Morphology And Status Transition Monitoring For Mouse Embryonic Stem Cell Colonies In Vitro By LSTM Networks With Progressive Training Using Fluorescence Microscopy Images

Slo-Li Chu  
Chung Yuan Christian University

Kuniya Abe  
RIKEN: Rikagaku Kenkyujo

Hideo Yokota  
RIKEN Center for Advanced Photonics: Rikagaku Kenkyujo Koryoshi Kogaku Kenkyu Center

Ming-Dar Tsai  (mingdar@cycu.edu.tw)  
Chung Yuan Christian University  https://orcid.org/0000-0002-4171-0450

Research Article

Keywords: Recurrent Neural Networks, LSTM, mouse embryonic stem cells, colony morphology and status transition monitoring, cell tracking

DOI: https://doi.org/10.21203/rs.3.rs-739229/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Morphology and Status Transition Monitoring for Mouse Embryonic Stem Cell Colonies in vitro by LSTM Networks with Progressive Training Using Fluorescence Microscopy Images

Slo-Li Chu¹, Kuniya Abe², Hideo Yokota³ and Ming-Dar Tsai*¹

¹ Department of Information & Computer Engineering, Chung Yuan Christian University, Taiwan

² RIKEN BioResource Research Center, Tsukuba, Japan

³ Center for Advanced Photonics, RIKEN, Japan

Reprint requests and correspondence to:
Ming-Dar Tsai  Professor
Institute of Information and Computer Engineering, Chung Yuan Christian University, Taiwan

TEL: (+886) 3-256-4718  FAX: (+886) 3-256-4799  E-MAIL: mingdar@cycu.edu.tw
Abstract

Purpose  Embryonic stem (ES) cells represent as a cellular resource for basic biological studies and for their uses as medically relevant cells in *in vitro* studies. Fluorescence microscopy images taken during cell culture are frequently used to manually monitor time-series morphology changes and status transitions of ES cell (ESC) colonies, and to study dynamical pattern formation and heterogeneity distribution within ESC colonies, intrinsic fluctuation and cell-cell cooperativity. Therefore, tracking and furthermore predicting morphology changes and status transitions of ESC colonies is an effective method to monitor culture medium for maintaining ES cells in undifferentiated or early differentiated stage.

Methods  A P-LSTM (Progressive Long Short-Term Memory) structure is proposed to incorporate some new time-lapse images real-time taken from incubators for a new RNN (Recurrent Neural Networks) training. The P-LSTM can achieve adaptive long- and short- term memories to generate accurate predicted images. On the time-lapse images, entropy and bi-lateral filtering are used to extract the range of every colony to calculate colony morphology. Colony status transitions between consecutive images are calculated by mapping the calculated colony centers and ranges.

Results  Accuracies for the colony status transition, area and roundness for the 15 predicted (five-hour) future frames calculated from 1500-2500 colonies for respective frames show the effectiveness of the proposed method.

Conclusion  We proposed an efficient and automatic method to predict and monitor status transitions and morphology changes of mouse ESC colonies in culture using time-lapse fluorescence microscopy images.

Keyword: Recurrent Neural Networks, LSTM, mouse embryonic stem cells, colony morphology and status transition monitoring, cell tracking,
1. Introduction

Embryonic stem cells (ESCs) are gaining attention as a cellular resource for basic biological studies and for their uses as medically relevant cells in *in vitro* studies. ESCs are known as undifferentiated and self-renewing stem cells. Upon induction of differentiation, they are going to differentiate into multiple cell types of distinct cell lineages. ESCs in a colony are not a group of homogenous cells, in which cells with different cell differentiation status co-exist and shift between the statuses. Although the nature of such heterogeneities of ESC population remains elusive, different types and levels of heterogeneity were revealed under different culture conditions [1, 2]. ESC colonies that have an internal structure with respect to cell size, marker expression or biomechanical properties were usually calculated using microscopy images. For example, heterogeneities were represented by entropy [3] and differentiation was by the colony shape [4] on the microscopy images.

