Computerized Medical Imaging and Graphics 60 (2017) 22–34

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

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

A Hierarchical Convolutional Neural Network for vesicle fusion event classification

Haohan Li[a,1], Yunxiang Mao[a,1], Zhaozheng Yin[a,*,1], Yingke Xu[b,2]

[a] Department of Computer Science, Missouri University of Science and Technology, Rolla 65409, USA
[b] Department of Biomedical Engineering, MOE Key Laboratory of Biomedical Engineering, Zhejiang Provincial Key Laboratory of Cardio-Cerebral Vascular Detection Technology and Medicinal Effectiveness Appraisal, Zhejiang University, Hangzhou 310027, China

ARTICLE INFO

Article history:
Received 1 May 2016
Received in revised form 15 February 2017
Accepted 5 April 2017

Keywords:
Vesicle fusion event
TIRFM Image
Hierarchical Convolutional Neural Network

ABSTRACT

Quantitative analysis of vesicle exocytosis and classification of different modes of vesicle fusion from the fluorescence microscopy are of primary importance for biomedical researches. In this paper, we propose a novel Hierarchical Convolutional Neural Network (HCNN) method to automatically identify vesicle fusion events in time-lapse Total Internal Reflection Fluorescence Microscopy (TIRFM) image sequences. Firstly, a detection and tracking method is developed to extract image patch sequences containing potential fusion events. Then, a Gaussian Mixture Model (GMM) is applied on each image patch of the patch sequence with outliers rejected for robust Gaussian fitting. By utilizing the high-level time-series intensity change features introduced by GMM and the visual appearance features embedded in some key moments of the fusion process, the proposed HCNN architecture is able to classify each candidate patch sequence into three classes: full fusion event, partial fusion event and non-fusion event. Finally, we validate the performance of our method on 9 challenging datasets that have been annotated by cell biologists, and our method achieves better performances when comparing with three previous methods.

1. Introduction

Vesicle exocytosis is an essential cellular trafficking process, by which materials (e.g., transporters, receptors and enzymes) are transported from one membrane-bounded organelle to another or to the plasma membrane for growth and secretion. Vesicle exocytosis needs to be highly regulated since its dysregulation is related to many human diseases (e.g., neurodegenerative disease, cancer and diabetes) (Hou et al., 1997; Jahn et al., 2012). Different modes of vesicle exocytosis have been found and characterized in mammalian cells. These include the *full fusion* where a vesicle collapses completely when it fuses with the plasma membrane, and the *partial fusion* or "kiss-and-run" fusion where a vesicle transiently fuses with the plasma membrane without the full collapse (Rizzoli et al., 2007; Xu et al., 2011). In cell biology research, it is of great importance to detect vesicle fusion events and also to classify different modes of vesicle exocytosis. Because the quantitative analysis of these biological processes can provide insights into cellular behaviors in normal and disease conditions.

Total Internal Reflection Fluorescence Microscopy (TIRFM), which illuminates the aqueous phase immediately adjacent to a glass interface with an exponentially decaying excitation (about 100 nm in *z*-axis), has been used widely to visualize single vesicle exocytosis at the cell surface (Axelrod et al., 1981; Schneckenburger et al., 2005). A pH-sensitive mutant of GFP, pHluorin, was developed and expressed to visualize vesicle exocytosis (Gero et al., 1998). Usually, pHluorin is targeted to the lumen of the vesicle, which is quenched and non-fluorescent in acidic environment, but becomes brightly fluorescent when the vesicle exposes to the extracellular neutral environment as the vesicle fuses with the plasma membrane (Xu et al., 2011, 2016). In this study, we imaged a variety of vesicle exocytosis in different types of mammalian cells. These include constitutive exocytosis (transferrin receptor-pHluorin exocytosis in endothelial cells and 3T3-L1 adipocytes) and regulated exocytosis (VAMP2-pHluorin labeled insulin granule in MIN-6 cells and VAMP2-pHlurin labeled GLUT4 vesicle in 3T3-L1 adipocytes). Quantitative analysis of the vesicle exocytosis in these

* Corresponding author.
E-mail addresses: hl87c@mst.edu (H. Li), ym8r8@mst.edu (Y. Mao), yinz@mst.edu (Z. Yin), yingkexu@zju.edu.cn (Y. Xu).

[1] H. Li, Y. Mao and Z. Yin were supported by NSF CAREER award IIS-1351049, NSF EPSCoR grant IIA-1355406, Intelligent System Center and Center of Biomedical Science and Engineering at Missouri University of Science and Technology.
[2] Y. Xu was supported by the National Basic Research Program of China (2015CB352003), National Natural Science Foundation of China (31571480, 31301176 and 61427818) and National Key Research and Development Program of China (2016YFF0101406).

http://dx.doi.org/10.1016/j.compmedimag.2017.04.003
0895-6111/

typical examples will strengthen our understanding of how vesicle exocytosis is regulated and how its dysregulation triggers human disease (e.g., insulin resistance and diabetes) (Bornemann et al., 1992; Leney et al., 2009; Xu et al., 2011).

Usually, the membrane fusion between pHluorin labeled vesicles and the plasma membrane can be represented by 2 significant stages in a continuous video sequence, as illustrated in Fig. 1. In stage 1, the vesicle is invisible in the *pre-appearance frame* (quenched), and then suddenly appears in the *first-appearance frame* as a brightly fluorescent circle spot. In stage 2, after being immobilized for some frames (from about 100 ms to a few seconds), the vesicle will either fuse completely with the plasma membrane with a visible bright "halo" (full fusion event), or remain its circular shape and gradually fade (partial fusion event), which can be observed in the *last appearance frame*, respectively. At the end of this process, the vesicle under the full or partial fusions will disappear in the *disappearance frame*. Note that, since the moving trajectory of vesicles during the exocytosis process is almost perpendicular to the cell membrane, the trajectory projected onto the cell surface (i.e., the image plane in the TIRFM) only has a small spatial displacement. In this movement process, the appearance variation pattern of the vesicle fusion event is a critical characteristic that is able to generate representative features to distinguish the vesicle fusion event from the background. Specially, the *pre-appearance frame*, *first-appearance frame*, *last-appearance frame* and *disappearance frame* are the 4 key moments of the vesicle fusion event, which represent the significant appearance change of a given fusion event.

A typical time-lapse TIRFM movie consists of thousands of individual frames with hundreds of vesicle fusion events. Unfortunately, so far the vesicle fusion detection and classification are performed mainly in a manual manner, which is a very time-consuming process, and likely to introduce personal biases. Therefore, there is a great demand to develop effective computational tools to automatically extract the vesicle fusion event information in TIRFM video sequences, which will aid the quantitative analysis on the vesicle exocytosis process.

1.1. Related work

When the computer-based microscopy image analysis is used to relieve human from the tedious manual labeling (Basset et al., 2014, 2015; Godinez et al., 2009), it is unsurprising that lots of challenges, such as the uncontrollable noise interference of TIRFM images and the high variability of fusion events' properties (e.g., intensity profiles, lifetime length and movement patterns), hinder the automated image processing. Furthermore, some of the bright spots (endocytic vesicles or vesicles from other non-acidic compartments) in TIRFM image sequences are moving in and out of the TIRFM field, which is a great challenge for designing automated algorithms for vesicle fusion detection. In order to detect fusion events, one needs to use specific detection algorithms considering both spatial and temporal features of individual objects.

Based on the bright circular appearance of vesicle fusions under the TIRFM, some approaches have been proposed to perform automated fusion identification, such as the pixel intensity thresholding methods in Huang et al. (2007), Yuan et al. (2015) and the intensity distribution analysis methods in Smith et al. (2011), Dosset et al. (2015). However, these methods are sensitive to the variation of vesicle fusion intensity profiles (shown in Fig. 2(a–c)). In order to improve the tolerance to the variation, some automated approaches were developed to model the moving process of fusion events. Based on both the temporal and spatial features, a template matching method was proposed to identify the fusion events with high correlation to a standard fusion event template in Vallotton et al. (2007). In another study, a Gaussian model was used to fit typical fusion events in Bai et al. (2007), where the parameters in the Gaussian model are used to classify fusion events. However, due to the frequent background intensity fluctuations (as shown in Fig. 2(d–f)) introduced by the TIRFM system and intracellular activities, it is hard to build a standard template or a general model to represent all fusion events.

Because of the large variations of the fusion events' properties (e.g., intensity profiles, lifetime length and movement patterns) and frequent background fluctuations, the robustness of a vesicle fusion detection and classification method is highly important. A robust detection method was proposed in Lorenz et al. (2010), which first detects candidate fusion events that suddenly appear in the TIRFM field. Then, a diffusive model is developed to analyze the intensity distribution variation pattern of the fusion event for the classification. Based on the visible "puff" phenomenon of the full fusion event, the diffusive fusion model effectively distinguishes full fusion events from non fusion regions, leaving a large amount of partial fusion events unrecognized. In addition, a Layered Probabilistic Approach was proposed in Godinez et al. (2012) to identify full fusion events by exploring three abstractions: the intensity over time, the underlying temporal intensity model and the high level behavior. Each of these three abstractions corresponds to a layer and these layers are represented via stochastic hybrid systems and hidden Markov models. However, partial fusion events are not considered in this work.

