Recurrent Neural Networks for Classifying Human Embryonic 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,
Baltimore, USA
cpachec2@jhu.edu

Abstract. Classification of human embryonic stem cell-derived cardiomyocytes (hESC-CMs) is important for many applications in cardiac regenerative medicine. However, a key challenge is the lack of ground truth labels for hESC-CMs: Whereas adult phenotypes are well-characterized in terms of their action potentials (APs), the understanding of how the shape of the AP of immature CMs relates to that of adult CMs remains limited. Recently, a new metamorphosis distance has been proposed to determine if a query immature AP is closer to a particular adult AP phenotype. However, the metamorphosis distance is difficult to compute making it unsuitable for classifying a large number of CMs. In this paper we propose a semi-supervised learning framework for the classification of hESC-CM APs. The proposed framework is based on a recurrent neural network with LSTM units whose parameters are learned by minimizing a loss consisting of two parts. The supervised part uses labeled data obtained from computational models of adult CMs, while the unsupervised part uses the metamorphosis distance in an efficient way. Experiments confirm the benefit of integrating information from both adult and stem cell-derived domains in the learning scheme, and also show that the proposed method generates results similar to the state-of-the-art (94.73%) with clear computational advantages when applied to new samples.

Keywords: Semi-supervised learning · LSTM
Embryonic cardiomyocytes

1 Introduction

The unquestioned cardiomyogenic potential of human embryonic stem cells [1] and the well-established protocols for their isolation and maintenance [2] make them one of the most promising sources of cardiomyocytes (CMs) for applications such as cell-based cardiac repair [3] and drug screening [4]. However, their use is still hampered by the current limited understanding of the phenotypic traits of

human stem cell-derived cardiomyocytes (hESC-CMs) and their relationship to the phenotypes of adult CMs [3]. To characterize the phenotype of hESC-CMs, prior work [5] studies the expression of specific genes and ion channel-encoding subunits. Alternatively, [6,7] apply thresholding to simple features extracted from the cell's action potential (AP). However, simple classification methods based on handcrafted features and subjective criteria not only discard most of the information contained in the AP, but also are hardly transferable.

Recently, automatic methods have been proposed to analyze the heterogeneity of hESC-CMs APs using the whole AP as an input. For instance, the existence of different clusters was studied via a spectral grouping-based algorithm in [8], and the metamorphosis distance proposed in [9] was adapted in [10] to classify embryonic APs by computing their distances to adult APs with known phenotype. While this new method shows better interpolation and clustering results [11], it is too computationally intensive to be applicable to large-scale datasets.

In this paper we propose a new method for classifying hESC-CMs APs based on recurrent neural networks (RNNs) with long short term memory (LSTM) units [12]. LSTMs have recently re-gained popularity for time series classification because of their great performance in applications to speech recognition [13] and activity recognition [14]. However, while LSTMs have also been successfully applied to the analysis of physiological signals [15,16], standard LSTMs are not directly applicable to the classification of hESC-CMs because of the lack of labels for embryonic APs. In this context, the contribution of this paper is to propose a semi-supervised approach that exploits the abundance of labels for adult APs, which can be obtained via simulation of electrophysiological models for the typical adult phenotypes (atrial, ventricular, etc.). The proposed semi-supervised approach uses a novel loss function to train an LSTM that combines a classification loss for adult APs (supervised part) and a contrastive loss for embryonic APs (unsupervised part). For the supervised part we use synthetic APs obtained from computational models of adult atrial [17] and ventricular [18] CMs, while for the unsupervised part we compute similarities between APs, making efficient use of Euclidean and metamorphosis distances.

Experiments on a dataset of 6940 hESC-CMs APs show that our semi-supervised approach provides smooth clustering results that are comparable to those presented in [11] in terms of Davies-Bouldin Index (DBI), and also confirm the benefit of integrating information from both adult and embryonic APs. Furthermore, the semi-supervised approach is able to use the Euclidean metric more effectively than previous methods, considerably outperforming the 1-nearest neighbor scheme (87.88% vs 62.90% of agreement with the best result published in [11]). When the metamorphosis distance is used, our method achieves 94.73% of agreement with the best results published in [11], but it is significantly less computationally expensive when applied to new data.

