PATTERNING MOTOR NEURONS IN THE DROSOPHILA VENTRAL NERVE CORD USING LATENT STATE CONDITIONAL RANDOM FIELDS

X. Chang[1], *M.D. Kim*[2], *A. Chiba*[3], *and G. Tsechpenakis*[1]

[1]Computer and Information Science Department, Indiana University-Purdue University Indianapolis, USA
[2]Molecular and Cellular Pharmacology Department, University of Miami Miller School of Medicine, USA
[3]Biology Department, University of Miami, USA

ABSTRACT

Type-specific dendritic arborization patterns dictate synaptic connectivity and are fundamental determinants of neuronal function. We exploit the morphological stereotypy and relative simplicity of the Drosophila nervous system to model the diverse neuronal morphologies of individual motor neurons (MNs) and understand underlying principles of synaptic connectivity in a motor circuit. In our analysis, we use images depicting single neurons labeled with green fluorescent protein (GFP) and serially imaged with laser scanning confocal microscopy. We model morphology with a novel formulation of Conditional Random Fields, a *latent state CRF*, to capture the highly varying compartment-based structure of the neurons (soma-axon-dendrites). We integrate a multi-class logistic model as the local potential function for combining compartment features. All parameters are learned in a single procedure, while L1-norm logistic model parameters are added in the maximum pseudo-likelihood model for learning with better scalability. The regularization hyper-parameters are chosen with a minimum cross-validation generalization error model.

Index Terms— neuron morphology, Drosophila, latent state Conditional Random Fields

1. INTRODUCTION

Normal locomotive behavior is fundamentally determined by precise patterns of motor neuron (MN) connectivity that are dictated by the selective connection of motor axons with muscle targets as well as specific dendritic input from presynaptic neurons. In the vertebrate spinal cord, MNs are organized into columns based on axonal projection patterns and within these columns, groups of MNs that target individual muscles are clustered into MN pools [8]. Different MN pools elaborate distinct dendritic arborization patterns and respond to sensory stimulation with different latencies, demonstrating that the selectivity of synaptic input is directly influenced by the differential patterning of MN dendrites in the spinal cord [19]. Therefore, determining how different MN subtypes pattern and organize their dendrites (Fig. 1) will be a crucial step toward understanding how motor circuits are assembled to control locomotion.

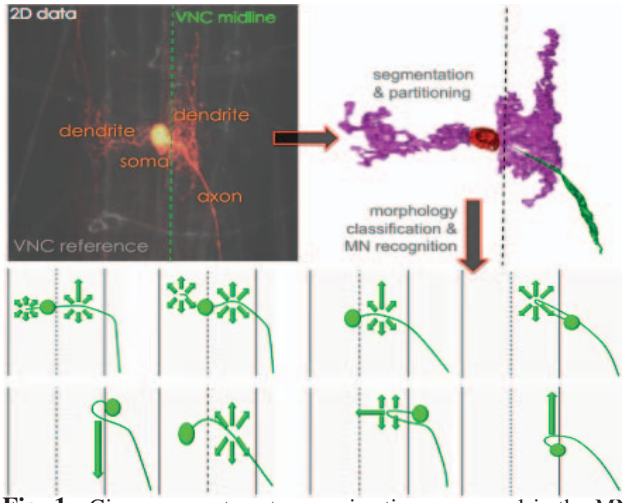


Fig. 1: Given compartment approximations, our goal is the MN subtype classification into distinct morphologies, shown with the sketches and described in [4]: the arrows show dendrites positions and extends, the disks illustrate somata, the curved lines show axons, while the dotted lines represent the VNC midline. The sketches correspond to different motor neurons from four main nerve branches (ISN, ISNb, ISNd, SNa): **top row, left to right:** [ISN: MN1-lb], [ISN: MN9-lb, MN10-lb], [ISN: MN2-lb], [ISN: MN3-lb, MN4-lb]; **bottom row, left to right:** [ISN: MN11-Ib, MN18-Ib, MN19-Ib, MN20-Ib], [ISNb: MN6/7-Ib, MN12-Ib, MN13-Ib,MN14-Ib, MN30-Ib, MNISNb/d-Is, MNISNb/d-II; ISNd: MN15/16/17-lb], [ISNb: MN15/16-lb], [SNa: MN21/22-Ib, MN22/23-Ib, MN23/24-Ib]. For further details we refer the reader to the study in [4].