Unlike the full fusion event, which can be distinguished by its "puff"/spread signal, the partial fusion event is resembled to other bright spots (Fig. 2(d–f)) on the background, which is problematic in most of the existing detection and classification methods. In order to reveal the unique variation pattern of the fusion events, a learning based method was developed in our previous work (Li et al., 2015). An adaptive detection and tracking method is first applied to TIRFM images to search for potential fusion patches through video frames, then a Gaussian Mixture Model (GMM) is fitted on each individual fusion event. Using the estimated parameters of this model as features, a classifier is trained to distinguish full fusion events, partial fusion events and non-fusion events. However, in this GMM-based method, the handcrafted features ignore the discriminative appearance information from the 4 key moments of a fusion event, which leads to miss-detection problems in short fusion events (shown in Fig. 2).

1.2. The major challenges

According to the observation of our own datasets and the review of previous works, the major challenges to the task of detecting and classifying vesicle fusion events are summarized as follows:

The high variability of the vesicle fusion events. Some typical partial fusion events and full fusion events are shown in Fig. 2(a) and (b) respectively, from which we can observe the characteristics of vesicle fusion events. For example, normally partial fusion events present the momentary appearance and disappearance, and full fusion events present a sudden appearance and a gradual disappearance with their signals fading away. However, in practical cases, the vesicle fusion event has large variations in its intensity profile, lifetime and movement pattern. For instance, compared with a typical full fusion event in Fig. 2(b), the full fusion event in Fig. 2(c) has a much shorter lifetime and a much more blurry intensity profile. These variations yield challenges in modeling the various visual patterns of fusion events.

Complex background interferences. Besides vesicle fusion events, there exist a large amount of other bright circular spots on the background, which are challenges for automated fusion event detection and classification. For instance, the circular background intensity fluctuation (Fig. 2(d)) is similar to a partial fusion event. Some moving bright spots, which are temporarily immobile near

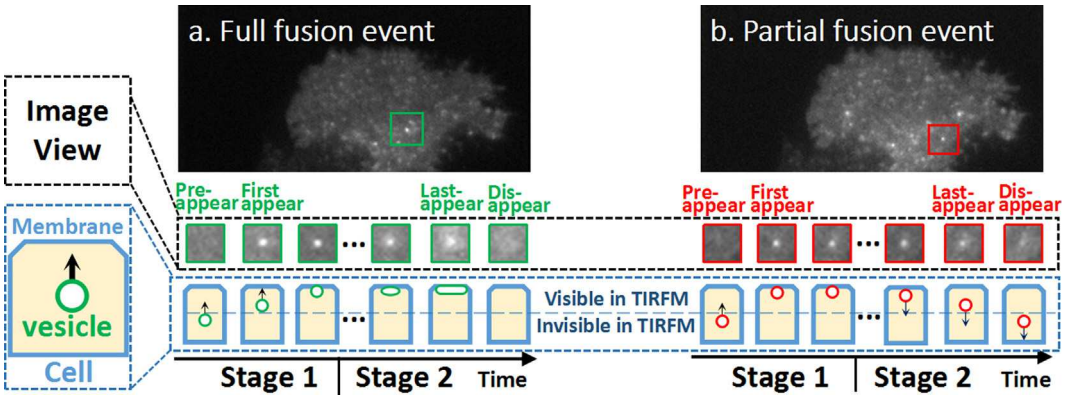

Fig. 1. The 2 significant stages of vesicle fusion processes and the related 4 key moments. 3T3-L1 adipocytes were transfected with VAMP2-pHluorin to label the GLUT4 vesicles. pHluorin is a pH-sensitive fluorescent protein that is invisible in the lumen of acidic vesicles, which becomes much more fluorescent when a vesicle fuses with the plasma membrane and exposes to a neutral environment. After a vesicle touches the cell membrane, it either fully collapses and fuses with the plasma membrane ((a) Full fusion event), or partially fuses with the plasma membrane and then is retrieved rapidly by the clathrin-dependent process ((b) Partial fusion event).

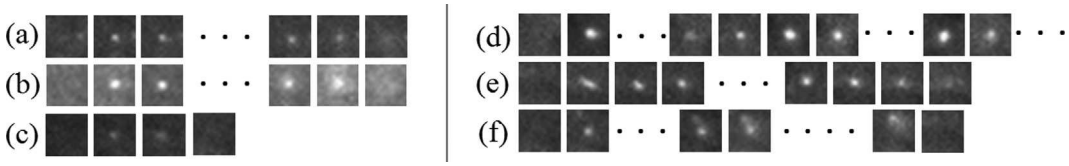

Fig. 2. Some samples of partial fusion event (a), full fusion events (b and c), and non-fusion events (d–f). (a) A typical partial fusion event; (b) A typical full fusion event with the "puff" phenomenon; (c) A short full fusion event is characterized by its "puff" phenomenon; (d) A bright circular object caused by the background intensity fluctuation; (e) A moving bright spot, which only moves in the first several frames then stays immobile, is similar to a partial fusion event when it stops moving; (f) A background fluctuation, which is really similar to standard full fusion events in the early stage, then gradually moves out of the field of view.

the cell membrane for several frames (Fig. 2(e and f)), can be mistakenly classified as partial fusion events. These interferences yield challenges in selecting effective features to build discriminative classifiers.

1.3. Our proposal and contributions

Rather than designing handcrafted visual models or features, Convolutional Neural Networks (CNN) that can learn the discriminative features from big training data have been widely used in different real world classification tasks, such as image recognition (Krizhevsky et al., 2012; Lawrence et al., 1997), video analysis (Yue-Hei et al., 2015; Karpathy et al., 2014) and natural language processing (Hu et al., 2014; Kim et al., 2014). CNN is a promising learning based method to handle classification challenges on microscopy images, such as cell detection (Mao et al., 2016a,b). Therefore, in order to enhance the tolerance to the variation of fusion events and the unpredictable background interferences, we propose to develop a novel CNN-based application which applies a Hierarchical Convolutional Neural Network (HCNN) to explore both appearance features and temporal cues for the vesicle fusion event classification. First, we extract fusion event candidate sequences and their appearance features from the input video data by using a newly developed iterative tracking algorithm. Secondly, a center-surrounded Gaussian Mixture Model (GMM) is fit on each patch of the patch sequence using the RANSAC algorithm (Martin et al., 1981) to remove outliers during the fitting process. The patch sequences are aligned with the same time length and time-series intensity change features corresponding to the Gaussian models' parameters are extracted over time. Thirdly, based on the time-series parameters from Gaussian Mixture Models and 4 key moments of the fusion event candidate sequence, a HCNN is developed to automatically select discriminative temporal and appearance features for the classification of the fusion event candidates in challenging datasets with low Signal-to-Noise-Ratio and frequent background fluctuations.

Our contributions in this paper include: (1) A novel application is proposed to detect and classify vesicle fusion events. The Hierarchical Convolutional Neural Network (HCNN) is utilized to learn discriminative appearance features from 4 key moments of a fusion event and combine them with the temporal features from the parametric Gaussian Mixture Models over time; (2) A center-surrounded Gaussian Mixture Model is used to model the intensity profile change of a fusion event in its entire lifetime; (3) A newly developed vesicle fusion event tracking algorithm is applied for the appearance feature extraction.

The rest of this paper is organized as follows: in Section 2, we briefly introduce our newly developed vesicle fusion event tracking algorithm, which contributes to appearance feature extraction for fusion event classification; in Section 3, the classification of the fusion event candidates by HCNN is presented; in Section 4, we validate our method on 9 challenging datasets and compare it with the previous methods and other neural network architectures. The paper concludes with Section 5.

2. Detection and tracking algorithm

Based on our preliminary work on detecting and tracking vesicle candidates in video sequences (Li et al., 2015), we improved the tracking algorithm to accurately measure the lifetimes of vesicle fusion events, which is important for the feature extraction task in fusion event classification. The major goal of our new tracking algorithm is to find the *first-appearance frame* and the *last-appearance frame* of a potential fusion event and every patch center between the *first-appearance frame* and the *last-appearance frame*. We utilize Fig. 3 to illustrate how to iteratively search in the forward direction to find the *last-appearance frame* (the search in the backward direction to find the *first-appearance frame* is similar).

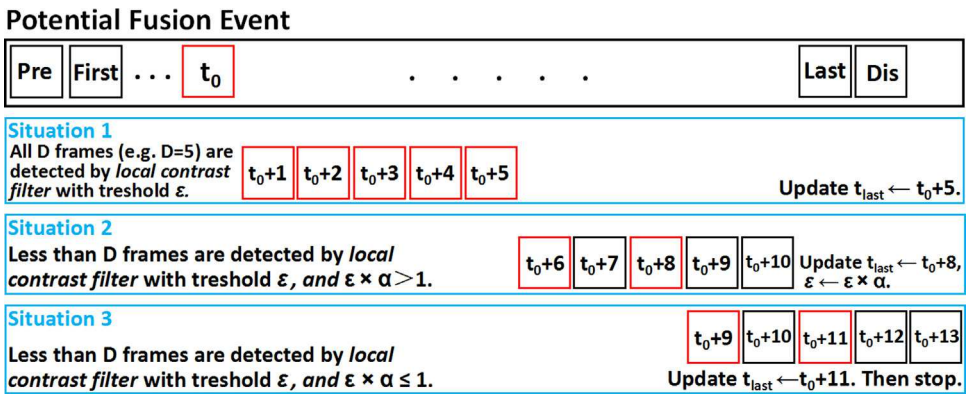

Fig. 3. An example to search the candidate patch sequence S in the forward temporal direction.

