

A Wavelet-Based Bayesian Framework for 3D Object Segmentation in Microscopy

Kangyu Pan^a, David Corrigan^a, Jens Hillebrand^b, Mani Ramaswami^b and Anil Kokaram^a

^aElectronic & Electrical Engineering Department, Trinity College Dublin, Dublin-2, Ireland.

^bInstitute for Neuroscience, Trinity College Dublin, Dublin-2, Ireland. *

ABSTRACT

In confocal microscopy, target objects are labeled with fluorescent markers in the living specimen, and usually appear with irregular brightness in the observed images. Also, due to the existence of out-of-focus objects in the image, the segmentation of 3-D objects in the stack of image slices captured at different depth levels of the specimen is still heavily relied on manual analysis. In this paper, a novel Bayesian model is proposed for segmenting 3-D synaptic objects from given image stack. In order to solve the irregular brightness and out-of-focus problems, the segmentation model employs a likelihood using the luminance-invariant ‘wavelet features’ of image objects in the dual-tree complex wavelet domain as well as a likelihood based on the vertical intensity profile of the image stack in 3-D. Furthermore, a smoothness ‘frame’ prior based on the a priori knowledge of the connections of the synapses is introduced to the model for enhancing the connectivity of the synapses. As a result, our model can successfully segment the in-focus target synaptic object from a 3D image stack with irregular brightness.

1. INTRODUCTION

Formation of long term memory requires new protein synthesis at specific synapses of the neuron cells. The investigations of the memory formation process require the knowledge of the distribution of the proteins involved in the synthesis as well as the shapes of the synapses where the protein synthesis occurs.^{1,2} The proteins have been successfully detected and abstracted from the observed images,³ even with merged clustering appearance in the images, by our spot detection model⁴ which is based on Gaussian mixture models. The goal of the algorithm proposed in this paper is segmenting the 3-D shapes of the synapses from a given stack of 2-D images (as shown in Fig.1) observed from the same living specimen but captured at different depth by a confocal microscope.

Typically, the synapses are labelled by immunostaining with fluorescent markers in a living specimen, and a fluorescent image of the synapses is captured by confocal microscopy. However, the fluorescent markers are usually distributed unevenly over the synapses and, as a result, the synaptic objects have irregular brightness in the observed image. Furthermore, due to the physical limitation of microscope and because the synapses are 3-D objects in the living specimens, each observed image slice contains objects that are out-of-focus at the corresponding depth level. These out-of-focus objects are usually dimmer than the in-focus objects in the same image. However, since the synaptic objects have irregular brightness in the observed images, focus problem becomes even more complicated, while the object captured in-focus may have a dimmer appearance in the image but at the same time an out-of-focus object is brighter.

Several automatic and semi-automatic segmentation systems in confocal microscopy have been developed that apply on different tools, e.g. the K-means clustering,⁵ the watershed method,⁶ the level-set framework⁷ and the histogram-based method.⁸ However, in our work, due to the co-existing of irregular brightness and focusing issue in this case, the existing techniques are unable to select the in-focus synaptic features from the given image stacks. Since the in-focus objects have sharper features in the image regardless of the local luminance, the key to resolving this problem is to consider the overall probability of the object based on both the features and intensity profile on the vertical dimension of the image stack. Therefore, we propose a segmentation model using a Bayesian framework consists of a wavelet-likelihood based on the object features and a local vertical intensity likelihood derived from the vertical intensity profile.

This work was supported by the Science Foundation Ireland PI Programme: SFI-PI 08/IN.1/I2112

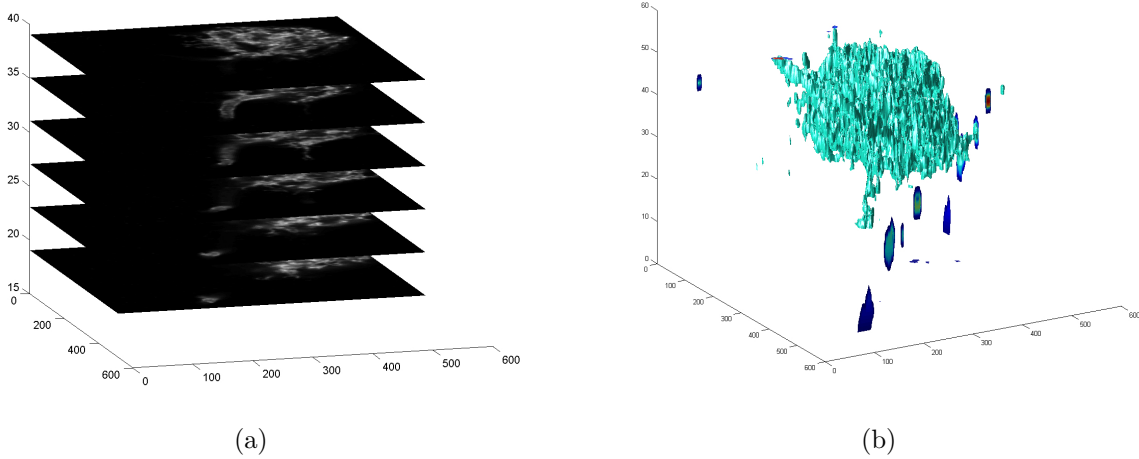


Figure 1. Sample of image stack and the result of 3D segmentation. (a) is a stack of 2-D image slices captured at different depth of a neuron specimen. (b) is a 3D illustration of the segmentation result of (a) using our model

In Section 2, we introduce the Bayesian frame work and the likelihoods are described in Section 3. The wavelet-likelihood considers the luminance-invariant ‘wavelet features’ of the synaptic objects in the wavelet domain through dual-tree complex wavelet transform,⁹ and the detail of description is given in Section 3.1. In addition, The wavelet-likelihood is co-operated with a global intensity likelihood (described in Section 3.2) in order to prevent faulty segmentation due to the existence of background noise with similar ‘wavelet features’ to the target synaptic objects. The local vertical intensity likelihood measures the probability of an image object at particular image slice being in-focus while considering the propagation of intensity of that local region through the image stack. This intensity likelihood term is discussed in Section 3.3. Furthermore, a priori is applied to the segmentation model in order to enhance the linkage of the connected synapses where the detail is given in Section 4. Then some segmentation results of our model are shown in Section 6.

2. BAYESIAN SEGMENTATION MODEL

Given the intensity information I of an image stack with M image slices, the goal of our Bayesian segmentation model is to assign labels to each slice by maximizing the probability

$$p(\mathbf{b}_1, \dots, \mathbf{b}_M | \mathbf{I}_1, \dots, \mathbf{I}_M) \propto p(\mathbf{I}_1, \dots, \mathbf{I}_M) p(\mathbf{b}_1, \dots, \mathbf{b}_M) \quad (1)$$

Note \mathbf{I}_s is the intensity matrix of image slice s while \mathbf{b}_s is the corresponding label field and the label ($b_s(\mathbf{x})$) for a pixel (\mathbf{x}) has the value

$$b_s(\mathbf{x}) = \beta = \begin{cases} 1 & \text{for } \mathbf{x} \text{ in 'object'} \\ 0 & \text{for } \mathbf{x} \text{ in 'background'} \end{cases}$$

In this work, we refer the pixels in the synapses as ‘object’ and all other pixels as ‘background’.

In order to make the optimization problem more tractable, we consider the optimization for each individual slice separately, Thus, we seek to maximize the posterior

$$p(\mathbf{b}_s | \mathbf{I}_1, \dots, \mathbf{I}_M, \mathbf{b}_1, \dots, \mathbf{b}_M) \propto p_L(\mathbf{I}_1, \dots, \mathbf{I}_M | \mathbf{b}_s) p_S(\mathbf{b}_s | \mathbf{b}_1, \dots, \mathbf{b}_M) \quad (2)$$

where $p_L(\cdot)$ and $p_S(\cdot)$ form the likelihoods and prior for the model respectively.