Monitoring of stem cell pluripotency and differentiation has been considered as important, where spectroscopy or electrochemical sensing [5], or fluorescence markers [6] were used. Time-lapse fluorescence microscopy images taken during culture were frequently used to track and thus monitor morphology changes (entropy, area, roundness, etc.) and status transitions (moving, merging, splitting, etc.) of an ESC colony or between colonies [7, 8]. The tracking results assist study of time-development of the cell-state transition of undifferentiated or early differentiated ES cells in colony-levels. For example, the morphology change of a mouse ESC colony could be used to calculate dynamical pattern formation and heterogeneity distribution at the colony, thus reveal the interaction of both stochastics and deterministic components, namely, intrinsic fluctuation and cell-cell cooperativity [9].

Recently, a machine learning method, RNN (Recurrent Neural Networks) learns sequential data of recurrent hidden (or internal) states at each time frame (such as consecutive time-lapse images) to
predict future time-series states for efficiently monitoring, such as for tool wearing [10] or for early characterization of drug-induced cardiotoxicity [11]. Many RNN structures have been proposed to improve the predicted accuracy and speed or to accommodate the application. For example, LSTM (Long Short-Term Memory) networks [12] were proposed to solve vanishing and exploding gradient problems in capturing long-term dependencies in classical RNNs [13]. LSTM achieved successes in many fields and applications such as handwriting recognition [14] and medical care predictions [15]. In some applications, not the whole images, but simplified pixels representing probability vectors or position vectors (classified by CNNs (convolution neural networks)) were used as LSTM inputs to predict probability of future hiPS reprogramming [16] or cell motility [17].

The simplified pixels or other signals such as spectroscopy or electrochemical sensing, or fluorescence markers; however, cannot represent morphological features or patterns of cell colonies at the original fluorescence microscopy images. This study, therefore, processes whole time-lapse fluorescence images to predict status transitions and shape changes of ESC colonies. However, the morphology usually changes apparently and the status transits abruptly (finishes the status transition in a few consecutive frames of time-lapse images). We propose a progressive LSTM (P-LSTM) structure that incorporates few new images and excludes the same number of the earliest images for a new repetition of training to deal with the colony short-term changes on transitions. A prediction can be made after any new training. However, too many repetitions cause retained long-term memory that brings image retention or disturbance in predicted images. Therefore, a P-LSTM network re-starts after certain repetitions and multiple P-LSTM networks are implemented concurrently to keep there is a P-LSTM network with optimal times of repetitions to solve retained long-term-memory problems.

The P-LSTM thus includes adaptive long- and short-term memories to generate predicted images with accurate status transitions and morphology changes of respective ESC colonies. Cell colony tracking is then implemented to calculate the colony morphology changes and status transitions for
monitoring future behavior of these colonies. The P-LSTM network structure is introduced in Section 2.2. Section 2.3 introduce the method of evaluating morphology of each ES colony on the predicted and their corresponding real time-lapse images. Section 2.4 introduces the method of tracking status transitions of ES colonies based on the convex-hull colonies on consecutive real and predicted images. The calculated morphology and status transitions for ES colonies on real time-lapse images are used to evaluate the accuracy of the morphology changes and status transitions on the predicted images.

2. Material and Methods

2.1 fluorescence microscopy images of mouse ES cell cultures

Time-lapse fluorescence microscopy images of mES cells were acquired by using a confocal microscope, CV1000-EH with built-in incubator (Yokogawa Electric Corp., Tokyo, Japan). The time-lapse images with 20 min interval were taken by the equipped EMCCD camera (512×512 pixels, 16 μm pixel width). The mES cells harboring the Venus (green) fluorescence protein gene driven by Mvh gene promoter and mRFP (monomeric red fluorescent protein) gene fused with human histone H2B gene driven by a ubiquitous Pgk1 promoter. Mvh-Venus gene showed heterogeneous expression among the cells in the colonies, while the H2B-mRFP exhibited ubiquitous expression in all the cells.