Assume we find the pixel (x^*, y^*) with the local maximum of local contrast as the center of the potential fusion event and crop an $n \times n$ image patch around it. Since we use fixed size patches, we only need to record the coordinates of the patch center in the fusion event candidate patch sequence, which are denoted as $S = \left\{x_t^*, y_t^* | t \in [t_{first}, t_{last}]\right\}$ where t_{first} and t_{last} denote the first and last frame index of the patch sequence, respectively. At the beginning, $t_{first} = t_{last} = t_0$. During each iteration, we search the *last-appearance frame* in a sliding temporal window of D frames. Three situations are considered during the iterative search:

Situation 1, if the maximums of the local contrast in all D frames around location $(x^*_{t_{last}}, y^*_{t_{last}})$ are larger than ε, so we can update $S = \left\{x_t^*, y_t^* | t \in [t_{first}, t_{last}]\right\}$ by setting $t_{last} \longleftarrow t_{last} + D$. Then, we continue the search from frame $t_{last} + 1$ to frame $t_{last} + D$.

Situation 2, if not all of the maximums of the local contrast in D frames around location $(x^*_{t_{last}}, y^*_{t_{last}})$ are larger than ε, while $\varepsilon \times \alpha > 1$ (α is a decay rate on the threshold), we update t_{last} as the last frame within the D frame whose maximal local contrast is larger than ε and the patch centers are updated accordingly. The threshold is updated as $\varepsilon \longleftarrow \varepsilon \times \alpha$. Then, we continue the search from frame $t_{last} + 1$ to frame $t_{last} + D$.

Situation 3, if not all of the maximums of the local contrast in D frames around location $(x^*_{t_{last}}, y^*_{t_{last}})$ are larger than ε and $\varepsilon \times \alpha \leq 1$, we update the patch sequence similar to situation 2, then we stop the iteration.

By applying this iterative tracking algorithm to the TIRFM image sequence, we can obtain the whole lifetimes of potential fusion events in the format of candidate patch sequences, each of which records the coordinates of the patch center from the *first-appearance frame* to the *last-appearance frame*. For each potential fusion event, we compute the pairwise Euclidean distance between each consecutive pair of patch centers within the candidate patch sequence. If any of these distances is larger than the neighborhood size n, this candidate patch sequence is highly possible to be a non-fusion event caused by a moving object from the background, and we remove it from the candidate list.

3. Classification of fusion event candidates

In this section, we will introduce the classification of fusion event candidates by using a novel Hierarchical Convolutional Neural Network (HCNN). Compared with the Support Vector Machine-based classification method in Li et al. (2015), HCNN is able to automatically select discriminative features which can provide the comprehensive representation of the fusion event. In order to enhance the tolerance to the variation of fusion events and the unpredictable background interferences, the proposed HCNN architecture considers both spatial and temporal information. The input of our HCNN consists of the time-series parametric information from the Gaussian Mixture Model fitting, and the visual appearance information from the 4 key moments of the fusion event candidate. The former is aiming at revealing the unique hidden variation pattern of the vesicle fusion event in its entire lifetime. The latter is proposed to extract the extraordinary visual appearance features of the vesicle fusion event. Moreover, the hierarchical architecture is able to exploit the high-level abstraction of intensity profiles of individual frames and the high-level temporal features from the entire fusion event lifetime to accurately distinguish fusion events from the other similar circular bright spots in Fig. 2.

3.1. Data preparation

Because of the frequent background interferences in the TIRFM video data, directly thresholding the candidate patch sequence might not be a good option to present its intensity profile variation. Therefore, we adopt the data preparation strategy in our previous work (Li et al., 2015). First, a robust Gaussian Mixture Model (GMM), which consists of two center-surrounded 2D Gaussian models (**Area**$_p$ and **Area**$_f$ in Fig. 4), is adopted to fit the intensity profile of each fusion event candidate, where a Random Sample Consensus algorithm (Martin et al., 1981) is applied to robustly estimate the parameters of Gaussian models without the outlier effect. Second, since most of the fusion events have their lifetimes less than 24 frames in the datasets we used in this study, we extract 24 image patches from each fusion event candidate starting from the *first-appearance frame*. For those fusion event candidates whose lifetimes are shorter than 24 frames, we will zero-padding them. For those fusion event candidates with longer lifetimes, they will be cut into the time length. Third, for each fusion event candidate, there are 24 extracted image patches in the patch sequence, where each image patch is represented by a set of GMM parameters (λ_{peak}[3], μ_{peak}, σ_{peak} of **Area**$_p$, and λ_{flat}, μ_{flat}, σ_{flat} of **Area**$_f$). Thus, the time-series intensity profile change of a vesicle fusion event candidate, which is represented by 24 sets of GMM parameters, can be utilized for fusion event classification.

3.2. The variation pattern in GMM image

In order to explore the hidden correlations among the image patches in each fusion event candidate, we generalized the vectorization process in our previous work (Li et al., 2015) by transforming the parameter sets of a fusion event candidate into a 2D image,

[3] λ is the weighting coefficient of each Gaussian component in the GMM.

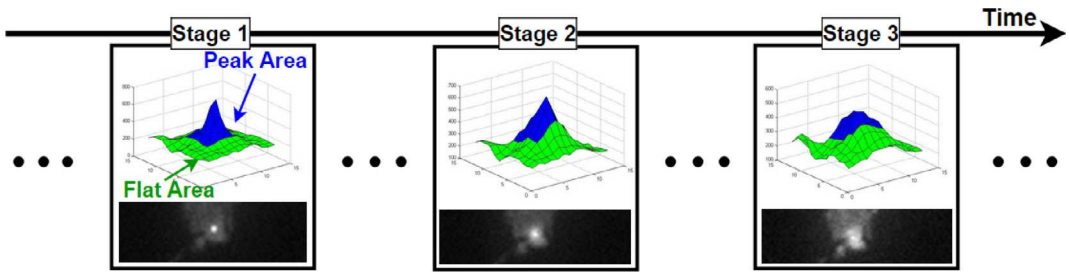

Fig. 4. The Gaussian Mixture Model consists of a 5 × 5 "peak area" and a 13 × 13 "flat area".

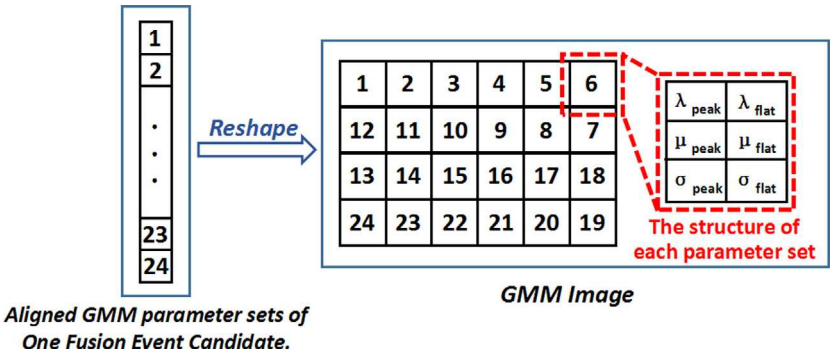

Fig. 5. Transforming the time-series Gaussian fitting parameter sets to a 2D array (Gaussian Mixture Model image, *GMM image*). In the *GMM image*, each cell represents a parameter set for one image patch of the fusion event candidate. In each cell, the 6 parameters are organized as a 3 × 2 matrix (λ_{peak}, λ_{flat};μ_{peak}, μ_{flat};σ_{peak}, σ_{flat}). So the *GMM image*, which contains 24 cells, is a 12 × 12 matrix.

which concatenates the time-series parameter sets into a 2D array in a special order, as shown in Fig. 5. We call this 2D array of Gaussian Mixture Model fitting parameters as *GMM image* that allows the HCNN to discover the hidden correlation among the parameter sets. Furthermore, in Fig. 5, we design the *GMM image* to be a square image, so each parameter set has more chances to be neighboring to other parameter sets. For example, given 24 parameter sets to stitch, if they are concatenated into a 24 × 1 matrix pattern, there is no 4- or 8-connected neighborhood relationship among the parameter sets. However, if we stitch them into a 12 × 2 matrix pattern, the relationship among the parameter sets will increase a little. Thus, in this work, we concatenate the 24 parameter sets into a 4 × 6 matrix pattern, many 4- or 8-connected neighborhood relationships can be built among the parameter sets.

3.3. The visual appearance in 4 key moments

In addition to the *GMM image*, which contains the high-level abstraction of intensity profiles of individual frames, we also consider the appearance features in the 4 key moments of a fusion event candidate. As described in Fig. 1, the movement of vesicles can be well represented in the 4 key moments: *pre-appearance frame*, *first-appearance frame*, *last-appearance frame* and *disappearance frame*. By using our newly developed vesicle fusion event tracking method, the whole entire of each fusion event candidate is able to be obtained. Therefore, for each candidate, we extract image patches in these 4 key moments. The *first-appearance frame* patch and *last-appearance frame* patch are extracted from the first frame and the last frame in the fusion event lifetime, respectively. The *pre-appearance frame* patch is extracted from the previous frame of the *first-appearance frame*. The *disappearance frame* patch is extracted from the next frame of the *last-appearance frame*. Both the parametric information from the *GMM image* and the 4 image patches of the 4 key moments will be input to the HCNN.