Furthermore, the assumption is made that the probabilities for each pixel in a given slice are independent and identically distributed (i.i.d.). Therefore, we rewrite the posterior as

$$p(\mathbf{b}_s | \mathbf{I}_1, \dots, \mathbf{I}_M, \mathbf{b}_1, \dots, \mathbf{b}_M) \propto \prod_{\mathbf{x}} \{p_L(I_1(\mathbf{x}), \dots, I_M(\mathbf{x}) | b_s(\mathbf{x})) p_S(b_s(\mathbf{x}) | b_1(\mathbf{x}), \dots, b_{s-1}(\mathbf{x}), b_{s+1}(\mathbf{x}), \dots, b_M(\mathbf{x}))\} \quad (3)$$

3. LIKELIHOODS

The expression for the likelihood in Eq.(3) is factorized further into 3 separate likelihoods given by

$$p_L(I_1(\mathbf{x}), \dots, I_M(\mathbf{x}) | b_s(\mathbf{x})) \propto [p_W(\mathbf{W}_s | b_s(\mathbf{x}))]^{\alpha_W(\beta)} \times [p_G(I_s(\mathbf{x}) | b_s(\mathbf{x}))]^{\alpha_G(\beta)} \times [p_V(I_1(\mathbf{x}), I_s(\mathbf{x}), \dots, I_M(\mathbf{x}) | b_s(\mathbf{x}))]^{\alpha_V(\beta)} \quad (4)$$

The first likelihood is the wavelet-based likelihood designed to classify the image features with synaptic texture as ‘object’ regardless to the luminance of the features in the given slice. The second global intensity likelihood is intended to discriminate against regions of low-intensity in order to remove the false alarms caused by noise features with similar texture to the synaptic objects. Finally, the vertical intensity likelihood is designed to detect in-focus features and is derived from the vertical intensity profile at a given location. Furthermore, these likelihood are controlled by the weights $\{\alpha_W(\cdot), \alpha_G(\cdot), \alpha_V(\cdot)\}$.

3.1 Wavelet-likelihood

Since the target synaptic object has irregular intensity in the image stack, a simple thresholding would be infeasible for the segmentation. Instead, we wish to classify pixels based on the luminance invariant texture content. The wavelet transform is ideally suited for this task since the discrete wavelet transform (DWT) decomposes an image into a luminance dependent low frequency band and a number of luminance invariant high frequency sub-bands. We propose a likelihood term using ‘wavelet features’ of image objects extracted from the high frequency sub-bands in the wavelet domain. The basic idea is to use manually labelled slice to train Gaussian mixture models (GMMs) of wavelet features for both ‘object’ and ‘background’. These GMMs provide the likelihood model for all other slices in the stack.

3.1.1 Dual-tree complex wavelet transform (DTCWT)

In this work, we use the DTCWT instead of the traditional DWT to abstract the ‘wavelet features’ from the given image slices. The DTCWT is introduced by Kingsbury *et al.*⁹ which has resolved the shift-variance problem and the deficiency of frequency orientation of the classic DWT, and hence provide stable wavelet features for small offsets in synapse location.

Like the basic DWT, the DTCWT is a multi-scale transform, where each scale of the transformed image consists of one low frequency band (lo-band) and six band-pass high frequency sub-bands (hi-bands). The lo-band is simply a down-sampled version of the input image whereas the hi-bands at a given scale correspond to the response of the image to the wavelet kernel at six different orientations ($\pm 15^\circ$, $\pm 45^\circ$ and $\pm 75^\circ$) in a frequency band dictated by the scale of the wavelet. The hi-band wavelet coefficients have complex values that indicate both the magnitude and phase responses of the wavelet at the corresponding orientation.

Using the DTCWT, the image feature at each pixel of the given image is then represented by the corresponding wavelet coefficients in the lo-band as well as the hi-bands of all the scales in the wavelet domain. This means that the ‘sharp’ and ‘thin’ image features would have large wavelet coefficients in the hi-bands and be approximately invariant to the low frequency intensity features which are encoded in the lo-band. We call these wavelet coefficients in the hi-bands the ‘wavelet features’ of the corresponding image objects, and these features are invariant to the global intensity and are sensitive to the shape and thickness of the object features.

3.1.2 Wavelet features and the Gaussian mixture model

Based on the DTCWT metric, we introduce the wavelet-likelihood that estimates the likelihood probability of each pixel in the image stack based on the ‘wavelet features’ associated to that pixel. Here, the ‘wavelet feature’ of a pixel is a 12-D vector for one scale of the DTCWT that contains the entries of the corresponding complex wavelet coefficients from all the 6 hi-bands at that scale, and denote as $W_{s,h}(\mathbf{x})$ at scale h for pixel \mathbf{x} at slice s .

We assume the ‘object’ wavelet features are well separated from the ‘background’ wavelet features at each scale in the DTCWT domain, which means the clusters of ‘object’ wavelet features are distant from the clusters of ‘background’ wavelet features. Here, we fit a Gaussian kernel to each of cluster, and hence the ‘object’ wavelet features and ‘background’ wavelet features at each scale in the DTCWT are fitted with two different Gaussian mixture models (GMMs). Therefore, the wavelet-likelihood probability $p_w(\mathbf{W}_s|b_s(\mathbf{x}))$ in Eq.(4) of pixel k at slice l being the ‘object’ (or ‘background’) is the product of the density values of the ‘object’ (or ‘background’) GMMs for the corresponding wavelet-features of all the scales, and the

$$p_w(\mathbf{W}_s|b_s(\mathbf{x})) = \prod_{h=1}^H G(W_{s,h}(\mathbf{x})|\Theta(\beta, h)) \quad \beta \in [0, 1] \quad (5)$$

where H is the number of scales of the DTCWT being used in the model, $G(\cdot)$ is the GMM with the parameter set $\Theta(\cdot)$ for label β at scale h . In this work, we use the first 2 scales of the DTCWT. Furthermore, in practice, since hi-bands at lower frequency scales have lower resolutions, the pixels that share the same wavelet coefficients at lower frequency scales would also share the same GMM values at that scale.

3.1.3 GMM initialization

In order to generate the GMMs, a set of segmented training data is manually selected from same sample images by the ground-truth evaluation of biological experts. For each segmented image slice, a binary mask of the ‘object’ pixels and the background pixels is provided. These binary masks are used to generate a set of down-sampled masks that have resolutions corresponding to the resolution of the equivalent DTCWT scale. The training wavelet features of each DTCWT scale are extracted from the sample images by performing the DTCWT and by labelling each location as ‘object’ or ‘background’ according to the value of the binary mask at that location and scale. An initial parameter estimation for the ‘object’ or ‘background’ GMM of each scale is first generated from the extracted ‘object’ or ‘background’ training features using the mean-shift clustering algorithm.¹⁰ The GMM parameters are then refined by applying the expectation maximization (EM) algorithm⁴ to the initial estimation.

However, as shown in Fig.2(c), due to the existence of the background noise in the image stack (as shown in the ‘noise boosted’ image Fig.2(b)), it is possible to get false alarms in the segmentation (Fig.2(c)) if only the wavelet likelihood is used. This occurs because some noise features may have wavelet features similar to the object regions. In this work, we introduce the global intensity likelihood as a means of discriminating against low intensity noise regions.