Our goal is to understand how connectivity patterns are formed within a normal motor circuit and determine whether a 'connectivity map' can be deduced by estimations of neuronal morphology. This will greatly facilitate our investigation into how different MN subtypes pattern their dendritic arbor to establish synaptic connectivity, which remains an understudied yet important aspect of motor circuit formation. Lastly, understanding the normal process of MN development will ultimately be important in devising strategies to repair or restore MN connectivity after spinal cord injury or disease.

Drosophila larvae as a model. The Drosophila embryonic

This work is supported in part by NSF/DBI [#1062405]: 'ABI Innovation: Modeling the Drosophila Brain with Single-neuron Resolution using Computer Vision Methods', awarded to G. Tsechpenakis, NIH/NINDS [5RC2NS069488-02]: 'In Situ Protein-Protein Interaction Networks of Neurons (isPIN)', awarded to A. Chiba, and NIH/NINDS [NS072588]: 'Genetic Mechanisms of Motor Neuron Degeneration', awarded to M.D. Kim.

central nervous system (CNS) comprises the developing brain and ventral nerve cord (VNC). The VNC, which can be considered functionally orthologous to the mammalian spinal cord, is segmentally reiterated and bilaterally symmetrical with respect to the ventral midline. There are approximately 400 neurons, including an estimated 38 MNs, within each hemisegment (or half-segment) of the VNC . Abdominal hemisegments in the embryo comprise 30 stereotyped body wall muscles, each of which is innervated by one or more of the 38 MNs. The muscle innervation pattern of individual MNs is further highly stereotyped making embryonic MNs uniquely identifiable [7, 4]. During larval development, dendrites of these abdominal MNs undergo considerable growth and dendritic branching is dramatically increased, likely reflecting extensive changes in synaptic connectivity that are required for more complex larval behaviors such as the peristaltic movements required for normal locomotion. Whether larval MNs project their dendrites to stereotyped regions of the VNC to create a connectivity map is unclear.

Imaging at single-cell resolution. We use mosaic analysis with a repressible cell marker (MARCM) [4, 10], a genetic technique that allows us to label and image individual MNs. The use of fluorescent proteins allows for high-resolution in vivo imaging with minimal photobleaching, reduced phototoxicity and enhanced labeling of the neuronal membrane. Although labeled MNs generated by MARCM can be imaged live in the intact animal, muscle contraction hinders the acquisition of confocal images through consecutive z positions. The brain and VNC are therefore exposed by dissection and the tissue is fixed with formaldehyde before immunostaining with antibodies directed against mCD8 and a secondary marker, Fasciclin II (FasII). FasII labels axon fascicles that divide the VNC into distinct territories and provides a frame of reference in which to map the relative positions of the MN soma and dendrites. The entire morphology of single MNs is then imaged with laser scanning confocal microscopy [4].

Morphology classification. Conditional Random Fields (CRFs) [6] have been proven very robust in part-based object [5] and action [18] recognition, as well as image labeling [3]. In some applications, however, higher level semantics are necessary to improve classification, which is the main motivation for using the hierarchical CRFs [20, 11]. To estimate the parameters in these models, the information of the hidden layers should be available for training, thus making ground-truthing labor-intensive. To tackle this issue the hidden Conditional Random Field (HCRF) [14] was introduced. HCRFs have been successfully applied in gesture [16], voice [15], image [13], and human action [21] recognition.

In this paper we describe a latent Condition Random Field, a variation of HCRFs, for part-based neuron classification (Fig. 2). The objective is the estimation of the neuron type y, given the unknown compartments t_i (soma, axon, dendrites) and their corresponding morphology features r_i. Such model incorporates topology information, e.g.,