3.4. The architectures of our HCNN

The overall architecture of our Hierarchical Convolutional Neural Network (HCNN) is shown in Fig. 6. In the first layer, the inputs of the first 4 Convolutional Neural Networks CNN_1^j ($j \in [1, 4]$) are the cropped image patches from 4 key moments, which provide the detailed visual appearance information of fusion event candidates. Each of these four CNNs takes a single cropped image patch. The input of the CNN_1^5 is the *GMM image* which provides the time-series intensity change information of the fusion process (a high-level abstraction using the parameters from Gaussian Mixture Model fitting). In the second layer of our HCNN, we design the CNN_2^6 to learn joint features of the CNN_1^j ($j \in [1, 4]$), which indicate the correlation of fusion event patches in the 4 key moments. In the third layer, the combined appearance and time-series intensity change features are fed into the CNN_3^7 to make the final prediction. In our notation of CNN_i^j, i denotes the layer in our HCNN and j indexes the CNN out of the total 7 CNNs in our proposed HCNN architecture.

The design of our proposed HCNN architecture has three motivations. First, the intensity variation pattern of a fusion event, which is different from other bright circular spots in TIRFM image sequences, is a significant characteristic to classify fusion events. Instead of directly using the consecutive image patch sequence to provide this time-series intensity change information, the time-series parameter sets from Gaussian Mixture Model fitting, which can avoid outlier pixels with undesired intensity fluctuations, are more reliable and the proposed *GMM image* can further explore hidden relations among the time-series parameters. Second, the characteristics of a fusion event's appearances can be well represented in the 4 key moments, thus utilizing these appearance characteristics and the correlation among the 4 key moments should boost the classification performance. Third, our proposed HCNN architecture is able to learn the correlation among the 4 key moments before combing the appearance and temporal features, which can reveal the unique variation pattern of the fusion event.

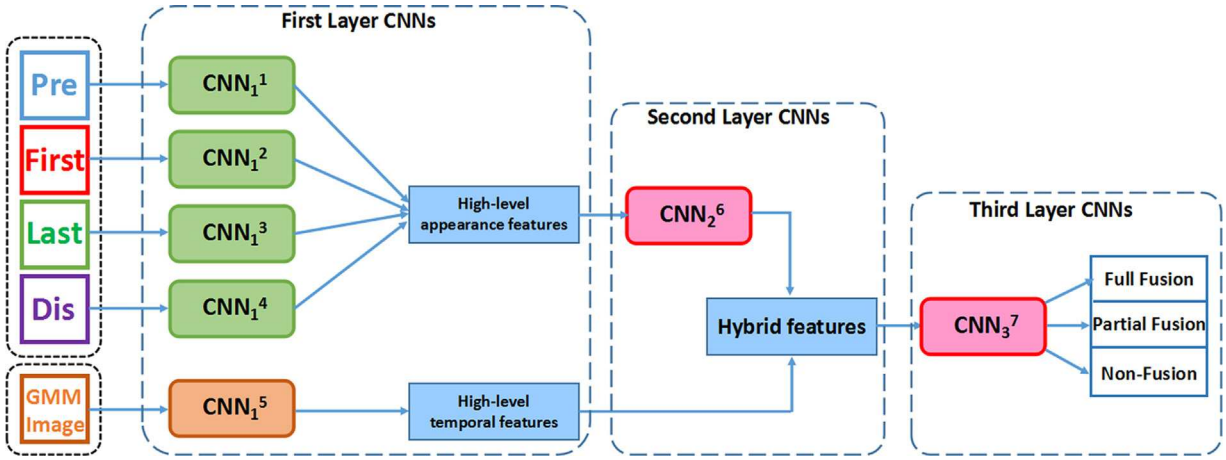

Fig. 6. The overall architecture of our proposed Hierarchical Convolutional Neural Network (HCNN).

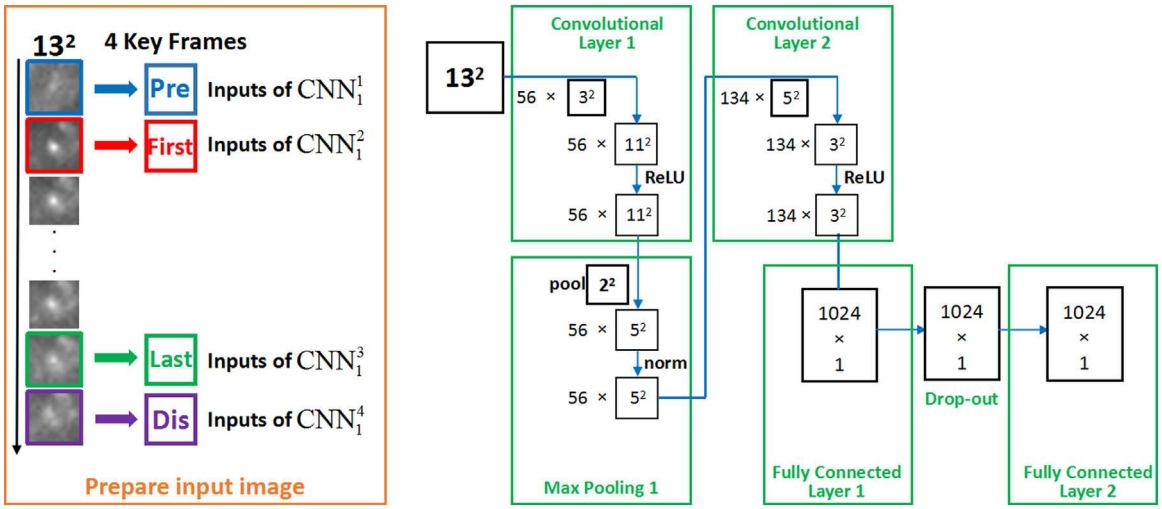

Fig. 7. The architecture of CNNs ($CNN_1^j, j \in [1, 4]$) in the first layer. The inputs of this architecture are image patches which are centered at the maximum intensity pixels of the fusion event in the 4 key moments respectively. In the Convolutional Layer 1, we set the number of the 3×3 kernels as 56. In the Convolutional Layer 2, we set the number of the 5×5 kernels as 134. In the Max Pooling 1, there is a 2×2 max pooling layer with stride 2. The number of neurons in each Fully Connected Layer is 1024.

The first layer of our HCNN contains 5 CNNs ($CNN_1^j, j \in [1, 5]$). The first 4 CNNs ($CNN_1^j, j \in [1, 4]$), each of which takes a cropped image patch (13×13) of the fusion event in one of the 4 key moments as the input, share the same architecture as shown in Fig. 7. In the architecture of CNN_1^j ($j \in [1, 4]$), there are two Convolutional Layers where each of them is connected to a Rectified Linear Unit (ReLU) for sparse representations. The first Convolutional Layer is followed by a 2×2 Max Pooling Layer with stride 2. The major goal of adding Max Pooling Layer is to enhance the robustness of the classifier by bringing invariance to the training process. We add a Drop-out Layer (Srivastava et al., 2014) between the two Fully Connected Layers to avoid the over-fitting.

The CNN_1^5, whose architecture is shown in Fig. 8, learns the high-level time-series features from the intensity variation pattern introduced by the *GMM image*. There are 3 Convolutional Layers, where each Convolutional Layer is followed by a Rectified Linear Unit (ReLU) for sparse representations. Compared with the other 4 CNNs in the first layer, there is no Max Pooling Layer in CNN_1^5. Because we do not expect to loss any time-series variation information during the convolution. To avoid the over-fitting, we add one Drop-out Layer between the Fully Connected Layer 1 and Fully Connected Layer 2.

The architecture of the CNNs in the second and last layer of our HCNN (CNN_2^6 and CNN_3^7) is shown in Fig. 9. The input feature layer to CNN_2^6 is the combined feature from the Fully Connected Layer 2 of CNN_1^j ($j \in [1, 4]$). The design of CNN_2^6 is to study the correlation information among the 4 key moments before combining appearance features and time-series variation features. The input features to CNN_3^7 is the combined features of the time-series intensity variation features from the Fully Connected Layer 2 of CNN_1^5, and the visual appearance features from the Fully Connected Layer 2 of CNN_2^6. Between the Fully Connected Layer 1 and Fully Connected Layer 2, we add a Drop-out Layer to avoid the over-fitting.

4. Experiments

In this section, first we describe our datasets, experimental design and evaluation metrics. Then, we validate the effectiveness of our fusion event candidate extraction. Thirdly, we compare our method with the state-of-the-arts and our previous methods in Li et al. (2015). Finally, we validate our HCNN design by comparing it with 11 alternative neural network designs.

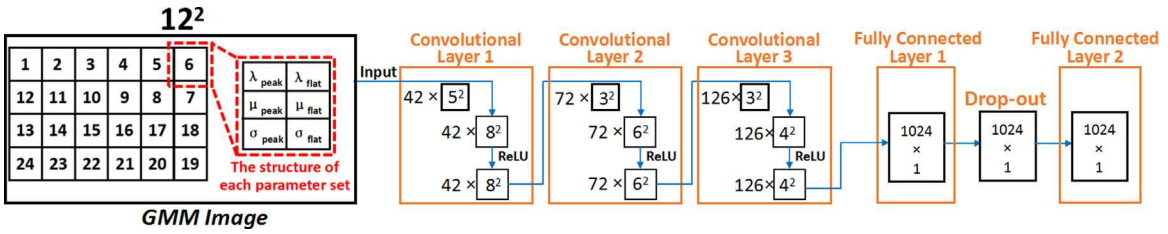

Fig. 8. The architecture of CNN_1^5 in the first layer of our HCNN. The input of this architecture is the *GMM image*. In the Convolutional Layer 1, we set the number of the 5×5 kernels as 42. In the Convolutional Layer 2, we set the number of the 3×3 kernels as 72. In the Convolutional Layer 3, we set the number of the 3×3 kernels as 126. The number of neurons in each Fully Connected Layer is 1024.