2.2 Progressive LSTM network for generating predicted fluorescence microscopy images

2.2.1 Progressive LSTM network

Figure 1 shows the three stages of P-LSTM (progressive-LSTM) network, progressive sampling, LSTM training stages and LSTM predicting stage. In progressive sampling stage, the P-LSTM network incorporates S sliding images to improve short-term memory in the LSTM structure. The first training can begin at any (t-th) frame using a number (D) of consecutive \{x_t, ..., x_{t+D-1}\} time-lapse microscopy images that are stored in the Current Input Buffer. \(x^{t+i}\) indicates the \(t+i\)-th time-
step image. Because $S$ images are added and the same number of the earliest images are deleted for a new repetition of training and prediction, the images in the Current Input Buffer are the ones of consecutive time steps $\{x^{t+D+S\times R}, ..., x^{t+D+S\times R-1}\}$ at the $R$ times of further repetitions. These images are then used to generate the same $b$ input image sets for the LSTM batch training. $R$ is 0 for the first training.

In the LSTM training stages, an image set $X_i$ of a constant size $w$ of consecutive images using a random number $\sigma$ as the first frame of image is selected from the Current Input Buffer to form an image set $X_i$. For example, at the last repetition, $X_i = \{x^{t+S\times R+\sigma}, ..., x^{t+S\times R+\sigma+w-1}\}$. $\sigma+w$ is not over $D$. The above selecting procedure are repeated to obtain $b$ image sets ($X_1, ..., X_b$) that are used to train $D$ consecutive LSTM units with a certain number of training steps. In the LSTM units, every pixel with three colors of an image is not simplified and trained independently.

In the LSTM predicting stage, parameters in the LSTM units are becoming optimal after training to make a prediction for every image with three color channels. Thus, prediction can be made after the first or every repetition of progressive training. $Y_0 = \{y_0^1, y_0^2, ..., y_0^P\}$ is the set of $P$ consecutive images predicted after the first training, actually predicts for the future $t+D$-th until $t+D+P$-1-th frames of images. Similarly, $Y_1 = \{y_1^1, y_1^2, ..., y_1^P\}$ is the set of the consecutive predicted images for the $t+S+D$-th until the $t+S+D+P$-1-th frames of time-lapse images after the first time of repetition (progressing training). After $R$ repetitions of progressive training and prediction, $P$ consecutive images $Y_R = \{y_R^1, y_R^2, ..., y_R^P\}$ can be predicted for the $t+S\times R+D$-th until the $t+S\times R+D+P$-1-th frames of images.

The values of $D$, $w$, and $b$ are set as 15, 6, and 3 according to our experiences. The training steps are 160. Larger values take more training time but could not improve prediction accuracy apparently. Meanwhile, the fewer sliding images $S$ the better effect of LSTM short-term memory but take more computation. $S$ is currently set as 3 to achieve high prediction accuracy and quick computation.
2.2.2 Concurrent progressive LSTM network

We implement two P-LSTM networks concurrently with the interval of 5 repetitions to keep there exist P-LSTM networks with further 5-9 times of repetitions after the first training. The prediction from any of these repetitions was tested as optimal to solve retained long-term-memory problems and possess sufficient long-term memory to obtain accurate predicted images. A P-LSTM network should be restart from the first training if the repetition of progressive training already reached nine.

For example, one P-LSTM network begins its first training using the first fifteen frames of time-lapse images taken in culture, and consecutively incorporates every three new frames into the progressive training until nine repetitions (from the 6-th until 42-nd frames) were implemented. The other P-LSTM network begins the first training using the 16-th until 30-th frames of time-lapse images, and consecutively incorporates every three new frames into the progressive training until the nine repetitions (31-st until 57-th frames) were implemented. For the next run of the first P-LSTM network, the 31-st until 45-th frames of time-lapse images are used for the first training and the 46-th until 72-th frames of time-lapse images are used for nine repetitions. For the next run of the other P-LSTM network, the 46-th until 60-th frames of time-lapse images are used for the first training and the 61-st until 87-th frames of time-lapse images are used until the nine repetitions. Prediction is possible after the first or any repetition of training, and the prediction with any of optimal 5-9 times of repetitions is possible after two concurrent P-LSTM networks were trained (time-lapse images after the 30-th frame).