3.2 Global Intensity Likelihood

From Fig.2(a&b), we can see that the intensity of the background noise is close to the background brightness and is significantly different from the object brightness. Also, as shown in Fig.3, the ‘background’ intensity and the ‘object’ intensity are separable from each other in the histogram. Intuitively, the ‘object’ and ‘background’ intensity in the image can be fitted by two different Gaussian-pdfs. The parameters of the pdfs are estimated by fitting two Gaussian pdfs separately to the ‘object’ and ‘background’ pixels collected from the segmented training images mentioned in Section 3.1.3. However, a pixel with higher intensity value is overwhelmingly likely to be the ‘object’, and conversely, a pixel is likely to be the ‘background’ when the intensity is low. Since Gaussian pdf for the object will be small for high intensities, we instead use two Gaussian cumulative density functions (cdf) to estimate the ‘relative probabilities’ of a given pixel being ‘object’ or ‘background’ as follows

$$p_G(I_s(\mathbf{x})|b_s(\mathbf{x})) = \begin{cases} 1 - \text{cdf}(I_s(\mathbf{x})|\theta_G(\beta = 0)) \\ \text{cdf}(I_s(\mathbf{x})|\theta_G(\beta = 1)) \end{cases}$$

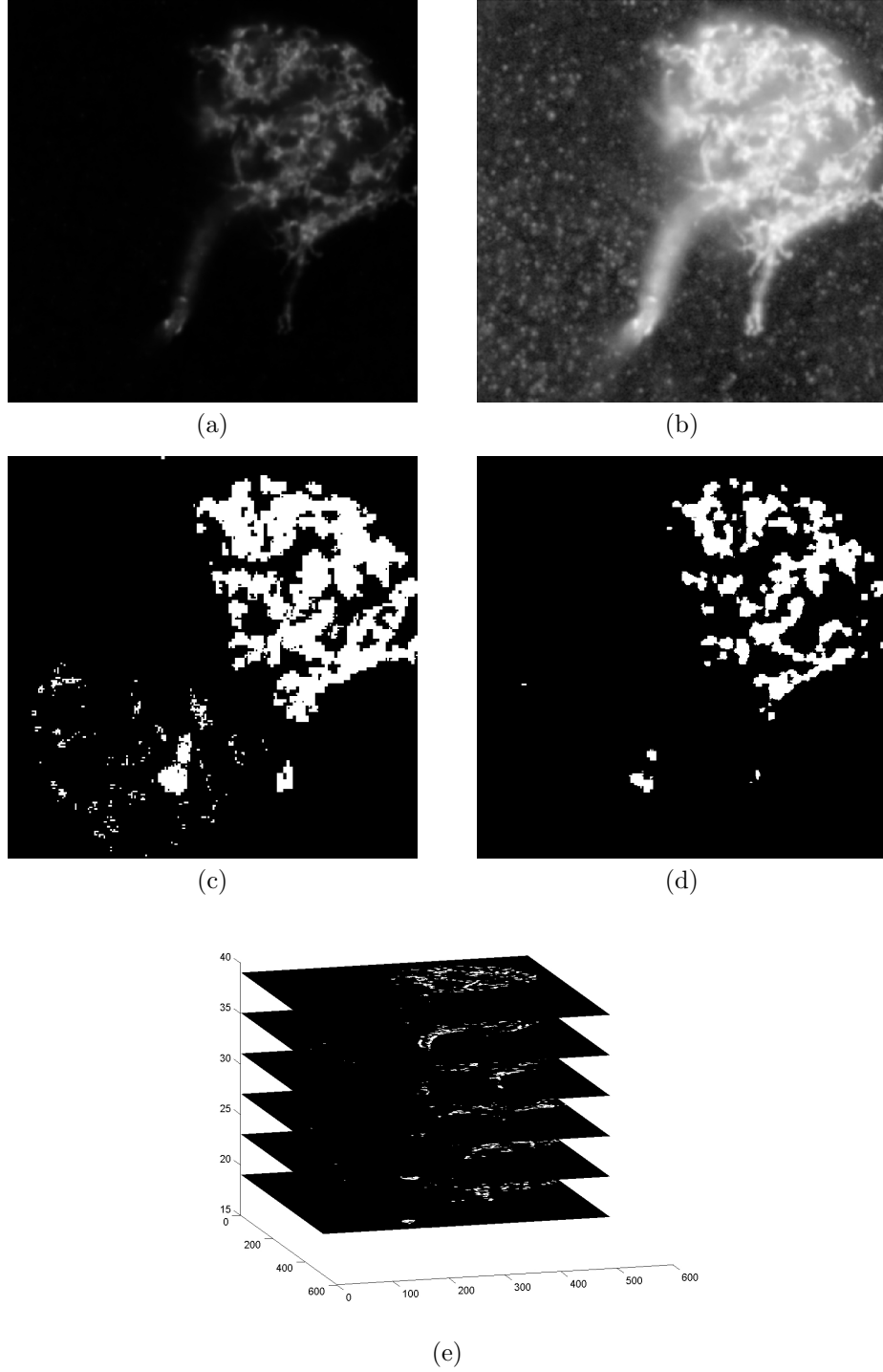


Figure 2. Segmentations based on Maximum Likelihood estimation. (a) is an image at slice $s = 30$ of the stack in Fig.1. (b) is 'noise boosted' version of (a) using a gamma operation. (c) is the segmentation result obtained by Maximum Likelihood estimation base only on the wavelet-likelihood $p_W(\cdot)$. (d) is the segmentation result obtained by Maximum Likelihood estimation base on both $p_W(\cdot)$ and the global intensity likelihood $p_G(\cdot)$. (e) is the stack of the segmentation result obtained by Maximum Likelihood estimation of every image slice in the stack, and used as the initial segmentation for our model.

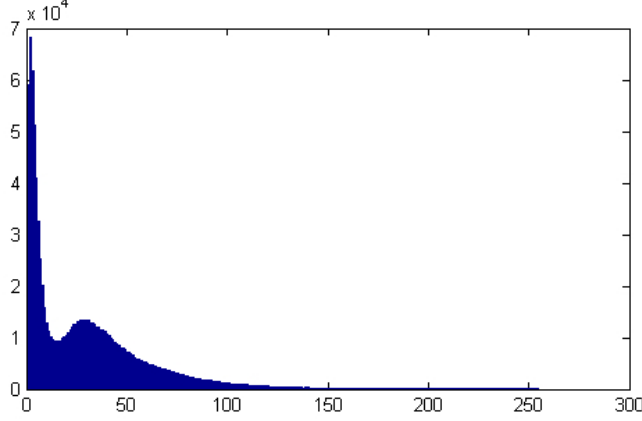


Figure 3. Histogram of global intensity of the image stack. *There are two peaks in the histogram where the peak with a mean at approximate 40 belongs to ‘object’ and the other narrower peak belongs to ‘background’. We model both peaks with Gaussian-pdfs.*

(6)

where $\theta_G(\beta) = \{\mu_G(\beta), \sigma_G^2(\beta)\}$ are the mean and the variance of the fitted Gaussain-pdf for ‘object’ or ‘background’ intensity. The result of the segmentation model with both the wavelet-likelihood and the global intensity likelihood is given in Fig.2(d), and obviously, the background noise features are removed while comparing to Fig.2(c) which is based on the wavelet-likelihood only.

The combination of the wavelet-likelihood and the global intensity likelihood forms the intra-slice likelihood in Eq.(4), and solves the irregular brightness problem while rejecting low intensity noise textures. The vertical likelihood in Eq.(4) that overcomes the out-of-focus problem is described in the following section.