2   Methods

Let the sequence $\mathbf{x}_j^e = \{x_j^e(k) \in \mathbb{R}\}_{k=1}^T$, where T is the total number of samples in one cycle length, represent the jth embryonic AP. Let $\mathbf{x}_i^a$ be the ith adult AP

and let $y_i^a \in \{0, 1\}$ be its ground truth label, where $y_i^a = 0$ denotes atrial and $y_i^a = 1$ denotes ventricular. We consider the problem of assigning a label $\hat{y}_j^e$ to each $\mathbf{x}_j^e$, where $\hat{y}_j^e = 0$ denotes atrial-like and $\hat{y}_j^e = 1$ denotes ventricular-like.

A simple approach is to use a 1 nearest-neighbor (1NN) classifier with the Euclidean distance $d_E(\mathbf{x}_j^e, \mathbf{x}_i^a) = \frac{1}{\sigma_M}\sqrt{\sum_{k=1}^{T} \left(x_j^e(k) - x_i^a(k)\right)^2}$, where σ_M is a normalization parameter. However, the Euclidean distance can be affected by nuisance factors such as changes in AP shape induced by the maturation process.

An alternative approach is to use 1NN classification with the metamorphosis distance, which generates an interpolation path $x(k, s)$ between an embryonic AP, $x(k, 0) = x_j^e(k)$, and an adult AP, $x(k, S) = x_i^a(k)$, that minimizes the amount of deformation between the two, which depends on a certain velocity $\mathbf{v}$:

$$d_M^2(\mathbf{x}_j^e, \mathbf{x}_i^a) = \min_{\mathbf{x},\mathbf{v}} \sum_{s=0}^{S-1} \|v(k, s)\|_V^2 + \frac{1}{\sigma_M^2}\|x(k + v(k, s), s + 1) - x(k, s)\|_2^2, \quad (1)$$

where $\|\cdot\|_V^2$ is a Sobolev norm and σ_M^2 is a balancing parameter (see [10,11]). However, the metamorphosis distance is computationally intensive to evaluate.

2.1 Classifier Architecture

Long Short-Term Memory (LSTM) units [12] are recurrent blocks whose key elements are input gates $i(k)$, forget gates $f(k)$ and output gates $o(k)$ that modulate the evolution of the state $c(k)$ and its output $h(k)$ at time k as follows

$$\begin{aligned}
i(k) &= \sigma\left(W_i x(k) + U_i h(k-1) + b_i\right) \in \mathbb{R}^p \\
f(k) &= \sigma\left(W_f x(k) + U_f h(k-1) + b_f\right) \in \mathbb{R}^p \\
o(k) &= \sigma\left(W_o x(k) + U_o h(k-1) + b_o\right) \in \mathbb{R}^p \\
c(k) &= f(k) \circ c(k-1) + i(k) \circ \tanh\left(W_c x(k) + U_c h(k-1) + b_c\right) \in \mathbb{R}^p \\
h(k) &= o(k) \circ \tanh\left(c(k)\right) \in \mathbb{R}^p
\end{aligned} \quad (2)$$

where p denotes the layer dimension, $\sigma(z) = \frac{1}{1+e^{-z}}$ is the sigmoidal function, $x(k)$ is the input sequence at time k and $\circ$ denotes the Hadamard product.

The proposed architecture for the classifier is depicted in Fig. 1: an RNN with one hidden LSTM layer of dimension $p = 3$, and one sigmoid unit as the output layer (64 parameters in total). This sigmoid unit operates only in the last value of the cell output $h(T)$, once all the sequence $x(k)$ has been processed by the LSTM layer.
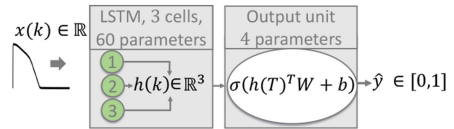

Fig. 1. Network architecture: one hidden LSTM layer ($p = 3$) and a sigmoid unit.