2.3 Morphology calculation for mESC colonies on fluorescence microscopy image

2.3.1 Morphology calculation for colonies on real fluorescence microscopy image
Figure 2 shows the colony region extraction for a real fluorescence microscopy image (Fig. 2(a)) taken from the incubator. For imaging live cells, the fluorescence image has relatively low signal-to-noise ratio. The image is converted as a gray-level image to calculate the entropy by the following equation [18].

\[
H(x,y) = -\sum_{i=-n}^{n} \sum_{j=-n}^{n} P(v((x+i),(y+j))) \cdot \log_2 P(v((x+i),(y+j)))
\] (1)

\(H(x,y)\) indicates the entropy of target pixel (with the position of \(x,y\)) on an 8-bit greyscale image. For each pixel value ranging from 0 to 255, the intensity of all the pixels was statistically independent, thus the intensity occurrence probability \(P(v((x+i),(y+j)))\) indicates the probability of the target pixel value \(v(x,y)\) appearing in a \(n\) (currently set as 8) square sub-image (17x17 pixels) surrounding the target pixel. The entropy image indicates the disorder of pixel values as shown in Fig. 2(b), where a pixel with high value indicates appearance of colony boundary.

Then, a binarization followed by a convex hull calculation ([19]) are implemented to fill holes (noises) inside every cell region, smooth the boundary, and label every obtained convex-hull boundary as shown in Fig. 2(c). The small regions (under a constant area) are considered as noises and not labeled. After the convex hull computation, the morphological features such as the circumcircle, area, and roundness for every labelled convex-hull colony are calculated and recorded. Fig. 2(c) shows background noises are well deleted and shapes of the labeled colonies are similar to their corresponding colonies on the real image (Fig. 2(a)) except the one during status transiting (as the merging colony labeled as “1” in the right-bottom quadrant).

2.3.2 Morphology calculation for colonies on predicted fluorescence microscopy image

Figure 3 shows the colony region extraction for an image predicted by the P-LSTM network (Fig. 3(a)) corresponding to the real image (Fig. 2(a)) it predicts. The images after the entropy and
the followed convex-hull computation are Fig. 3(b) and Fig. (c), respectively. Because of heterogeneous inherence of the predicted image (Fig. 3(a)), the obtained convex-hull colonies are generally larger than the ones on the original predicted image (Fig. 3(c)). Therefore, we binarize the predicted image (Fig. 3(d)), then the bi-later filtering [20] to delete the noises and finally the convex-hull algorithm to obtain the convex-hull colonies as shown in Fig. 3(e).

These convex-hull colonies are generally similar than to the ones on the original image (Fig. 3(a)). However, some small colonies become too small by the bi-later filtering to be dealt as noises and deleted. For example, the colony at the right-top part is too small to be judged as noise, and the area in-between the opposite parts of a merging colony is deleted so that the merged colony became two (Fig. 3(c)). Therefore, we use the colony centers to check if there are not labeled or multiply labeled colonies by the bi-lateral filtering but merged (as “1” colony in Fig. 3(c)) or labelled (as “2” and “3” colonies in Fig. 3(c)) by the entropy processing. These colonies are then used to replace the ones on the bi-later filtered images as shown in Fig. 3(f).

2.4 Status transition tracking for mouse ESC colonies using convex-hull colonies on consecutive time-lapse images

Two mappings are implemented to track the status transition of every mouse ESC colony between two (previous and current) real or predicted consecutive images. The recorded colony centers at the previous frame image are mapped to the current frame image at the first mapping, and the colony centers at the current frame are mapped onto the previous frame at the second mapping. The status transition of every colony is then determined as either of the following cases.