3.3 Local Intensity Likelihood

The 2nd challenge of the synaptic object segmentation is the presence of out-of-focus features in the image stack. The out-of-focus features are usually dimmer, thinner and more smooth than the in-focus features, however a synaptic object can have ‘branches’ with different thickness in the images and also, as discussed in Section 1, the fluorescent material is distributed unevenly over the synapse and hence appears with irregular brightness in the images. In other words, a relatively dimmer and thinner feature can be in-focus in the given image slice while a bright and thick feature is out-of-focus. Consequently, the determination of whether an object feature is in-focus or out-of-focus can not rely only on a single image slice, but instead, we should look through the image stack (as shown in Fig.4(a)) at the same local region of the object feature in the given image. Therefore, the determination is made based on local inter-slice information of the image stack rather than each individual horizontal planes. This key idea is injected to the Bayesian segmentation model as another likelihood term is capable of distinguishing the in-focus and out-of-focus features in a given image stack.

3.3.1 Vertical intensity profile

As shown in Fig.4, while considering the intensity distribution of a single point through the image stack vertically, the feature that is in-focus on a particular image slice corresponds to the highest value in the vertical intensity distribution (the ‘blue’ plot in Fig.4(b)), and the value decreases while the projection of the feature becomes more out-of-focus on the slices below or above the in-focus slice. Also, we assume that the value of the vertical intensity distribution will remain high while the feature is in-focus on multiple adjacent image slices. This happens when an object has a vertical orientation in the specimen. Consequently, the local intensity likelihood p_v of pixel \mathbf{x} is estimated by a Gaussian pdf as following

$$p_V(I_1(\mathbf{x}), I_s(\mathbf{x}), \dots, I_M(\mathbf{x})|b_s(\mathbf{x})) = \begin{cases} 1 & , \beta = 0 \\ \exp\left[-\frac{(I_s(\mathbf{x}) - \mu_v(\mathbf{x}, s))^2}{2\sigma_v^2}\right] & , \beta = 1 \end{cases}$$

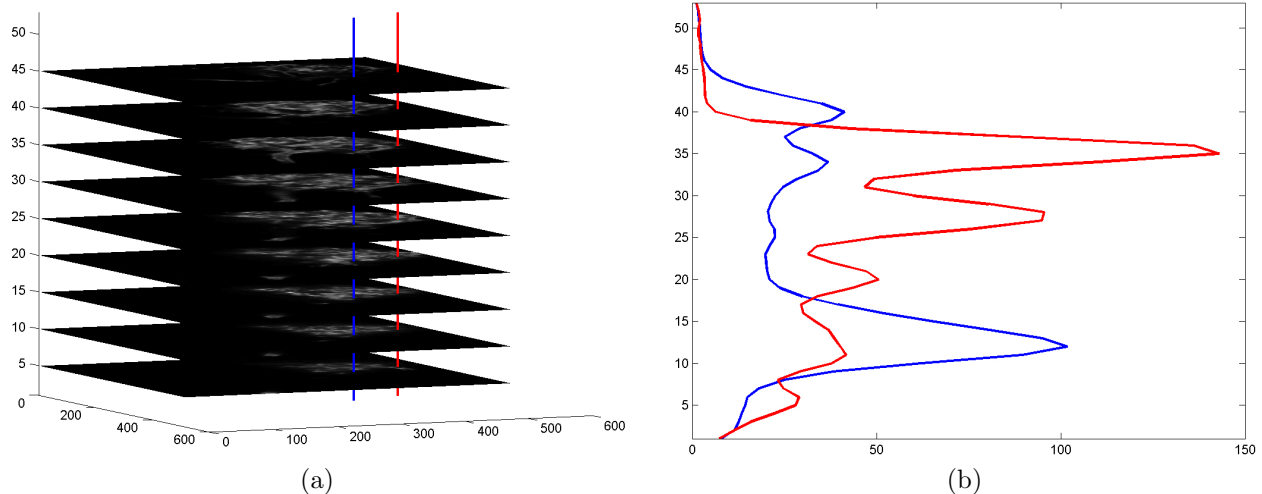


Figure 4. Local vertical intensity profiles. (a) is the image stack in Fig.1 but highlighted by two vertical inter-slice lines at different pixel-wise locations. (b) shows the vertical intensity profiles corresponding to the locations labeled in (a)

(7)

where the mean $\mu_v(\mathbf{x}, s)$ of the Gaussian-pdf is the maximum value of the local vertical intensity distribution of the same pixel site \mathbf{x} through the image slices. Since the intensity profile due to focus is assumed to be the same for all features in the same image stack, we use a fixed variance $\sigma_v^2 = 2$ for Gaussian-pdf. Note that the goal of the local intensity likelihood is to select the in-focus feature from the present of the out-of-focus version of the feature, hence this likelihood term is only applied to the ‘object’ probability and is ‘1’ for the ‘background’ probability.

In practice, when multiple features vertically coincide in an image stack, the vertical intensity profile of that local region would instead have multiple separate peaks in the intensity profile (‘red’ plot in Fig.4(b)). In these cases, we fit different Gaussian-pdfs to each peak. We find the peaks using the intra-slice likelihood (the combined wavelet-likelihood and the global intensity likelihood) to give a rough segmentation of the image stack into dark or bright regions.

4. SMOOTHNESS FRAME PRIOR

So far, the likelihood terms have enabled the Bayesian model to overcome the irregular-brightness and out-of-focus problems. In fact, a single synapse is an object with connected ‘branches’ in the specimen, and therefore the ‘object’ features should also be connected as the ‘branches’ in the image stack. However, the likelihood terms described in the previous sections do not include any a priori knowledge of the smoothness properties of the object features in the image stack. For the cases that some synaptic ‘branches’ are stained with little fluorescent material in the specimen, the ‘weak’ appearance of those ‘branches’ in the image stack makes it likely to be classified as ‘background’ by the likelihood terms of the model. As a result, the object of a single synapse in the image stack may be ‘broken’ into several disconnected objects after segmentation.

In order to enhance the connection of such ‘weak’ branches, we inject a spatial smoothness prior to the segmentation model. This idea of this prior is to label a pixel as ‘object’ or background based on the label of the neighboring pixels. The smoothness is constrained by the ‘frame’ of the synaptic object, which can be considered as the skeleton of the connected ‘branches’ in the image stack.

4.1 Synaptic object frame

To obtain the ‘frame’ of the synaptic object, a method (as shown in Fig.5) similar to the pyramid algorithm^{11,12} is applied to a given image stack. In¹² the pyramid algorithm is used to detect brush strokes in paintings. The image stack is ‘blurred’ to coarse versions at different scales and placed as a pyramid where the original image

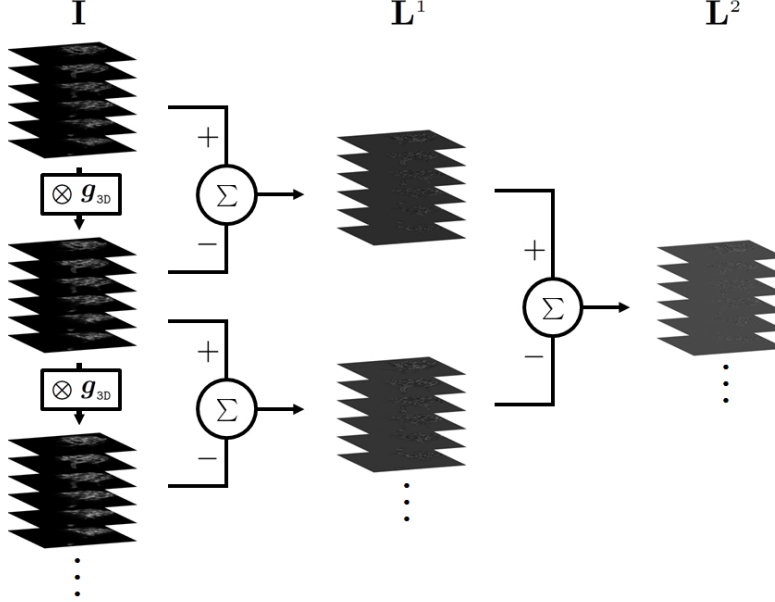


Figure 5. Illustrates the process of the Laplacian approach. *The left column shows the image stack with different ‘resolutions’ by convolving with a 3D Gaussian kernel g_{3D} . The stacks are arranged from fine to coarse. The middle column contains the 1st order Laplacian features which obtained by abstracting the coarse image stack from the ‘finer’ one ‘above’ it. The right column consists of the 2nd Laplacian features that obtained by subtracting the 1st order Laplacian features at different resolution levels.*