Fig. 9. The architecture shared by CNN_2^6 in the second layer and CNN_3^7 in the third layer. In CNN_2^6, the input feature layer contains the high-level appearance feature, which is extracted from the 4 key moments. In CNN_3^7, the input feature layer consists of visual and temporal information.

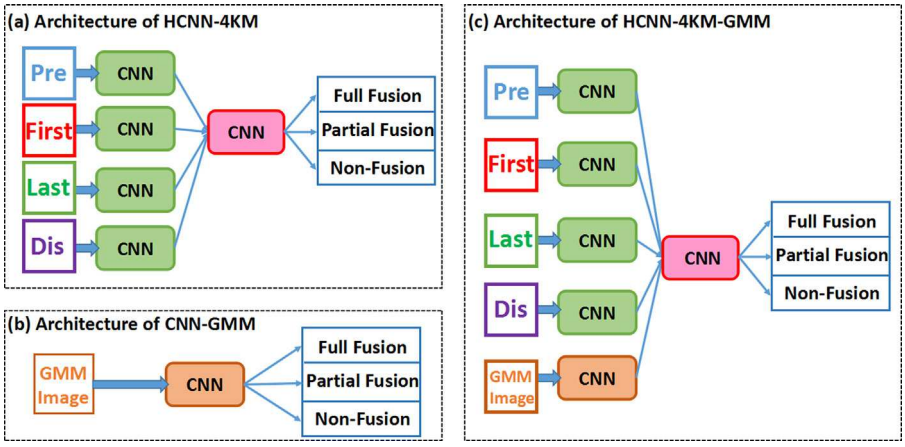

Fig. 10. The architectures of the HCNN-4KM (a), CNN-GMM (b) and HCNN-4KM-GMM (c).

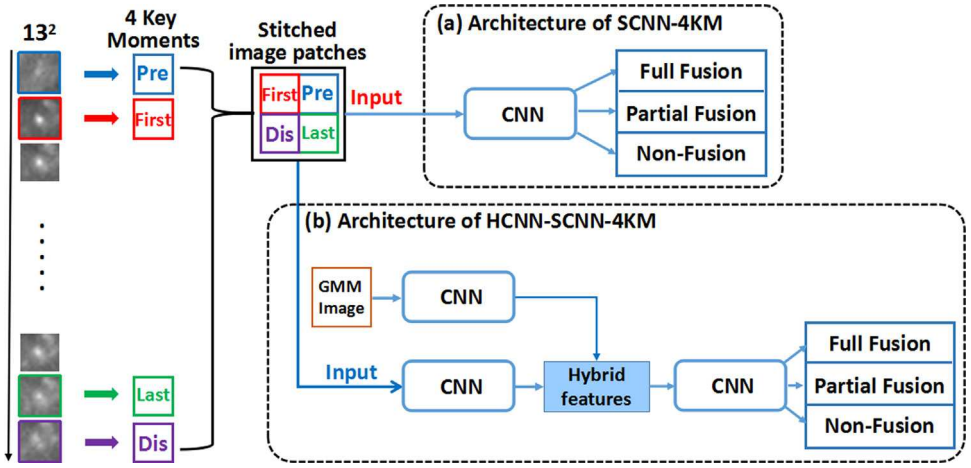

Fig. 11. The architectures of the SCNN-4KM (a) and HCNN-SCNN-4KM (b).

4.1. Datasets, experimental design and evaluation metric

Datasets. In the experiments, 9 TIRFM image sequences were captured at 5 frame per second (fps), which consist of 15718 frames and 1260 fusion events in total. The detailed information of our datasets is summarized in Table 1. All image sequences were well annotated by experienced cell biologists working in the field of vesicle trafficking analysis using TIRFM.

Experimental design and evaluation metric The leave-one-out strategy is adopted to evaluate the performance of our method, i.e., eight sequences are used for training while the last one is used for testing (the parameters in the detection and tracking process and the Gaussian Mixture Model (GMM) fitting are optimized by the 4-fold cross-validation using the eight training sets). There are totally 9 leave-one-out experiments are performed on the datasets. The average performance on the 9 experiments in terms of precision, recall and F-score is utilized as the evaluation metrics.

Table 1
The image size and the number of fusion events in each dataset.

DataSet	1	2	3	4	5
Full fusion event	118	169	31	132	48
Partial fusion event	28	64	56	6	10
Image size (pixels)	327 × 179	271 × 284	233 × 324	271 × 341	408 × 381
DataSet	6	7	8	9	
Full fusion event	16	19	76	193	
Partial fusion event	16	76	11	797	
Image size (pixels)	382 × 338	241 × 211	478 × 412	485 × 299	

Table 2
Comparing our method with 2 state-of-the-arts and our previous method on 9 challenging datasets: GMM-SVM (Li et al., 2015): Gaussian Mixture Model using Support Vector Machine classifier; SGM (Bai et al., 2007): Single Gaussian Model; LPA-FullFusion (Godinez et al., 2012): Layered Probabilistic Approach for full fusion detection.

	Full fusion			Partial fusion		
	Precision	Recall	*F* score	Precision	Recall	*F* score
Our method	**95.2%**	**96.2%**	**95.7%**	**96.1%**	**96.7%**	**96.4%**
GMM-SVM (Li et al., 2015)	76.9%	79.3%	78.1%	75.5%	76.0%	75.7%
SGM (Bai et al., 2007)	61.1%	64.7%	62.8%	64.6%	62.0%	63.3%
LPA-FullFusion (Godinez et al., 2012)	75.3%	72.3%	73.8%	N/A	N/A	N/A

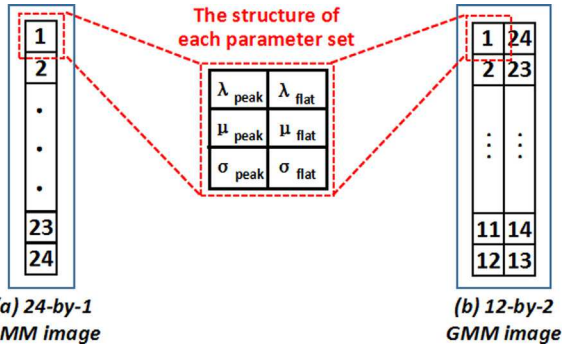

Fig. 12. The structures of 24 × 1 *GMM image* (a) and 12 × 2 *GMM image* (b).

4.2. Effectiveness of the fusion event candidate extraction

By using our newly developed detection and tracking method, we obtain 4642 candidate patch sequences on the 9 datasets. The candidate pool contains all the 1260 ground-truth vesicle fusion events from 15718 frames (i.e., our candidate sequence extraction achieves 100% recall and 27% precision). Instead of exhaustively selecting fusion event candidates from every volume of the TIRFM video sequences, the proposed detection and tracking method not only ensures all vesicle fusion events are included in the fusion event candidate pool, but also effectively improves the efficiency of the whole system. Note, data augmentation techniques (e.g., flipping, rotation and translation) were applied on our positive training samples to provide enough training data.

4.3. Comparison with the previous methods

Our algorithm is compared with the learning-based Gaussian Mixture Model using Support Vector Machine classifier (GMM-SVM, Li et al. (2015)), the intensity-based Single Gaussian Model (SGM, Bai et al. (2007)) and the Layered Probabilistic Approach (LPA-FullFusion, Godinez et al. (2012)). Note, the Layered Probabilistic Approach cannot detect partial fusion events. All the parameters in Bai et al. (2007), Godinez et al. (2012) and Li et al. (2015) are optimized to ensure that they can achieve their best performance in our TIRFM image sequences. As shown in Table 2, compared with the GMM-SVM (Li et al., 2015) that uses handcrafted features, our method achieved much better classification results for both the full fusion event and the partial fusion event in 9 datasets. It validates that the automatically selected features from the time-series intensity change introduced by GMM images and the visual appearance in 4 key moments by our HCNN architecture have a more comprehensive representation of the vesicle fusion event. Compared to the SGM (Bai et al., 2007) that only depends on the spatial radius of the Single Gaussian fit to the bright blob, our method achieved better classification results, which proves that the proposed Gaussian Mixture Model has a more precise representation to extract the intensity variation pattern of vesicle fusion events. Compared with the Layered Probabilistic Approach in Godinez et al. (2012), which uses three abstractions of fusion events as the feature for the classification, the temporal and spatial features extracted by GMM in our previous method (Li et al., 2015) achieved better classification results on full fusion events. Furthermore, our proposed HCNN architecture, which can automatically select the discriminative features from the whole lifetime of the fusion event, obtained the best performance. In short, besides visual appearance features, our HCNN based method can extract hidden variation patterns of the fusion event, which are qualified for the task of accurate fusion event classification. Fusion event classification samples of our proposed method are presented in Figs. 15 and 16 .

4.4. Comparison of different neural network designs

In this subsection, first we test different layouts in our overall architecture (Fig. 6) and compare the performance. Second, we test different input formats of the visual appearance features extracted from 4 key moments and compare the performance. Third, we test different input formats of the temporal features and compare the performance. Last, we test different designs in our individual CNNs (there are 7 CNNs in total, Fig. 6).