**Moving**: one-to-one mapping. A colony region on the current image is mapped by only one center of a colony on the previous image, and conversely this colony region on the previous image is also mapped by only one colony center of the current image. This case is defined as no status transition.
Moving in: one case of Moving where the colony center of the previous image is located near the image edge but the colony center of the current image was not. This case is defined as in a status transition.

Moving out: one case of Moving where the colony center of the current image is located near the image edge but the colony center of the previous image was not. This case is also defined as in a status transition.

Disappearance: none mapping. It is a colony region on the previous image that was mapped by no colony centers of the current image. This case is in a status transition.

Appearance: none mapped. It is a colony on the current image that was mapped by no colony centers of the previous image. This case is in a status transition.

Merging: Multiple to one mapping. Two or more centers of the previous image were mapped into the region of the same colony on the current image. This case is a status transition.

Splitting: One mapped by multiple. Two or more centers of the current image were mapped into the region of the same colony on the previous image. This case is also a status transition.

Disappearance not followed by a Moving out, or Appearance not followed by a Moving in is considered as an error caused the area of a convex-hull colony is near the threshold for dealing with this colony as a noise. Such Disappearance or Appearance occurs when a small colony is judged as noise and deleted but not in its neighboring frames.

3. Results

3.1 Real-time evaluation for P-LSTM structure

We developed the system on the PCs with Intel Core i9-9900K 3.60 GHz, 64GB RAM and graphics card of Asus RTX2080Ti (11G) and the software Tensorflow/Keras framework, Python programming language and the OpenCV library. The computation for calculating morphology and status transitions
of ES colonies takes much little time comparing to the P-LSTM training and prediction. Because the training and prediction is implemented pixel by pixel, independently, ideal parallel processes for divided subimages can be implemented for multi-core PCs. Eight processes (the most time-saving for Intel Core i9-9900K) can reduce more than half of the execution time than single process and achieved the following execution time.

Currently, the first or a progressive takes almost the same time (2240 seconds), and a prediction takes 455 seconds including generation of 15 prediction images and their processing (bilateral filtering, entropy computation, and colony transition and morphology computation). Meanwhile, generating more (than 15) frames for predicted microscopy image and processing these images takes only little more time. Therefore, it takes about 45 min to complete one repetition of P-LSTM training and prediction to monitor the future status transitions and morphology changes of mouse ESC colonies in culture. Real-time P-LSTM training and prediction of one repetition can be completed before incorporating the 3 images (1 h) newly generated by the incubator.

3.2 Method evaluation by comparing predicted accuracies colony status transition and morphology

We have implemented 110 times of real-time P-LSTM training and prediction for respective 5 sets of time-lapse images taken in mouse ESC culture. Totally, 550 times of P-LSTM training and prediction were implemented. Each prediction generated 15 frames of predicted images. Table I shows, at each predicted frame, the mean and standard deviation of the accuracies for the colony status transition, area and roundness (the ratio of the inscribed radius to circumradius) calculated the colonies on the 550 predicted images comparing to their corresponding the colonies on the real images. A predicted frame includes totally 1500-2500 colonies from the 550 predictions. Meanwhile, a colony predicted as **Moving** (no status transition) is counted as correct only its corresponding colony at the corresponding frame calculated from real time-lapse images is also **Moving**. Meanwhile, a colony
predicted as a status transition (Merging, Splitting, Appearance, Disappearance, Merging in, and Merging out) is counted as correct if either of its corresponding colony at the corresponding frame together with its previous or next three frames (currently as ±3) is the same type of status transition calculated from real time-lapse images. Figure 4 shows the accuracies for the status transition, area and roundness are the later predicted frames the lower, but all acceptable in the 15 frames (about 75% accuracy for status transition and 65% for area prediction, and over 80% for shape (roundness)). The result shows the P-LSTM structure helps monitoring mouse ESC colony status transition and shape changes in culture for five-hour future.