stacks with finest features is placed on top of the pyramid and the coarsest version is located at the bottom. Here, a coarse image stack is generated by convolving the given image stack with a 3-D Gaussian like kernel, i.e. a ‘low-pass’ filter, and this coarse image stack is used to generate a coarser version at higher scale by convolving with the same ‘low-pass’ filter. A Laplacian approach is applied to the image pyramid to get the 1st order Laplacian features by subtracting a coarse image stack from the adjacent finner image stack above it, and as a result, a pyramid of the 1st order Laplacian features at different scale is produced.

The frame of the synaptic object is then extracted from the Laplacian features by removing the negative values and filtering the positive values with the raw object mask (Fig.6(d)) mentioned in Section 3.3.1. A sample slice of the 1st order Laplacian features (Fig.6(b)) at scale 6 and the corresponding extracted object frame are given in Fig.6(e), however this frame seems too coarse for the fine details of the synaptic object on the image. Therefore, in order to get ‘finer’ frame of the thin objects, we perform the Laplacian approach again but on the 1st order Laplacian features and do the abstraction on the resulted 2nd order Laplacian features (Fig.6(c)). A sample of the resulted finer frame for the corresponding thin synaptic branches is shown in Fig.6(f).

4.2 Frame prior

Since the ‘likely’ connections between the synaptic objects are now identified by the object frame, the a priori information of a given pixel x at image slice s derived from pixels in a 3D neighborhood \mathbf{N}_{3D} in the image stack should be constrained by a binary object mask of the frame corresponding to the image stack, and hence the prior term is called ‘frame prior’.

The frame prior is designed to augment the standard isotropic smoothness prior by boosting smoothness along the axis of the synapse frame while penalizing smoothness perpendicular to the frame. In order to avoid over-biasing the segmentation B toward the frame B_F , the prior is designed in the way that, when the frame B_F and β agree, smoothness is suspended across the edges of the synapse frame, otherwise the standard isotropic smoothness is applied. Therefore, the ‘frame prior’ in Eq.(3) is expressed as

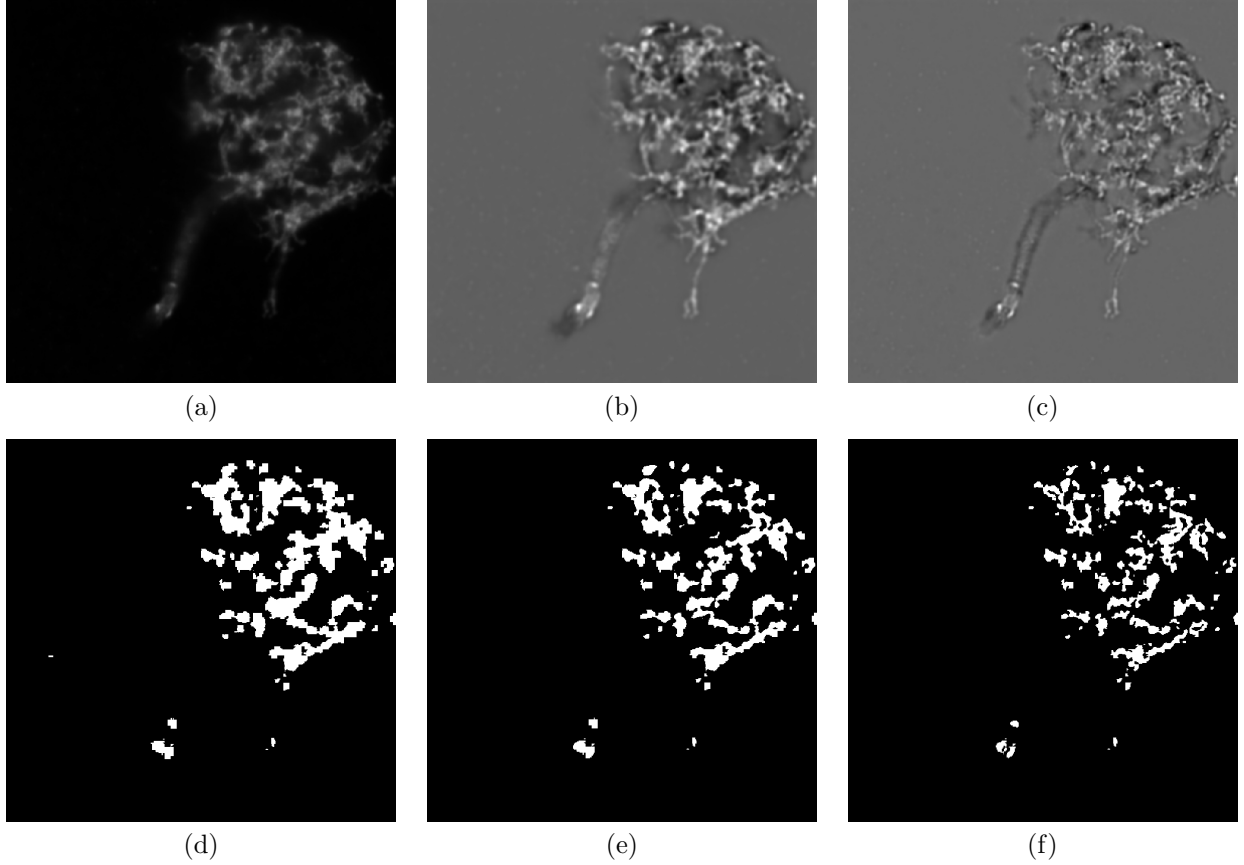


Figure 6. Illustration of the Laplacian process by slice. (a) is an image slice of the original image stack in Fig.1. (b) is the 1st order Laplacian features of (a) at resolution 'level 6'. (c) is the 2nd order Laplacian features of (b). (d) is the initial segmentation \mathbf{b}_s^0 for (c). (e) is the 1st order object frame of (a) by thresholding the positive 1st order Laplacian features in (b) and filtered by (d). (f) is the 2nd order object frame of (a) by thresholding the positive 2nd order Laplacian features in (c) and filtered by (d), it is also the synaptic object frame we use in this work.

$$p_S(b_s(\mathbf{x})|b_1(\mathbf{x}), \dots, b_{s-1}(\mathbf{x}), b_{s+1}(\mathbf{x}), \dots, b_M(\mathbf{x})) = \begin{cases} \lambda \sum_{\mathbf{y}_i \in \mathbf{N}_{3D}} \{\omega_i |1 - B_F(\mathbf{y}_i)| |1 - b_s(\mathbf{y}_i)|\} & , B_F(\mathbf{x}) = 0 \text{ \& } \beta = 0 \\ (1 - \lambda) \sum_{\mathbf{y}_i \in \mathbf{N}_{3D}} \{\omega_i |1 - b_s(\mathbf{y}_i)|\} & , B_F(\mathbf{x}) = 1 \text{ \& } \beta = 0 \\ (1 - \lambda) \sum_{\mathbf{y}_i \in \mathbf{N}_{3D}} \{\omega_i |b_s(\mathbf{y}_i)|\} & , B_F(\mathbf{x}) = 0 \text{ \& } \beta = 1 \\ \lambda \sum_{\mathbf{y}_i \in \mathbf{N}_{3D}} \{\omega_i |B_F(\mathbf{y}_i)| |b_s(\mathbf{y}_i)|\} & , B_F(\mathbf{x}) = 1 \text{ \& } \beta = 1 \end{cases}$$

$$\omega_i = \frac{|\mathbf{x} - \mathbf{y}_i|}{\sum_{\mathbf{y}_i \in \mathbf{N}_{3D}} |\mathbf{x} - \mathbf{y}_i|} \quad (8)$$

where ω_i is a set of normalized neighbor distances. Furthermore, the ‘strength’ of ‘frame constrained’ part is controlled by the ratio $\lambda \in (0, 1)$ in order to vary the ‘strength’ of the connections between the synapses.