4.4.1. Comparison of alternative overall architecture designs

We designed the HCNN-4KM (Fig. 10(a)) that only considers appearance features, and the CNN-GMM (Fig. 10(b)) that only considers temporal features. As shown in Table 3, our HCNN architecture outperformed HCNN-4KM and CNN-GMM, which validates that both appearance features and temporal features contribute significantly to the fusion event classification task.

In order to show the importance of the CNN_2^6 in our proposed HCNN architecture (Fig. 6), we designed HCNN-4KM-GMM (Fig. 10(c)) by removing the CNN_2^6 from our HCNN, and compared

Table 3
Comparing our HCNN with 3 alternative overall architecture designs on 9 challenging datasets, which include: HCNN-4KM (Fig. 10(a)): based on our HCNN, we remove the CNN_1^5 and CNN_2^6 so only appearance features are used; CNN-GMM (Fig. 10(b)): based on our HCNN, we only use the temporal features in *GMM images* for the classification; HCNN-4KM-GMM (Fig. 10(c)): based on our HCNN, we remove the CNN_2^6.

	Full fusion			Partial fusion		
	Precision	Recall	*F* score	Precision	Recall	*F* score
Our method	**95.2%**	**96.2%**	**95.7%**	**96.1%**	**96.7%**	**96.4%**
HCNN-4KM	79.3%	82.4%	80.8%	85.1%	84.7%	84.9%
CNN-GMM	84.8%	88.7%	86.7%	82.0%	85.6%	83.8%
HCNN-4KM-GMM	94.1%	95.0%	94.6%	90.0%	92.7%	91.3%

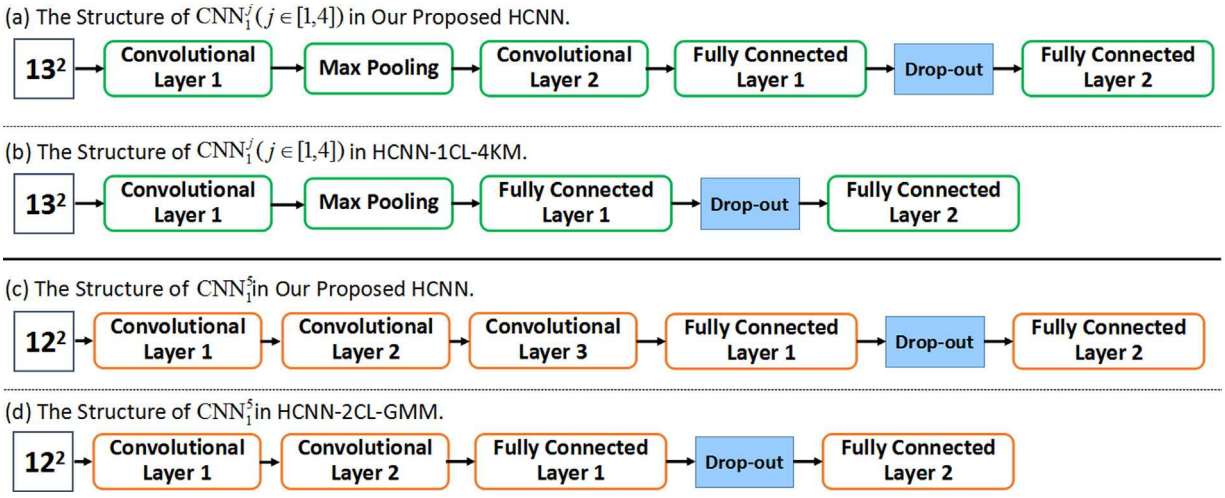

Fig. 13. (a) The CNN_1^j ($j \in [1, 4]$) structure in our proposed HCNN; (b) The CNN_1^j ($j \in [1, 4]$) structure in HCNN-1CL-4KM; (c) The CNN_1^5 structure in our proposed HCNN; (d) The CNN_1^5 structure in HCNN-2CL-GMM.

the classification results. As shown in Table 3, our proposed HCNN architecture achieved better classification results than HCNN-4KM-GMM, which proves that it is important to learn the correlation information among the 4 key moments before combining appearance and temporal features.

4.4.2. Comparison of alternative appearance feature input formats

In our HCNN architecture, we use 4 CNNs to learn the appearance features from the 4 key moments, where each CNN takes a single cropped image patch as input. In order to show the effectiveness of this design, we compared our HCNN architecture with SCNN-4KM (Fig. 11(a)), which is one single CNN whose inputs are the stitched image patches from 4 key moments, and HCNN-SCNN-4KM (Fig. 11(b)), which uses a CNN to learn appearance features from stitched image patches of 4 key moments and then combines with temporal features from GMM images for classification. As shown in Table 4, our proposed method achieved better classification results than SCNN-4KM and HCNN-SCNN-4KM, which validates the high-level appearance features extracted from 4 CNNs are more reliable for the fusion event classification task.

4.4.3. Comparison of alternative temporal feature input formats

In our proposed HCNN architecture, each *GMM image* consists of 24 parameter sets which are organized as a 4×6 matrix pattern (Fig. 5) to allow the HCNN to discover the hidden correlation among the parameter sets. In order to validate the effectiveness of our *GMM image* design, we compared our proposed 4×6 *GMM image* with the 24×1 *GMM image* (Fig. 12(a)) and the 12×2 *GMM image* (Fig. 12(b)). As shown in Table 5, our proposed HCNN architecture with 4×6 *GMM image* inputs achieved better classification results than HCNN-GMM(24×1) with 24×1 *GMM image* inputs and HCNN-GMM(12×2) with 12×2 *GMM image* inputs. It proves that the 4×6 matrix pattern *GMM image*, which contains many 4- or 8-connected neighborhood relationships, can provide comprehensive information to reveal the unique pattern of the fusion event.

4.4.4. Comparison of alternative CNN designs

To validate the effectiveness of the individual CNNs in our proposed HCNN architecture, we tested different number of Convolutional Layers and Fully Connected Layers and compared with our proposed HCNN. Since it is unpractical to test all possible CNN structures, we only tested some reasonable CNN designs in this work.

In our proposed HCNN architecture, the structure of CNN_1^j ($j \in [1, 4]$) has 2 Convolutional Layers (Fig. 13(a)) and the structure of CNN_1^5 has 3 Convolutional Layers (Fig. 13(c)). We designed HCNN-1CL-4KM (Fig. 13(b)) by setting only 1 Convolutional Layer to the structure of CNN_1^j ($j \in [1, 4]$) in our proposed HCNN, where the other CNNs in HCNN-1CL-4KM are exactly the same with the ones in our proposed HCNN. We also designed HCNN-2CL-GMM (Fig. 13(d)) by setting only 2 Convolutional Layers to the structure of CNN_1^5 in our proposed HCNN, where the other CNNs in HCNN-2CL-GMM are exactly the same with the ones in our proposed HCNN. As shown in Table 6, our proposed HCNN architecture outperformed HCNN-1CL-4KM and HCNN-2CL-GMM, which validates the effectiveness of the CNN_1^j ($j \in [1, 5]$) in our proposed HCNN architecture.

In our proposed HCNN architecture, the structure shared by CNN_2^6 and CNN_3^7 has 2 Fully Connected Layers (Fig. 14(a)). We designed HCNN-3FCL (Fig. 13(b)) by setting 3 Fully Connected Layers to the structure shared by CNN_2^6 and CNN_3^7, where the other settings in HCNN-3FCL are the same with our proposed HCNN. We also designed HCNN-1FCL (Fig. 13(c)) by setting only 1 Fully Connected Layer to the structure shared by CNN_2^6 and CNN_3^7, where the

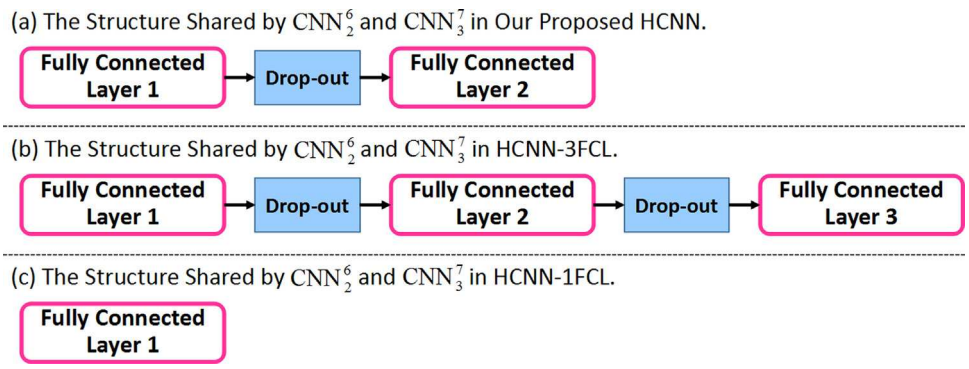

Fig. 14. (a) The structure shared by CNN_2^6 and CNN_3^7 in our proposed HCNN; (b) The structure shared by CNN_2^6 and CNN_3^7 in HCNN-3FCL; (c) The structure shared by CNN_2^6 and CNN_3^7 in HCNN-1FCL.

Table 4
Comparison of different input formats of the appearance features on 9 challenging datasets, which include: SCNN-4KM (Fig. 11(a)): we stitch the image patches from 4 key moments into an image, which will be the input to a CNN; HCNN-SCNN-4KM (Fig. 11(b)): based on our HCNN, instead of using 4 CNNs, we use a CNN to learn the appearance features from stitched image patches for the classification.