Figure 5 shows an example with a prediction for colony Merging. Figure 5(a), (b), (c) and (d) show selected frames of original real images, labelled convex-hull colonies (with circumcircles) on the real images, predicted images, and labelled convex-hull colonies (with circumcircles) on the predicted images, respectively. The convex-hull colonies shown in Fig. 5(b) are similar to the ones of the corresponding real images (Fig. 5(a)), thus can be used for evaluating the status transitions and shapes of the colonies on the real images. The convex-hull colonies as shown in Fig. 5(d) are compared with the ones in Fig. 5(b) to calculate the accuracies of predicted colony status transitions, shapes and areas. The system monitored (predicted) a Merging at the seventh frame (near the middle of the image) as shown in Figure 5(d) that really happen at the fourth frame after the prediction as shown in Fig. 5(b). Other colonies in this 15-frame prediction are all Moving and correctly predicted.

Figure 6 shows an example with a Moving in status transition. Figure 6(a), (b), (c) and (d) show selected frames of the real images, the labelled convex-hull colonies with their circumcircles on the real images, predicted images, and the labelled convex-hull colonies with their circumcircles on the predicted images, respectively. The convex-hull colonies shown in Fig. 6(b) are similar to the corresponding real fluorescence images (Fig. 6(a)). The system monitored (predicted) a Moving in at the sixth frame as shown in Figure 6(d), that really happen at the eighth frame after the prediction.
as shown in Fig.6(b). Therefore, the predicted **Moving in** is two-frames ahead of the real images, thus is considered as correct.

Figure 7 shows an example with an **Appearance** status transition. Figure 7(a) show three continuous frames of predicted images, and Figure 7(b) shows their labelled convex-hull colonies circumcircles calculated from the three respective predicted images. A convex-hull colony at the right-button part of the image shown on the second frame is recognized as a colony but not recognized on its neighboring (previous and next) frames because of small area. Therefore, it is a considered as an **Appearance** error caused by the image processing when determining the labelled colonies. Such **Appearance** and **Disappearance** errors did not appear in the convex-hull colonies processed from the real fluorescence images, but were about 0.25% of all convex-hull colonies processed from the predicted fluorescence images.

### 4. Conclusion and discussion

In this paper, we proposed an efficient method for monitoring status transitions and morphology changes of mouse ESC colonies in culture using time-lapse fluorescence images real-time taken from an incubator. To accurately learn apparent short-term changes of mouse ESC colonies on time-lapse fluorescence images, a new machine learning method, progressive LSTM structure was proposed that incorporates several non-simplified time-lapse fluorescence images to implement a prediction of new training. Meanwhile, the LSTM structure restarts to avoid retained long-term-memory problems. Besides, we developed entropy and bi-lateral filtering methods to extract the range and calculate morphology of every cell colony on a predicted image, and to track colony transitions using the colony centers on consecutive predicted images. The tested accuracy shows the prediction for (15 frames) five-hour future are reliable by comparing the original predicted and real fluorescence images, and cell colony morphology and status transition calculated based on extracted convex-hull cell colonies from the predicted and real fluorescence images.
The results show our new machine learning method is effective in generating predicted fluorescence images. However, the prediction for the frames over 15 is usually not precise according to our tests (not shown in the paper). The reason can be considered as no further new real images were learned. Besides, in some cases especially in status transitions, shapes of convex-hull colonies may be not similar to their corresponding the mouse ESC colonies on the real or predicted images. The method for calculating the colony morphology and centers should be improved to acutely evaluate morphology changes and status transitions.

Funding

Research supported by Ministry of Science and Technology (MOST), R.O.C., under Grant No. 108-2923-E-033-002-MY2, 109-2221-E-033-034, and by RIKEN Presidential Fund: Technology development for quantitative evaluation of cellular state based on image processing and machine learning.

Author contributions

**S.L. Chu:** Software conceptualization and development, Formal analysis, Investigation, Methodology, Validation, **K. Abe:** Conceptualization, Data curation, Investigation, Methodology, Validation, Writing-original draft. **H. Yokota:** Software confirmation, Formal analysis, Investigation, Methodology. **M.D. Tsai:** Conceptualization, Investigation, Project administration, Software supervision, Writing-original draft, Writing -review & editing.