2.2 Semi-supervised Objective Function

We use the binary crossentropy loss $\ell(y, \hat{y}) = -y\log(\hat{y}) - (1-y)\log(1-\hat{y})$ to quantify how close the LSTM prediction $\hat{y} = \sigma(h(T)^T W + b)$ is to label y. More

specifically, given N_a adult APs $\{\mathbf{x}_i^a\}$ and their labels $\{y_i^a\}$, our supervised loss is

$$
\frac{1}{N_a} \sum_{i=1}^{N_a} \{-y_i^a \log(\hat{y}_i^a) - (1 - y_i^a) \log(1 - \hat{y}_i^a)\}. \tag{3}
$$

Now, while we do not have labels for the embryonic APs $\{\mathbf{x}_j^e\}$, we can still use $\ell(\hat{y}_j^e, \hat{y}_{j'}^e)$ to compare the predicted labels for two different embryonic APs. Intuitively, we would like similar APs to have the same labels, and dissimilar APs to have different labels. Therefore, given N_e APs, we use a contrastive loss

$$
\frac{1}{N_e (N_e - 1)} \sum_{j=1}^{N_e} \sum_{j' \neq j} s_{(j,j')} \cdot \ell\left(\hat{y}_j^e, \hat{y}_{j'}^e\right) + \left(1 - s_{(j,j')}\right) \cdot \ell\left((1 - \hat{y}_j^e), \hat{y}_{j'}^e\right), \tag{4}
$$

where $s_{(j,j')}$ represents the similarity between AP $\mathbf{x}_j^e$ and AP $\mathbf{x}_{j'}^e$. We define the similary between two APs based on their distance $d\left(\mathbf{x}_j^e, \mathbf{x}_{j'}^e\right)$ (Euclidean or metamorphosis) as $s_{(j,j')} = \exp\left(-\frac{d^4\left(\mathbf{x}_j, \mathbf{x}_{j'}\right)}{\sigma_s^4}\right)$, where σ_s is chosen as $\sigma_s^4 = \overline{d^4}$, where d is the distance variable and the top bar denotes average operator.

After combining the supervised and unsupervised terms of the loss, we obtain

$$
\frac{1 - \lambda}{N_a} \left(\sum_{j=1}^{N_a} \ell\left(y_j^a, \hat{y}_j^a\right)\right) + \frac{\lambda}{N_e - 1} \sum_{j=2}^{N_e} s_{(j,j-1)} \cdot \ell\left(\hat{y}_j^e, \hat{y}_{j-1}^e\right)
$$
$$
+ \left(1 - s_{(j,j-1)}\right) \cdot \ell\left((1 - \hat{y}_j^e), \hat{y}_{j-1}^e\right), \tag{5}
$$

where λ is a balancing parameter between supervised and unsupervised parts. Instead of making pairwise comparisons between all APs, we propose to compare an AP $\mathbf{x}_j^e$ with the previous one $\mathbf{x}_{j-1}^e$, so fewer distance computations are needed.

2.3   Clustering Quality Index

Since no ground truth labels are available for embryonic APs, the Davies-Bouldin Index (DBI) [19] is considered as a measure of clustering quality. Let $\Omega_0 = \left\{\mathbf{x}_j^e \mid \hat{y}_j^e < 0.5\right\}$ and $\Omega_1 = \left\{\mathbf{x}_j^e \mid \hat{y}_j^e \geq 0.5\right\}$ be the sets containing the different clusters, let S_y be the mean distance from elements of class y to the average signal of the same class, $\mu_y(k) = \frac{1}{|\Omega_y|} \sum_{\mathbf{x}_j^e \in \Omega_y} x_j^e(k)$, and let M_{01} be the distance between the averages $\mu_0(k)$ and $\mu_1(k)$. The DBI is defined as the ratio between the intra-cluster dispersion and the distance between clusters

$$
DBI\left(\Omega_0, \Omega_1\right) = \frac{S_0 + S_1}{M_{01}}, \tag{6}
$$

and should be as small as possible. For computational reasons, and since the Euclidean distance d_E is a good approximation of the metamorphosis distance d_M for small distances, the intra-cluster dispersions S_0 and S_1 are computed using d_E, whereas the distance between clusters M_{01} is computed using d_M.