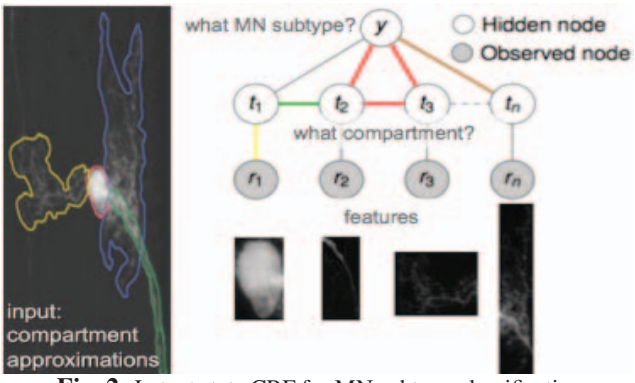


Fig. 2: Latent state CRF for MN subtype classification.

that axon is adjacent to soma. In this model we determine four types of cliques (indicated with different colors in Fig. 2) to define the model's energy potentials. Intuitively, we model the relationships between (a) the compartment label and its corresponding morphology features (discriminative four-class classifier)[yellow]; (b) the labels of two compartments [green]; (c) the labels of two compartments and the neuron type [red]; (d) a compartment label and the neuron type [brown]. During the optimization, we inherently impose relative topology constraints for the recognition of the unknown y. As observations we calculate the compartment shape histograms and distances from the VNC midline (Fig. 1). The difference of our approach with models that incorporate local classifiers, such as [5, 17], is that we do not use the compartment labels during training. Also, there are two key differences between our model and hidden CRFs: (a) we formulate features with continuous functions (as opposed to discrete variables); (b) global and local classification is achieved with a single training procedure.

2. LATENT STATE CONDITIONAL RANDOM FIELD

We build a model that integrates local compartment features and relative topology in the latent layer, as shown in Fig. 2, where the joint prediction of MN subtype y and compartment labels $\mathbf{t} = (t_1, t_2, \ldots, t_n)$ is formulated as,

$$p(y, \mathbf{t}|\mathbf{R}) = \frac{p(y, \mathbf{t}, \mathbf{R})}{\sum_{y,\mathbf{t}} p(y, \mathbf{t}, \mathbf{R})}, \quad (1)$$

where $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$, $\mathbf{r}_i$ is the feature vector of region i, and n is the total number of regions in the image. The joint distribution of all variables in the model is,

$$p(y, \mathbf{t}, \mathbf{R}) \propto \prod_c \psi_c(A_c) \quad (2)$$

where c is a clique containing a set A_c of nodes, and $\psi_c(A_c) = \exp\{-E(A_c)\}$, with $E(A_c)$ being the clique energy function. We model the distribution of eq. (2) with the clique energy function,

$$\Psi(y, \mathbf{t}, \mathbf{R}, \boldsymbol{\theta}) = -\sum_c E(A_c) =$$
$$= \sum_{C_1} \phi(r_i, t_i) + \sum_{C_2} \varphi(t_i, t_j) + \sum_{C_3} \tau(t_i, y) +$$
$$+ \sum_{C_4} \omega(t_i, t_j, y), \quad (3)$$

where θ is the set of model parameters, and (C_1, C_2, C_3, C_4) denote the four types of cliques (colored lines in Fig. 2).

Definition of $\phi(\mathbf{r}_i, t_i)$. This function formulates the compatibility between the i-th region features $\mathbf{r}_i$ and the compartment type $t_i \in T = \{$ *soma, axon, dendrite on soma, dendrite on axon*$\}$, based on multi-class logistic regression,

$$\phi(\mathbf{r}_i, t_i) = \eta \prod_{k=1}^{|T|} \left(\frac{\exp\{w_k^T \mathbf{r}_i\}}{\sum_j \exp\{w_j^T \mathbf{r}_i\}} \right)^{c_k^i}, \quad (4)$$

where η is a normalization parameter to be estimated, k encodes the compartment type, $|T| = 4$ is the size of the set T, and w_k is the parameter vector for class k in the logistic model; c_k^i is a $|T|$-element code vector, such that, e.g.: if $c^i = (1, 0, 0, 0)$, it is $c_{k=1}^i = 1$, which denotes that the i-th region is of the first class in T, namely soma.