5. OPTIMIZATION

The Bayesian framework using the proposed likelihoods and prior is optimized using the Iterated Conditional Modes (ICM) process.¹³ To initialize the optimization, an initial estimate of the segmentations for each slice, \mathbf{b}_s , are evaluated using the Maximum Likelihood (ML) estimate. We then apply the ICM procedure to update each slice segmentation, \mathbf{b}_s , in turn by finding the *Maximum a Posteriori* (MAP) estimate from Eq.(3). When estimating the prior probability at each iteration, a fixed estimate of the synapse frame $B_F(\mathbf{x})$ calculated at the start of the process is used. We use a total of 5 iterations over the entire stack. From our experience, this allows for the segmentation to converge sufficiently. Furthermore, in our experiments, the weights $\{\alpha_w(\cdot), \alpha_G(\cdot), \alpha_V(\cdot)\}$ in Eq.(4) are set as $\alpha_w(0) = 1.6$, $\alpha_w(1) = 1$, $\alpha_G(0) = 1$, $\alpha_G(1) = 2$, $\{\alpha_V(0) = 1$, $\alpha_V(1) = 1$ and λ is set to 0.8.

6. RESULTS

The 3D segmentation result of the a given sample image stack is given in Fig.1(b). As it is difficult to visualize a 3D segmentation, we have selected small samples of the image stack and illustrated the segmentations for a number of slices in the stack (Fig.7-Fig.10). Although this representation is not ideal for viewing the vertical relationships in the segmentation masks, it does allow us to demonstrate the effectiveness of aspects of our algorithm.

As shown in Fig.7 to Fig.10, our Bayesian mode has successfully segmented the in-focus synaptic features in both bright regions dim regions. For example, in Fig.7, the bright feature on the left of ‘slice 9’ is selected while the dimmer sharp in-focus features in ‘slice 15’ are also detected. The effectiveness of the vertical likelihood in detecting the in-focus regions can be seen in Fig.8. As shown in center of ‘slice 4’ to ‘slice 14’, when the object becomes brighter, there are more in-focus features being detected, but when it becomes dimmer in the later slices, there are less features being selected. However in some slices, like ‘slice 17’ in Fig.8 and ‘slice 44’ in Fig.9, the in-focus constraint is too strict so that the ‘relatively’ in-focus features are unselected. The reason of this ‘false negative’ is that we use a fixed variance $\sigma_v^2 = 2$ for the vertical inter-slice intensity profile. A potential approach to improve in-focus detection is using an adaptive variance. Furthermore, effectiveness of the frame prior is can seen in ‘slices 33 and 34’ in Fig.10, the in-focus features with ‘blob’ appearance would be detected as individual spots by only the likelihoods but are detected as elongated synaptic features.

In order to achieve optimum segmentation results, the likelihoods can be adjusted by adding different regularizing weights, so that the ‘ratio’ between the wavelet-likelihood and the global intensity likelihood can be tuned for the image stacks with severe irregular brightness problem, or the ‘judgment’ of out-of-focus feature can be changed. The smoothness ratio λ in the frame prior can also be adjusted in order to vary the strength of the connection between the synaptic objects. However, the parameter values outlined in Section 5 were found to give the best result over the entire stack.

7. CONCLUSION

In this paper, we introduce a Bayesian model for segmenting 3D synaptic objects from a given stack of specimen image slices captured using a fluorescent confocal microscope. The key intra-slice wavelet-likelihood provides a degree of luminance invariance in order to detect the synaptic features with irregular brightness. The other core likelihood based on the vertical inter-slice intensity profile enables the model to distinguish between in-focus and out-of-focus features. Furthermore, a novel frame prior is introduced to direct spatial smoothness along the skeleton of the synapses. Our model gives a plausible segmentation of synapses that requires manual segmentation in only one slice to train the Gaussian mixture models used for estimating the wavelet-likelihood. Further work on this problem will focus on using graph-cut optimization with our Bayesian framework. Using graph-cuts is advantageous as it finds the globally optimal labelling for a binary segmentation problem and hence is insensitive to any initialization.

REFERENCES

1. Barbee, S. A., Estes, P. S., Cziko, A.-M., Hillebrand, J., Luedeman, R. A., Collier, J. M., Johnson, N., Howlett, I. C., Geng, C., Ueda, R., Brand, A. H., Newbury, S. F., Wilhelm, J. E., Levine, R. B., Nakamura, A., Parker, R., and Ramaswami, M., “Staufen- and fmprp-containing neuronal rnps are structurally and functionally related to somatic p bodies,” *Neuron* **52**(6), 997–1009 (2006).
2. Bramham, C. R. and Wells, D. G., “Dendritic mrna: transport, translation and function,” *Nature reviews. Neuroscience* **8**(10), 776–789 (2007).
3. Hillebrand, J., Pan, K., Kokaram, A., Barbee, S., Parker, R., and Ramaswami, M., “The me31b dead-box helicase localizes to postsynaptic foci and regulates expression of a camkii reporter mrna in dendrites of drosophila olfactory projection neurons,” *Frontiers in Neural Circuits* **4**, 121 (2010).
4. Pan, K., Hillebrand, J., Ramaswami, M., and Kokaram, A., “Gaussian mixture models for spots in microscopy using a new split/merge em algorithm,” *IEEE International Conference on Image Processing (ICIP’10)*, 3645–3648 (2010).
5. Niemisto, A., Korpelainen, T., Saleem, R., Yli-Harja, O., Aitchison, J., and Shmulevich, I., “A k-means segmentation method for finding 2-d object areas based on 3-d image stacks obtained by confocal microscopy,” in *[Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE]*, 5559–5562 (aug. 2007).
6. Long, F., Peng, H., and Myers, E., “Automatic segmentation of nuclei in 3d microscopy images of c.elegans,” in *[Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on]*, 536–539 (april 2007).
7. Gouaillard, A., Mosaliganti, K., Gelas, A., Souhait, L., Obholzer, N., and Megason, S., “Streaming level set algorithm for 3d segmentation of confocal microscopy images,” in *[Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE]*, 3621–3624 (sept. 2009).
8. Garza-Jinich, M., Meer, P., and Medina, V., “Robust retrieval of three-dimensional structures from image stacks,” *Medical Image Analysis* **3**(1), 21–35 (1999).
9. Ivan W. Selesnick, R. G. B. and Kingsbury, N. G., “The dual-tree complex wavelet transform,” *IEEE Signal Processing Magazine* **22**(6), 123–151 (2005).
10. Comaniciu, D. and Meer, P., “Mean shift: A robust approach toward feature space analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, 603–619 (May 2002).
11. Burt, P. and Adelson, E., “The laplacian pyramid as a compact image code,” *Communications, IEEE Transactions on* **31**, 532–540 (apr 1983).
12. Hertzmann, A., “Painterly rendering with curved brush strokes of multiple sizes,” in *[Proceedings of the 25th annual conference on Computer graphics and interactive techniques]*, *SIGGRAPH ’98*, 453–460, ACM, New York, NY, USA (1998).
13. Besag, J., “On the statistical analysis of dirty pictures,” *Journal of the Royal Statistical Society. Series B (Methodological)* **48**(3), 259–302 (1986).