3   Experiments

3.1   Adult CMs APs Data

A population of 2000 synthetic adult APs was generated by using computational models. The O'hara-Rudy model (ORd) [18] and the Nygren model [17] were paced at 1.5 Hz with 1000 random sets of parameters each (varying between 80% and 120% of their nominal values) to generate ventricular and atrial mature CMs APs, respectively. The parameters varied were the maximum conductances and permeabilities of ion channels (g_{Na}, g_{NaL}, g_{t_0}, g_{Kr}, g_{Ks}, g_{K1}, g_{NC_X}, g_{Kb}, g_{pCa}, P_{Ca}, P_{NaK}, P_{Nab}, P_{Cab} in ORd model, and g_{CaL}, g_{Ks}, g_{Kr}, g_{K1}, g_{Nab} and g_{Cab} in Nygren model). Normalization was applied to each AP so that its maximum voltage and resting membrane potential are 1 and 0, respectively. The Sparse Modeling for Representatives Selection (SMRS) method [20] was then applied to select a subset of $N_a = 300$ templates shown in Fig. 2a.
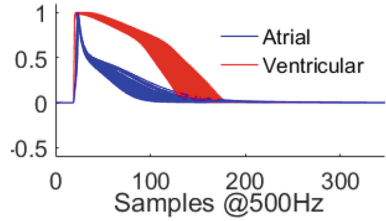
(a) APs from adult CMs
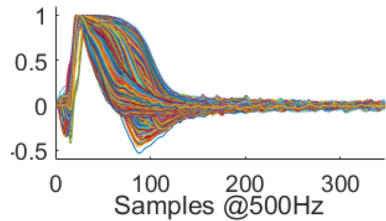
(b) APs from hESC-CMs

Fig. 2. Action potentials: (a) 300 adult CMs, (b) 6940 hESC-CMs.

3.2   hESC-CMs Data

A population of $N_e = 6940$ hESC-CMs APs obtained from 9 cell aggregates paced at 1.5 Hz and optically mapped at a sampling rate of 500 Hz was obtained in [21]. The APs were averaged over beating cycles, a 5×5 boxcar spatial filter was applied for denoising, and then they were normalized (see Fig. 2b). Only 1600 APs (fixed and coming from 2 cell aggregates) were used for training, but labels were predicted for the whole dataset.

3.3   Implementation Details

The classifier architecture was implemented in Keras [22] with TensorFlow backend and trained using the RMSProp optimizer (learning rate 0.003) using batches of 19 APs (3 adult and 16 embryonic). 90 batches were used for training and 10 batches for validation, completing $N_a = 300$ adult APs and $N_e = 1600$ embryonic APs in total. The metamorphosis parameter was set as $\sigma_M = 0.3$.

Three cases are studied: Supervised learning $\lambda = 0$ (Sup-LSTM), Semi-supervised learning $\lambda = 0.1$ with Euclidean distances (Semi-LSTM-E), and Semi-supervised learning $\lambda = 0.1$ with metamorphosis distances (Semi-LSTM-M). In

each case the network was trained 5 times (100 epochs for the Sup-LSTM case and 200 epochs for the Semi-LSTM cases), and the average of the classification results at the last epoch is analyzed.