Definition of $\varphi(t_i, t_j)$. We model the topology compatibility between two latent states (neuron compartments) as,

$$\varphi(t_i, t_j) = \gamma(t_i, t_j) g(t_i, t_j), \quad (5)$$

where $\gamma(t_i, t_j)$ is a normalization parameter for the given pair, and $g(t_i, t_j)$ encodes relative topology,

$$g(t_i, t_j) = \begin{cases} 1, & if\ (t_i, t_j)\ are\ adjacent \\ 0, & otherwise \end{cases} \quad (6)$$

Definition of $\tau(t_i, y)$. This function models the compatibility between compartment label and the type of neuron.We build the multi-class logistic regression model,

$$\tau(t_i, y) = \sum_{k=1}^{|T|} \delta(t_i = k) \xi_k \prod_{u=1}^{|\mathcal{Y}|} \left(\frac{\exp\{\alpha_{uk}^T \mathbf{r}_i\}}{\sum_l \exp\{\alpha_{lk}^T \mathbf{r}_i\}} \right)^{d_u^i}, \quad (7)$$

where ξ_k is the k-th compartment-class parameter, $|\mathcal{Y}|$ is the number of MN subtypes considered and indexed with u, and α_{uk} encodes the regression parameters for compartment class k in subtype u; d_u^i is a $|\mathcal{Y}|$-element vector, such that, e.g.: if $d^i = (1, 0, 0, \ldots, 0)$, it is $d_{u=1}^i = 1$, which denotes that the i-th region belongs to MN of subtype code 1.

Definition of $\omega(t_i, t_j, y)$. This function models the compatibility between two compartments and the motor neuron subtype, considering relative topology with respect to the reference midline in the ventral nerve cord (Fig. 1) and the neuron topological properties,

$$\omega(t_i, t_j, y) = \sum_{q=1}^{Q} \beta_q s_q(t_i, t_j, y), \quad (8)$$

where $s_q(t_i, t_j, y)$ encodes motor neuron property q in a set Q of observed properties. For example, as shown in the sketches of Fig. 1, the rightmost sketch in the first row corresponds to the morphology of MN3,4. A property q that we use for $y = MN3, 4$ is that 'all compartments are in the same side of the reference VNC with respect to the midline' (shown with dotted line),

$$s_q(t_i, t_j, y) = \begin{cases} 1, & t_i, t_j\ at\ the\ same\ side\ of \\ & midline\ \&\ y = MN3, 4 \\ 0, & otherwise \end{cases}$$

β_q is the weight of each property q.

Parameter estimation. In our training set $S = \{(\mathbf{R}^l, y^l)\}_{l=1}^N$, only the neuron labels and their features are given, while the compartments (latent classes) are unknown. We use the pseudo-likelihood model of [16], and we add the L1-norm of logistic classification [12] for local classification with better scalability. The parameter learning model is formulated as,

$$\theta^* = \arg\min_\theta - \sum_{l=1}^{N} \log p(y^l | \mathbf{R}^l, \boldsymbol{\theta}) + \frac{\|\boldsymbol{\kappa}\|_2^2}{2\lambda_1^2} + \lambda_2 \|\mathbf{w}\|_1 + \sum_{i=3}^{6} \lambda_i \|\boldsymbol{\alpha}_{i-2}\|_1, \quad (9)$$

where $\boldsymbol{\kappa} = (\eta, \gamma, \xi, \beta)$, with $\gamma = \{\gamma(t_i, t_j)\}_{(i,j)=1, i \neq j}^n$, $\xi = \{\xi_k\}_{k=1}^{|T|}$, $\beta = \{\beta_q\}_{q=1}^Q$, is the parameter set in the potential functions, and $\mathbf{w}$ and $\boldsymbol{\alpha}_{(\cdot)}$ is the multi-class logistic regression parameters in eqs. (4),(7) respectively. Also, we introduce six regularization hyper-parameters λ in eq. (9) to avoid over-fitting. For their estimation, we follow the gradient-based optimization approach of [9, 2] that minimizes the cross-validation error. The marginal distribution of hidden variables y is,

$$p(y | \mathbf{R}, \boldsymbol{\theta}) = \sum_{\mathbf{t}} p(y, \mathbf{t} | \mathbf{R}, \boldsymbol{\theta}) \quad (10)$$

To solve this objective, we employ the BFGS Quasi-Newton optimization method, which has been proven efficient even in non-smooth optimization. In our model, the topology of the latent variables is modeled using a distance-based minimum spanning tree among image patches, and the Sum-Product algorithm [1] is used for marginal probability inference.