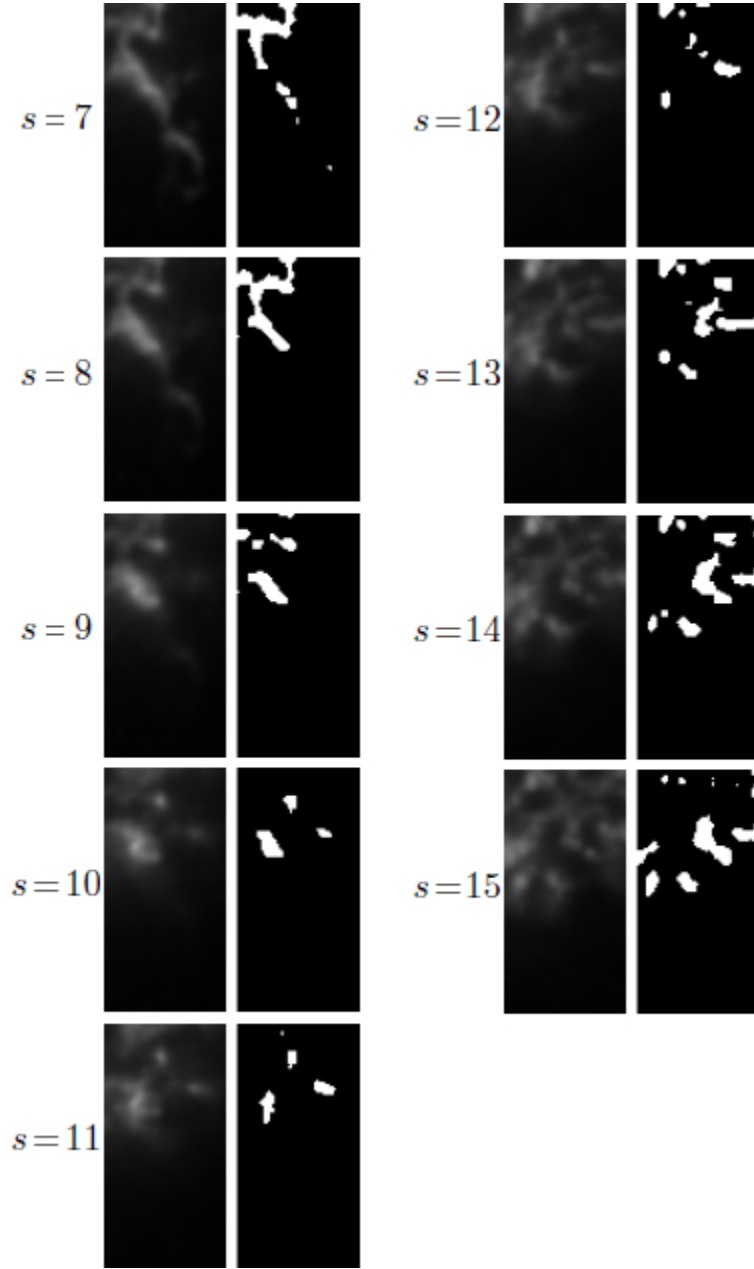


Figure 7. Results of Sample 1. *The left part of each image pair is the original image sample extracted from the corresponding slice ‘s’ of the image stack in Fig.1. The right part is the segmentation results of the left images. Through the slices, we can see that the both the bright and dim in-focus features are detected.*

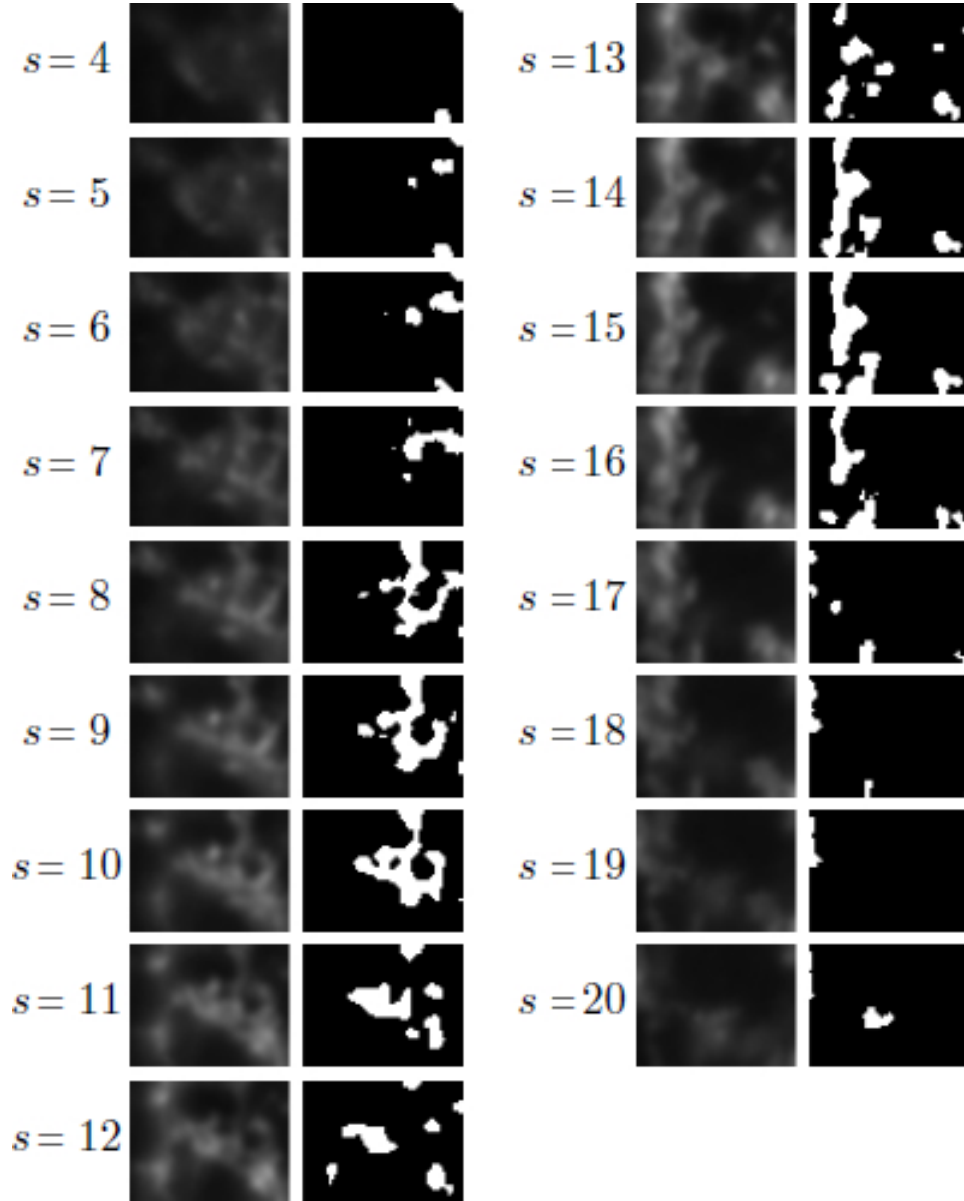


Figure 8. Results of Sample 2. *This sample shows that the proposed model can distinguish between in-focus and out-of-focus features. However, the in-focus constraint is too strict for the features in slice 17.*

