, O., Itskovitz-Eldor, J., Gepstein, L.: Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. *The Journal of clinical investigation* **108**(3) (2001) 407–414
4. He, J.Q., Ma, Y., Lee, Y., Thomson, J.A., Kamp, T.J.: Human embryonic stem cells develop into multiple types of cardiac myocytes: action potential characterization. *Circulation research* **93**(1) (2003) 32–39
5. Sartiani, L., Bettioli, E., Stillitano, F., Mugelli, A., Cerbai, E., Jaconi, M.: Developmental changes in cardiomyocytes differentiated from human embryonic stem cells: a molecular and electrophysiological approach. *Stem cells* **25**(5) (2007) 1136–1144
6. Zhu, R., Millrod, M.A., Zambidis, E.T., Tung, L.: Variability of action potentials within and among cardiac cell clusters derived from human embryonic stem cells. *Scientific reports* **6** (2016) 18544
7. Gorospe, G., Younes, L., Tung, L., Vidal, R.: A metamorphosis distance for embryonic cardiac action potential interpolation and classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer (2013) 469–476

8. Gorospe, G., Zhu, R., He, J.Q., Tung, L., Younes, L., Vidal, R.: Efficient metamorphosis computation for classifying embryonic cardiac action potentials. In: 5th Workshop on Mathematical Foundations of Computational Anatomy. (2015)
9. Pacheco, C., Vidal, R.: Recurrent neural networks for classifying human embryonic stem cell-derived cardiomyocytes. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer (2018) 581–589
10. Margolis, A.: A literature review of domain adaptation with unlabeled data. Technical Report (2011) 1–42
11. Gretton, A., Borgwardt, K.M., Rasch, M., Schölkopf, B., Smola, A.J.: A kernel method for the two-sample-problem. In: Advances in neural information processing systems. (2007) 513–520
12. Pan, S.J., Kwok, J.T., Yang, Q.: Transfer learning via dimensionality reduction. In: AAAI Conference on Artificial Intelligence. Volume 8. (2008) 677–682
13. Duan, L., Tsang, I.W., Xu, D., Maybank, S.J.: Domain transfer svm for video concept detection. In: IEEE Conference on Computer Vision and Pattern Recognition. (2009) 1375–1381
14. Dziugaite, G., Roy, D., Ghahramani, Z.: Training generative neural networks via maximum mean discrepancy optimization. In: The Conference on Uncertainty in Artificial Intelligence. (2015) 258–267
15. Chen, H.Y., Chien, J.T.: Deep semi-supervised learning for domain adaptation. In: IEEE International Workshop on Machine Learning and Signal Processing, IEEE (2015) 1–6
16. Davies, D.L., Bouldin, D.W.: A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2) (1979) 224–227
17. Nygren, A., Fiset, C., Firek, L., Clark, J., Lindblad, D., Clark, R., Giles, W.: Mathematical model of an adult human atrial cell: the role of k⁺ currents in repolarization. *Circulation research* **82**(1) (1998) 63–81
18. O’Hara, T., Virág, L., Varró, A., Rudy, Y.: Simulation of the undiseased human cardiac ventricular action potential: model formulation and experimental validation. *PLoS computational biology* **7**(5) (2011) e1002061
19. Iravanian, S., Tung, L.: A novel algorithm for cardiac biosignal filtering based on filtered residue method. *IEEE Transactions on Biomedical Engineering* **49**(11) (2002) 1310–1317
20. Chollet, F., et al.: Keras. <https://keras.io> (2015)