3.4   Experimental Results

The average classification results generated by the RNN LSTM in the studied cases are shown in Fig. 3 for the 9 cell aggregates. In all cases the proposed classifier generates smooth classification regions and suggests heterogeneity in most of the cell aggregates, which coincides with previous findings [11,21]. Observe that the classification result produced by semi-supervised learning is significantly different from the one produced by supervised learning, with the former being significantly better in terms of DBI. This emphasizes that adult and embryonic APs intrinsically belong to different domains, and therefore classifying embryonic APs with a network trained only with adult APs is not adequate.
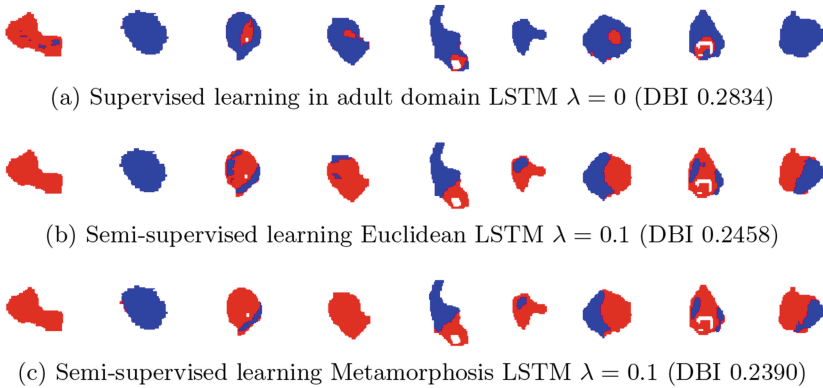

(a) Supervised learning in adult domain LSTM $\lambda = 0$ (DBI 0.2834)

(b) Semi-supervised learning Euclidean LSTM $\lambda = 0.1$ (DBI 0.2458)

(c) Semi-supervised learning Metamorphosis LSTM $\lambda = 0.1$ (DBI 0.2390)

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 compares our results to those of the method presented in [11] (1NN classifier with $N_a = 20$ synthetic adult AP templates). Observe that supervised learning shows significantly higher DBI than the rest, which is expected since it does not consider hESC-CMs data during training. On the other hand, the semi-supervised learning scheme outperforms the 1NN scheme when Euclidean distances are used (DBI 0.2458 vs 0.2558). 1NN with Euclidean distances was replicated with the same 300 adult AP templates used to train the network (see Table 1), confirming that the improvement in clustering quality observed in the semi-supervised scheme is not attributable to the number of templates used, but to the method itself: Euclidean metric is a good approximation of metamorphosis for small distances, so it performs better when distances within hESC-CMs domain are computed (proposed semi-supervised framework) than when distances between hESC-CMs and adult CMs domains are computed (1NN).

Table 1. Comparing the results of the proposed method (LSTM) with the results presented in [11]. Accuracy* is computed assuming 1NN classification with metamorphosis distance as ground truth (E: Euclidean, M: Metamorphosis).

Method	1NN	1NN	1NN	Sup-LSTM	Semi-LSTM	Semi-LSTM
Templates	20 [11]	20 [11]	300 SMRS	300 SMRS	300 SMRS	300 SMRS
Metric	M	E	E		E	M
DBI	0.2297	0.2558	0.2566	0.2834	0.2458	0.2390
Accuracy*	1	0.6488	0.6290	0.4723	0.8788	0.9473

1NN metamorphosis results presented in [11] show the best clustering quality (DBI 0.2297), followed by the Semi-LSTM-M (DBI 0.2390). The classification accuracy assuming 1NN metamorphosis as the ground truth was computed and plotted vs the DBI in Fig. 4. The use of metamorphosis distance in semi-supervised learning not only produces lower DBI but also consistently generates bet-

Fig. 4. Accuracy* vs DBI. 1NN M as ground truth (E: Euclidean, M: Metamorphosis).

ter classification accuracy than when the Euclidean distance is used (small dots in Fig. 4 represent single trials results, and squares represent the average classification per case). An improvement of 24.98% in the classification accuracy is observed between 1NN and the semi-supervised learning scheme when 300 templates and only Euclidean distances are used, achieving 87.88% accuracy without any metamorphosis distance computation.
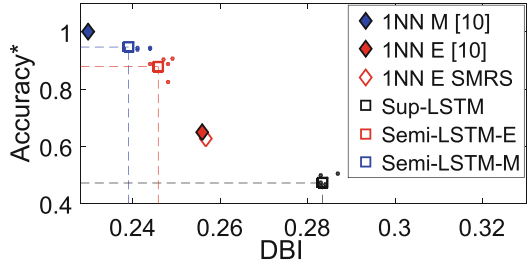
4   Conclusion