Declaration of Competing Interest

The authors declare no competing financial or personal relationships that could be viewed as influencing the work reported in this paper.

References
1. Tamm, C., Galito´, S. P., & Annere´n, C. (2013). A comparative study of protocols for mouse embryonic stem cell culturing. *PLOS One, 8*, e81156.

2. Canham, M. A., Sharov, A. A., Ko, M.S.H. & Brickman, J. M. (2010), Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcrip. *PLoS Biology, 8*, e1000379.

3. Herberg, M., Zerjatke, T., Back, W., Glauche, I., & Roeder, I. (2015). Image-based quantification and mathematical modeling of spatial heterogeneity in ESC colonies. *Cyrometry Part A, 87*(6), 481-490.

4. Rosowski, K. A., Mertz, A. F., Norcross, S., Dufresne, E. R., & Horsley. V. (2015). Edges of human embryonic stem cell colonies display distinct mechanical properties and differentiation potential. *Sci. Rep. 5*: 14218.

5. Suhito, I. R., Angeline, N., Choo, S. S., Woo, H. Y., Paik, T., Lee, T., & Kim, T. H. (2018). Nanobiosensing platforms for real-time and non-invasive monitoring of stem cell pluripotency and differentiation. *Sensors, 18*, 2755.

6. Pascut, F. C., Goh, H.T., Welch, N., Buttery, L. D., Denning, C., & Notingher, I. (2011). Noninvasive detection and imaging of molecular markers in live cardiomyocytes derived from human embryonic stem cells. *Biophysical Journal, 100*(1), 251–259.

7. Scherf, N., Herberg, M., Thierbach, K., Zerjatke, T., Kalkan, T., Humphreys, P., Smith, A., Glauche, I., & Roeder I. (2012). Imaging, quantification and visualization of spatio-temporal patterning in mESC colonies under different culture conditions. *Bioinformatics, 28*, i556-i561, 2012.

8. Chang, Y. H., Yokota, H., Abe, K., Tang C. T. & Tasi, M. D. (2017). Automated detection and tracking of cell clusters in time-lapse fluorescence microscopy images. *J. Med. Biol. Eng. 37*, 18-25.
9. Okamoto, K., Germond, A., Fujita, H., Furusawa, C., Okada, Y. & Watanabe, T. M. (2018). Single cell analysis reveals a biophysical aspect of collective cell-state transition in embryonic stem cell differentiation, *Sci. Rep.* 8, 11965.

10. Zhao, R., Yan, R., Wang, J. & Mao, K. (2017). Learning to monitor machine health with convolutional bi-directional LSTM networks, *Sensors 2017, 17, 273.*

11. Golgooni, Z., Mirsadeghi, S., Baghshah, M. S., Ataee, P., Baharvand, H., Pahlavan S., & Rabiee, H.R. (2019). Deep learning based proarrhythmia analysis using field potentials recorded from human pluripotent stem cells derived cardiomyocytes. *Cardiovascular Devices and Systems, 7, 1900203.*

12. Gers, F.A., Schmidhuber, J. A., & Cummins, F.A. (2000). Learning to Forget: Continual Prediction with LSTM. *Neural Computation, 12, 2451-2471.*

13. Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 6(2), 107-116.*

14. Graves, A., Liwicki, M., Fernández, S., Bertolami, R., Bunke, H., & Schmidhuber, J. (2009). A Novel connectionist system for unconstrained handwriting recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 31, 855–868.*

15. Choi, E., Bahadori, M. T., Schuetz, A., Stewart, W. F., & Sun J. (2016). Doctor AI: predicting clinical events via recurrent neural networks. *arXiv: 1511.05942.*

16. Chang, Y.H., Abe, K., Yokota, H., Sudo K., Nakamura Y., Chu, S.L., Hsu, C.H., & Tasi, M.D., (2019). Human Induced Pluripotent Stem Cell Reprogramming Prediction in Microscopy Images using LSTM based RNN. *The 41th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2416-2419.*
17. Kimmel, J., Brack, A. & Marshall, W. (2019). Deep convolutional and recurrent neural networks for cell motility discrimination and prediction. *IEEE/ACM transactions on Computational Biology and Bioinformatics, 13*, 600-612.