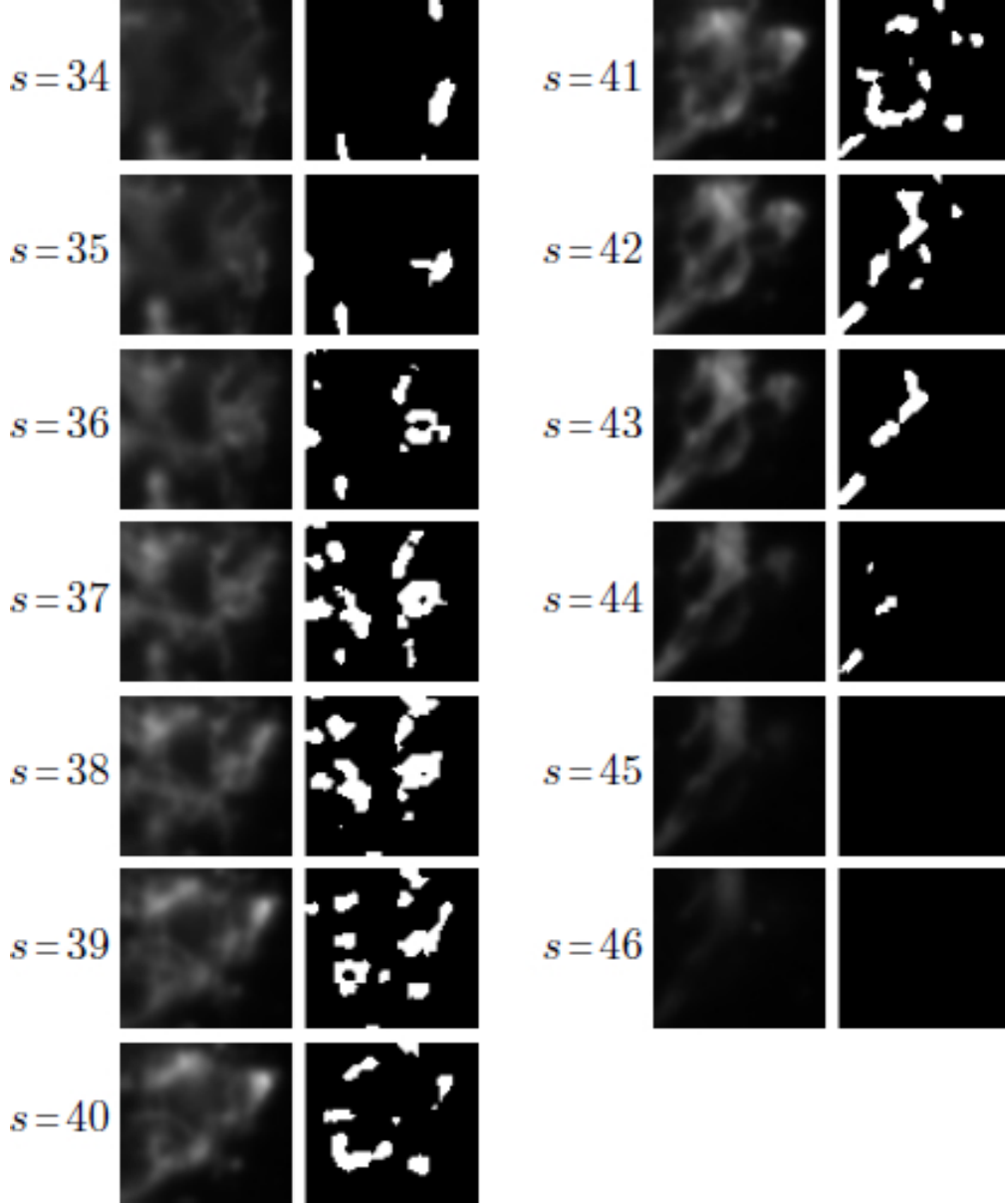


Figure 9. Results of Sample 3. *This sample also shows that the proposed model can detect in-focus features while the out-of-focus features are also presented. However, the in-focus constraint is too strict for the features in slice 36 – 39&44.*

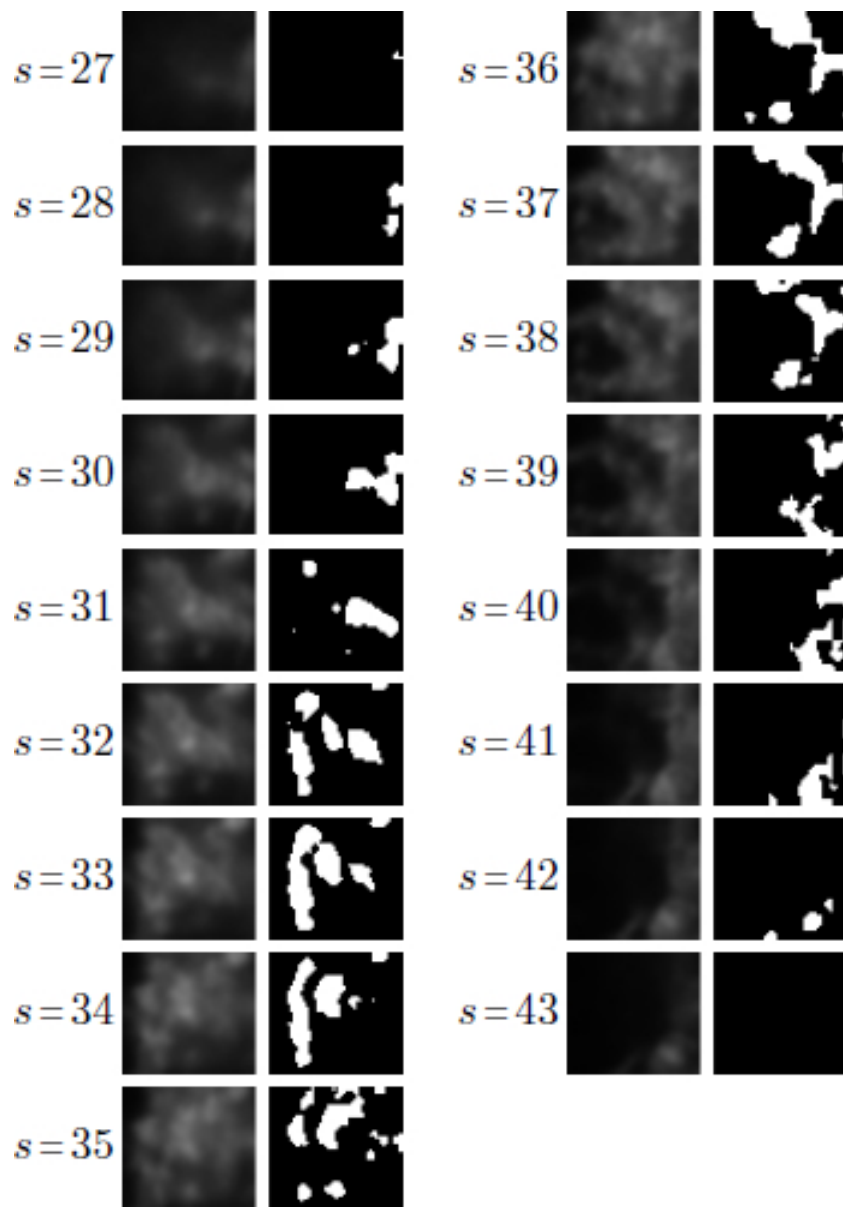


Figure 10. Results of Sample 4. *The sample shows the effectiveness of frame prior by detecting the in-focus features with ‘blob’ appearance as connected synaptic features in slice 33&34.*