	Full fusion			Partial fusion		
	Precision	Recall	*F* score	Precision	Recall	*F* score
Our method	**95.2%**	**96.2%**	**95.7%**	**96.1%**	**96.7%**	**96.4%**
SCNN-4KM	93.7%	94.9%	94.3%	91.0%	93.2%	92.1%
HCNN-SCNN-4KM	94.8%	94.0%	94.4%	92.1%	91.7%	91.9%

Table 5
Comparison of 2 alternative *GMM image* designs on 9 challenging datasets, which include the 24×1 *GMM image* (Fig. 12(a)) in HCNN-GMM(24×1) and the 12×2 *GMM image* (Fig. 12(b)) in HCNN-GMM(12×2).

	Full fusion			Partial fusion		
	Precision	Recall	*F* score	Precision	Recall	*F* score
Our method	**95.2%**	**96.2%**	**95.7%**	**96.1%**	**96.7%**	**96.4%**
HCNN-GMM(24×1)	91.3%	94.0%	92.6%	92.3%	93.0%	92.7%
HCNN-GMM(12×2)	93.3%	93.5%	93.4%	92.3%	94.3%	93.3%

Table 6
Comparison of 4 alternative CNN designs in our proposed HCNN architecture on 9 challenging datasets, which include: HCNN-1CL-4KM, HCNN-2CL-GMM, HCNN-3FCL and HCNN-1FCL. These architectures are described in Section 4.4.4 in details.

	Full fusion			Partial fusion		
	Precision	Recall	*F* score	Precision	Recall	*F* score
Our method	**95.2%**	**96.2%**	**95.7%**	**96.1%**	**96.7%**	**96.4%**
HCNN-1CL-4KM	86.0%	84.1%	85.0%	82.9%	85.4%	84.1%
HCNN-2CL-GMM	82.7%	80.2%	81.4%	81.4%	80.3%	80.9%
HCNN-3FCL	94.8%	95.0%	94.9%	95.9%	96.3%	96.1%
HCNN-1FCL	91.0%	93.5%	92.2%	93.2%	95.2%	94.2%

other settings in HCNN-1FCL are the same with our proposed HCNN. As shown in Table 6, our proposed HCNN architecture achieved the best performance, which validates the effectiveness of the CNN_2^6 and CNN_3^7 in our proposed HCNN architecture.

4.5. Discussion

According to the classification results of our proposed method, there are two main failure cases in our experiments. First, during our data collection, the Total Internal Reflection Fluorescent Microscope (TIRFM) sometimes was out of focus for several frames, as shown in Fig. 17. Our proposed tracking method can still detect the image patches while the TIRFM is out of focus, but the intensity variation pattern of the fusion event is largely interfered by the out-of-focus problem, which misleads the HCNN to make a wrong classification. Second, some fusion events have extremely short lifetimes which are as short as 2 frames. For the short event process, the time-series intensity variation information from Gaussian fitting and the patches from the key moments are not very informative for the classification. Refining our current TIRFM hardware and increasing the image acquisition rate will be our future work to overcome the current drawbacks.

5. Conclusion

Accurately detecting and classifying vesicle-plasma membrane fusion events from TIRFM images is an essential research problem on cellular trafficking processes. In this paper, we proposed a novel Hierarchical Convolutional Neural Network (HCNN) based application to solve the fusion event detection and classification task. An adaptive detection and tracking method is developed to extract fusion event candidates and their time-series intensity variation information. By using the time-series intensity variation pattern introduced by Gaussian Mixture Models and the appearances in 4

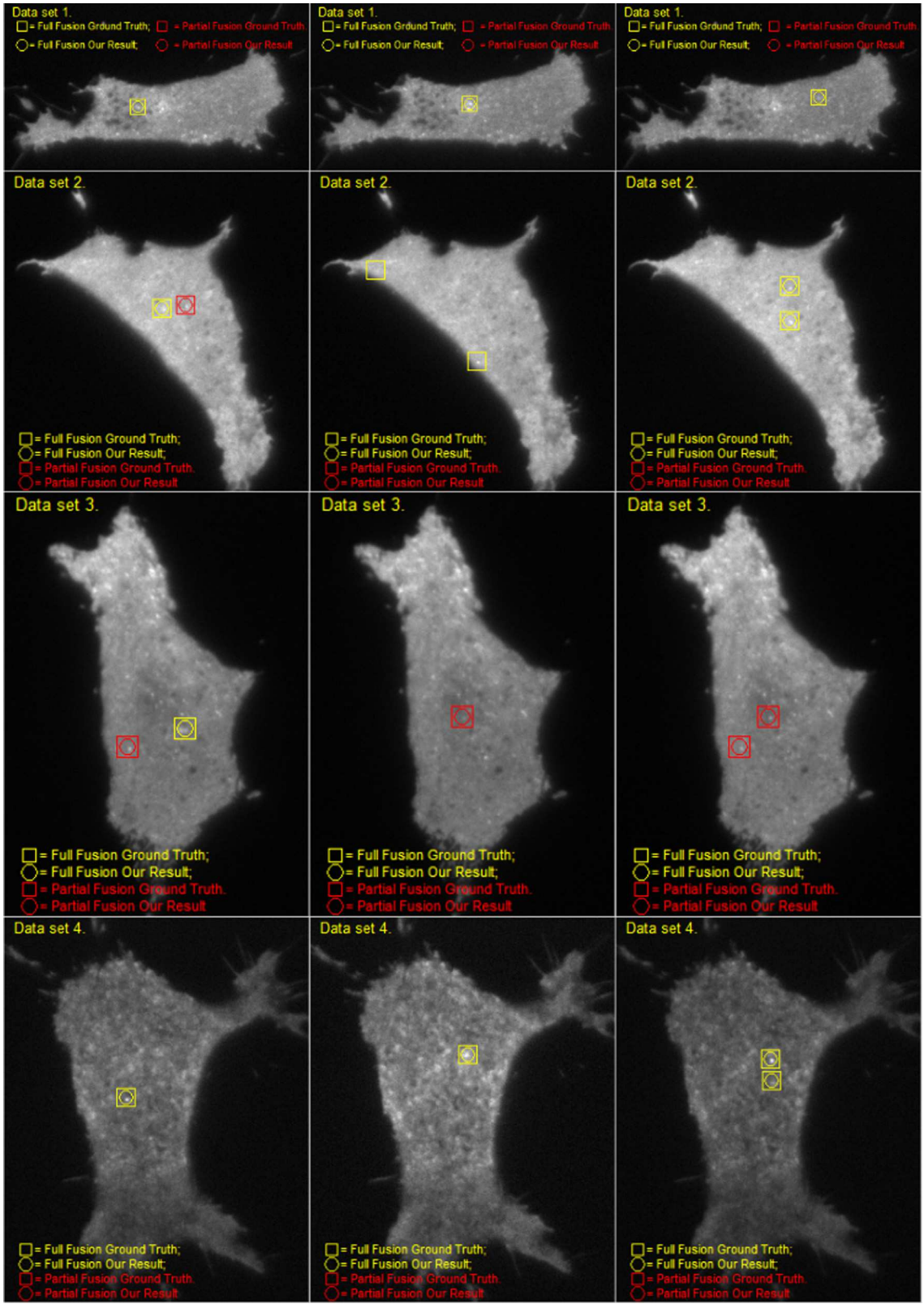

Fig. 15. Fusion event classification samples of dataset 1, 2, 3, 4 (yellow: full fusion; red: partial fusion; square: ground truth; circle: our result). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