18. Elad, M. (2002). On the origin of the bilateral filter and ways to improve It. *IEEE Trans Image Process, 11*, 1141–1151.

19. Kim, D., Choi, J., & Nam, J. (2016). Entropy-assisted image segmentation for nano-and micro-sized networks. *Journal of microscopy, 262*(3), 274-294.

20. Liu-Yu, S. & Thonnat, M. (1992). Using apparent boundary and convex hull for the shape characterization of foraminifera images. *Proc. 11th IAPR Int. Conf. on Pattern Recognition. Vol. III, Conference C: Architectures for Vision and Pattern Recognition*. 569-572.
Fig. 1 P-LSTM structure
Fig. 2 Colony labeling and morphology calculation for real fluorescence image. (a) Real fluorescence image, (b) image after entropy computation, and finally by (c) convex-hull computation to obtain colonies; red circles represent circumcircles of labelled colonies, red points are their centers; small regions not in circumcircles are dealt as noises.

Fig. 3 Colony labeling and morphology calculation for the P-LSTM predicted fluorescence image. (a) Predicted image, (b) Image after entropy computation and then (c) convex-hull computation; (d) binarized image and then bi-lateral filtering from predicted image and then (e) convex-hull computation; (f) mapping labeled lost colonies and areas on (c) to (e).
Table 1  prediction accuracy vs 15 predicted frames

| Predicted frame | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Roundness (mean) | 0.867 | 0.860 | 0.855 | 0.851 | 0.845 | 0.842 | 0.843 | 0.837 | 0.835 | 0.836 | 0.830 | 0.827 | 0.829 | 0.829 |
| Roundness (SD)  | 0.867 | 0.860 | 0.855 | 0.851 | 0.845 | 0.842 | 0.843 | 0.837 | 0.835 | 0.836 | 0.830 | 0.827 | 0.829 | 0.829 |
| Status (mean)   | 0.742 | 0.801 | 0.791 | 0.778 | 0.776 | 0.769 | 0.773 | 0.765 | 0.756 | 0.763 | 0.760 | 0.748 | 0.755 | 0.745 |
| Status (SD)     | 0.210 | 0.185 | 0.194 | 0.203 | 0.203 | 0.208 | 0.206 | 0.200 | 0.204 | 0.220 | 0.208 | 0.213 | 0.211 | 0.214 |
| Area (mean)     | 0.727 | 0.712 | 0.707 | 0.703 | 0.695 | 0.691 | 0.688 | 0.685 | 0.677 | 0.678 | 0.673 | 0.665 | 0.666 | 0.658 | 0.643 |
| Area (SD)       | 0.207 | 0.210 | 0.215 | 0.219 | 0.264 | 0.221 | 0.224 | 0.220 | 0.220 | 0.227 | 0.224 | 0.226 | 0.244 | 0.243 | 0.299 |

Fig. 4 Predicted accuracies by P-LSTM for colony shapes, status transitions, and areas.
Fig. 5 Comparison of predicted and real time-lapse images with their labelled convex-hull colonies for a Merging case. (a) Real fluorescence images; (b) labelled convex-hull colonies with their circumcircles (red circles) and centers (red points) on the real images; (c) predicted fluorescence images and (d) labelled convex-hull colonies with their circumcircles and centers on the predicted images.
Fig. 6  Comparison of predicted and real time-lapse images with the labelled convex-hull colonies for a **Moving-in** case. (a) labelled convex-hull colonies with their circumcircles (red circles) and centers (red points) on the real images; (c) predicted fluorescence images and (d) labelled convex-hull colonies with their circumcircles and centers on the predicted images.
Fig. 7. Consecutive frames of predicted mESC images with the corresponding labelled convex-hull colonies for an Appearance case. (a) Consecutive predicted fluorescence images and (b) labelled convex-hull colonies with circumcircles (red circles) and centers (red points).