The proposed method not only successfully integrates labeled data from a different domain to solve the task, but also proves to be a powerful framework to improve the performance of Euclidean-based methods in the classification of hESC-CMs APs. Moreover, it reaches 94.73% of agreement with the state-of-the-art, trading off accuracy with computational complexity: whereas the classification of a new sample in state-of-the-art method requires to solve 20 computationally intensive optimization problems (6.74 sec/sample in 2 8-core computer nodes with 8 2.3 GHz CPUs per node [11]), in the proposed method it just needs to be processed by a small RNN with fixed weights (less than 6 sec for the whole 6940 APs dataset in one 2.2 GHz CPU with 2 cores, 4 threads).

References

1. Kehat, I., et al.: Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. J. Clin. Investig. **108**(3), 407–414 (2001)
2. Turksen, K.: Human embryonic stem cell protocols. vol. 331. Springer (2006)
3. Zhu, W.Z., Hauch, K.D., Xu, C., Laflamme, M.A.: Human embryonic stem cells and cardiac repair. Transp. Rev. **23**(1), 53–68 (2009)
4. Braam, S.R.: Prediction of drug-induced cardiotoxicity using human embryonic stem cell-derived cardiomyocytes. Stem Cell Res. **4**(2), 107–116 (2010)
5. Sartiani, L.: Developmental changes in cardiomyocytes differentiated from human embryonic stem cells: a molecular and electrophysiological approach. Stem Cells **25**(5), 1136–1144 (2007)
6. Peng, S., Lacerda, A., Kirsch, G., Brown, A., Bruening-Wright, A.: The action potential and comparative pharmacology of stem cell-derived human cardiomyocytes. J. Pharmacol. Toxicol. Methods **61**(3), 277–286 (2010)
7. 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. Circul. Res. **93**(1), 32–39 (2003)
8. Gorospe, G., Zhu, R., Millrod, M., Zambidis, E., Tung, L., Vidal, R.: Automated grouping of action potentials of human embryonic stem cell-derived cardiomyocytes. IEEE Trans. Biomed. Eng. **61**(9), 2389–2395 (2014)
9. Trouvé, A., Younes, L.: Metamorphoses through Lie group action. Found. Comput. Math. **5**(2), 173–198 (2005)
10. Gorospe, G., Younes, L., Tung, L., Vidal, R.: A metamorphosis distance for embryonic cardiac action potential interpolation and classification. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013. LNCS, vol. 8149, pp. 469–476. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40811-3_59
11. Gorospe, G., et al.: Efficient metamorphosis computation for classifying embryonic cardiac action potentials. In: 5th Workshop on Mathematical Foundations of Computational Anatomy (2015)
12. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
13. Graves, A., Jaitly, N.: Towards end-to-end speech recognition with recurrent neural networks. In: International Conference on Machine Learning, pp. 1764–1772 (2014)
14. Donahue, J., et al.: Long-term recurrent convolutional networks for visual recognition and description. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2625–2634 (2015)
15. Lipton, Z.C., Kale, D.C., Elkan, C., Wetzel, R.: Learning to diagnose with LSTM recurrent neural networks. arXiv:1511.03677 (2015)
16. Zihlmann, M., Perekrestenko, D., Tschannen, M.: Convolutional recurrent neural networks for electrocardiogram classification. arXiv:1710.06122 (2017)
17. Nygren, A., et al.: Mathematical model of an adult human atrial cell: the role of k+ currents in repolarization. Circul. Res. **82**(1), 63–81 (1998)

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 Comput. Biol. **7**(5), e1002061 (2011)
19. Davies, D.L., Bouldin, D.W.: A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. **2**, 224–227 (1979)
20. Elhamifar, E., Sapiro, G., Vidal, R.: See all by looking at a few: Sparse modeling for finding representative objects. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1600–1607 (2012)
21. 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. Sci. Rep. **6**, 18544 (2016)
22. Chollet, F., et al.: Keras (2015). https://keras.io