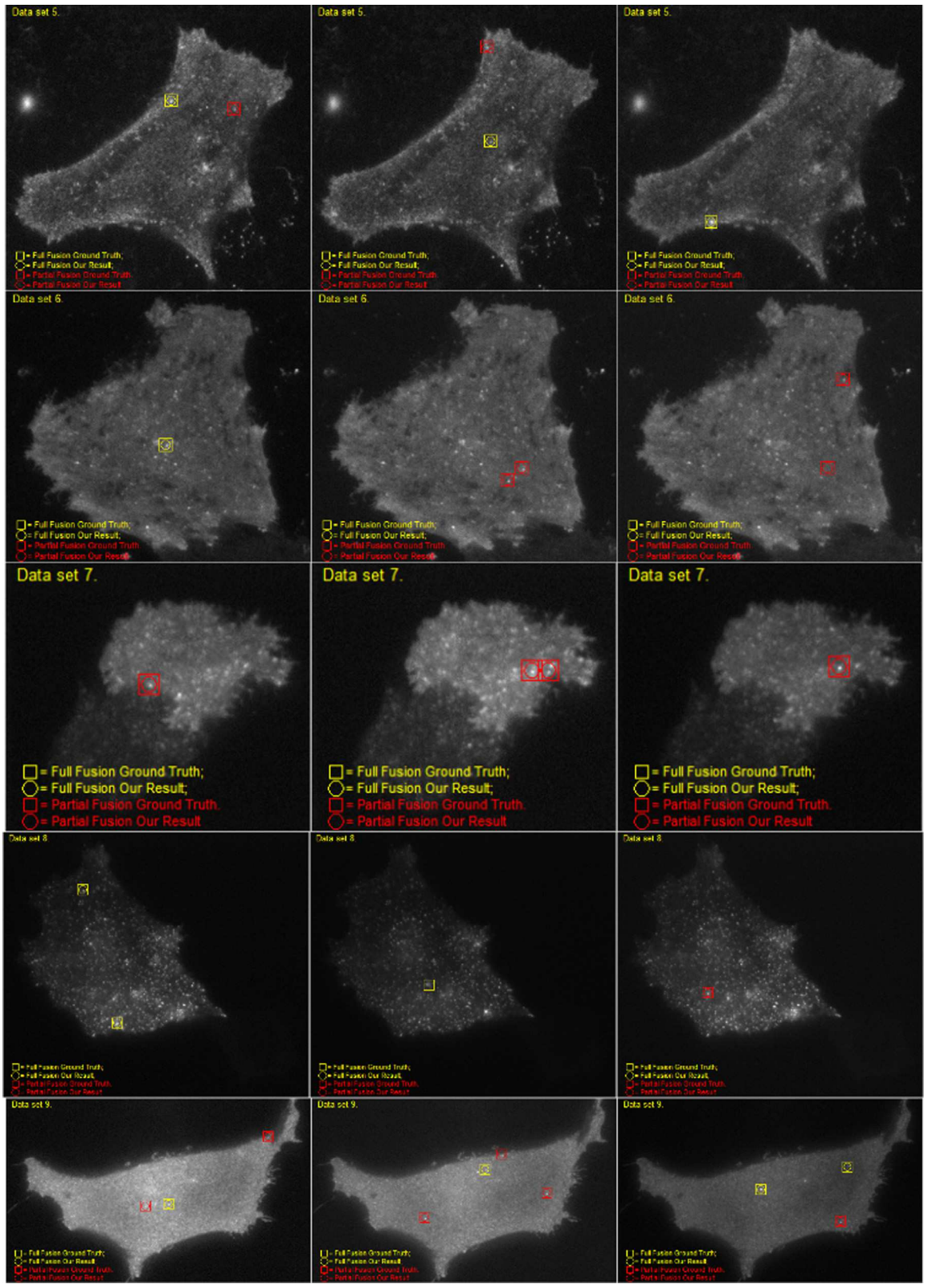

Fig. 16. Fusion event classification samples of dataset 5, 6, 7, 8, 9 (yellow: full fusion; red: partial fusion; square: ground truth; circle: our result). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

key moments of the process of a fusion event, a HCNN architecture is proposed to classify fusion event candidates into three classes: full fusion, partial fusion and non-fusion. Our method showed its competitive performance and outperformed our previous work, two state-of-the-arts and eleven alternative neural network architectures on nine challenging datasets with low signal to noise ratio and frequent background fluctuations.

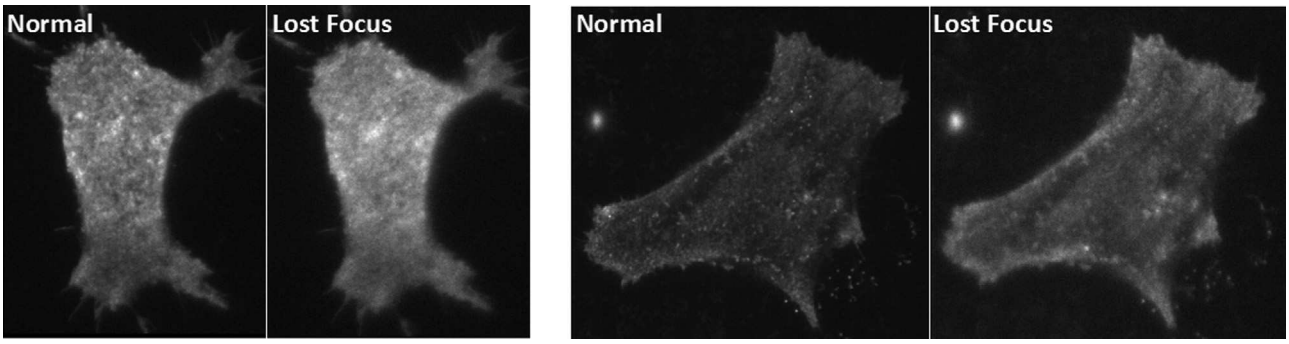

Fig. 17. The TIRFM image samples which are affected by out-of-focus.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at 10.1016/j.compmedimag.2017.04.003.

References

Axelrod, D., et al., 1981. Cell-substrate contacts illuminated by total internal reflection fluorescence. J. Cell Biol., 141–145.

Bai, L., et al., 2007. Dissecting multiple steps of GLUT4 trafficking and identifying the sites of insulin action. Cell Metab. 5 (1), 47–57.

Basset, A., et al., 2014. Localization and classificaiton of membrane dynamic in TIRF microscopy image sequences. International Symposium on Biomedical Imaging, 830–833.

Basset, A., et al., 2015. Detection and estimation of membrane diffusion during exocytosis in TIRFM image sequences. International Symposium on Biomedical Imaging.

Bornemann, A., et al., 1992. Subcellular localization of GLUT4 in nonstimulated and insulin-stimulated soleus muscle of rat. Diabetes, 215–221.

Dosset, P., et al., 2015. Automatically identifying fusion events between GLUT4 storage vesicles and the plasma membrane in tirf microscopy image sequences. Comput. Math. Methods Med., 1–7.

Gero, M., et al., 1998. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. Nature, 192–195.

Godinez, W., et al., 2009. Identifying fusion events in fluorescence microscopy images. International Symposium on Biomedical Imaging, 1170–1173.

Godinez, W., et al., 2012. Identifying virus-cell fusion in two-channel fluorescence microscopy image sequences based on a layered probabilistic approach. IEEE Trans. Med. Imag., 1786–1808.

Hou, J., et al., 1997. Ins (endocytosis) and outs(exocytosis) of GLUT4 trafficking. Curr. Opin. Cell Biol. 19 (4), 466–473.

Hu, B., et al., 2014. Convolutional neural network architectures for matching natural language sentences. Adv. Neural Inf. Process. Syst., 2042–2050.

Huang, S., et al., 2007. Insulin stimulates membrane fusion and GLUT4 accumulation in clathrin coats on adipocyte plasma membranes. Mol. Cell. Biol. 27 (9), 3456–3469.

Jahn, A., et al., 2012. Molecular machines governing exocytosis of synaptic vesicles. Nature, 201–207.

Karpathy, A., et al., 2014. Large-scale video classification with convolutional neural networks. IEEE conference on Computer Vision and Pattern Recognition, 1725–1732.

Kim, Y., et al., 2014. Convolutional Neural Networks for Sentence Classification, arXiv preprint arXiv:1408.5882.

Krizhevsky, A., et al., 2012. ImageNet classification with deep convolutional neural networks. Neural Inf. Process. Syst., 1097–1105.

Lawrence, S., et al., 1997. Face recognition: a convolutional neural-network approach. IEEE Trans. Neural Netw., 98–113.

Leney, S., et al., 2009. The molecular basis of insulin-stimulated glucose uptake: signalling trafficking and potential drug targets. J. Endocrinol. 203 (1), 1–18.

Li, H., et al., 2015. A Gaussian mixture model for automated vesicle fusion detection and classification. Comput. Methods Mol. Imag.

Lorenz, B., et al., 2010. A diffusion model for detecting and classifying vesicle fusion and undocking events. International Conference on Medical Image Computing and Computer-Assisted Intervention, 329–336.

Mao, Y., et al., 2016a. A hierarchical convolutional neural network for mitosis detection in phase-contrast microscopy images. International Conference on Medical Image Computing and Computer-Assisted Intervention, 685–692.

Mao, Y., et al., 2016b. A deep convolutional neural network trained on representative samples for circulating tumor cell detection 2016. IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 1–6.

Martin, F., et al., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM, 381–395.

Rizzoli, S., et al., 2007. Kiss-and-run, collapse and 'Readily Retrievable' vesicles. Traffic, 1137–1144.

Schneckenburger, H., et al., 2005. Total internal reflection fluorescence microscopy: technical innovations and novel applications. Curr. Opin. Cell Biol., 13–18.

Smith, B., et al., 2011. Interactive, computer-assisted tracking of speckle trajectories in fluorescence microscopy: application to actin polymerization and membrane fusion. Biophys. J., 1794–1804.

Srivastava, N., et al., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res., 1929–1958.

Vallotton, P., et al., 2007. Towards fully automated identification of vesiclemembrane fusion events in TIRFM. AIP Proceedings, 3–10.

Xu, Y., et al., 2011. Dual-mode of insulin action controls GLUT4 vesicle exocytosis. J. Cell Biol. 193 (4), 643–653.

Xu, Y., et al., 2016. Optogenetic activation reveals distinct roles of PIP3 and Akt in adipocyte insulin action. J. Cell Sci., 2085–2095.

Yuan, T., et al., 2015. Spatiotemporal detection and analysis of exocytosis reveal fusion "Hotspots" organized by the cytoskeleton in endocrine cells. Biophys. J., 251–260.

Yue-Hei, J., et al., 2015. Beyond short snippets: deep networks for video classification. IEEE Conference on Computer Vision and Pattern Recognition, 4694–4702.

Further reading

Chessel, A., et al., 2001. A detection-based framework for the analysis of recycling in TIRF microscopy. International Symposium on Biomedical Imaging 1281-1284.

Dosset, P., et al., 2016. Automatic detection of diffusion modes within biological membranes using backpropagation neural network. BMC Bioinform.

Matthew, S., et al., 2011. Interactive, computer-assisted tracking of speckle trajectories in fluorescence microscopy: application to actin polymerization and membrane fusion. Biophys. J. 1794-1804.