Prediction. Our goal is to find the most likely motor neuron class label y. In the latent Conditional Random Field, the most likely latent state combination $\mathbf{t}$ is estimated simultaneously. The state prediction model,

$$(y^*, \mathbf{t}^*) = \arg\max_{y, \mathbf{t}} p(y, \mathbf{t} | \mathbf{R}', \boldsymbol{\theta}^*) \quad (11)$$

is solved by first fixing the state of y and finding the most likely latent state combination $\mathbf{t}^*$, and then finding y^* that corresponds to $\mathbf{t}^*$.

3. EXPERIMENTAL RESULTS

Here we examine three of the most challenging motor neuron subtypes, namely MN1, 9 and 15, which appear with similar morphologies, as shown in Fig. 1 (first and second sketches in first row, second sketch in second row). We used 43 images for the three subtypes: 14, 22 and 7 images depicting MN1, MN9 and MN15 respectively. We manually assigned subtype labels and annotated the neuron compartments (Fig. 3) to obtain ground-truth. Within each compartment patch, we applied intensity segmentation to derive better shape approximations[1]. Five-fold cross validation was used for estimating the model hyper-parameters and evaluating the method performance. In Table 1 we show the confusion matrix for the three subtypes, while in Table 2 we report precision, recall

[1]We applied intensity clustering (k-means, k-harmonic means and fuzzy c-means) and in each case we used the best segmentation result

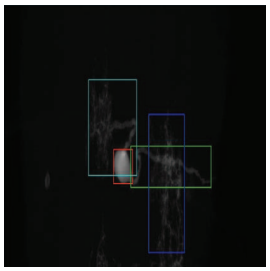


Fig. 3: Approximations for neuron compartment regions: (blue, cyan): dendrites; red: soma; green: axon.

and F1-scores. As shown in these tables, and observing the morphology illustrations in Fig. 1, the increased similarities between the three motor neuron subtypes make their classification challenging. Although we did not focus on the accurate segmentation of the compartments, our latent state CRF showed robustness with respect to shape and relative topology features. Fig. 3 shows a representative example where two latent variables (dendrite on soma: cyan box; dendrite on axon: blue box) could cause misclassification, due to appearance similarity between the axon and the dendrite branch adjacent to soma; we successfully tackle this issue considering relative topology in the model potentials.

4. CONCLUSIONS

In this paper we described a latent state Conditional Random Field, applied to the problem of motor neuron morphology classification in the Drosophila ventral nerve cord. There are two main advantages of our model: (a) it integrates local image features and spatial structure for better prediction, and (b) it does not require labels for the latent states during training. We validated our method using indicatively an image pool from three MN subtypes that are challenging to be classified. Our experimental results here show the increased robustness and scalability of our model with respect to observation approximations. This motivates the morphology classification of all ($\sim$)38 MN subtypes, among which approximately 12 distinct morphologies have been so far observed.

Actual Predicted	MN1	MN15	MN9
MN1	12	0	2
MN15	0	6	1
MN9	1	0	21

Table 1: The confusion matrix of neuron image classification.

	MN1	MN15	MN9	Macro-Average
Precision	92.31	100	87.50	93.27
Recall	85.71	85.71	95.46	88.96
F1	88.89	92.30	91.31	90.83

Table 2: Precision, Recall and F1 of three classes.

5. REFERENCES

[1] C. M. Bishop, Pattern Recognition and Machine Learning (Chapt. 8), M. Jordan, J. Kleinberg, B. Schölkopf (Eds.), Springer, 2006.

[2] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, 'Choosing multiple parameters for support vector machines,' *Machine Learning*, 46(1-3):131-159, 2002.

[3] Q. Huang, M. Han, B. Wu, and S. Ioffe, 'A hierarchical conditional random field model for labeling and segmenting images of street scenes,' *CVPR*, 2011.

[4] M. D. Kim, Y. Wen, and Y.-N. Jan, 'Patterning and organization of motor neuron dendrites in the drosophila larva,' *Developmental Biol.*, 336:213- 221, 2009.

[5] S. Kumar and M. Hebert, 'Discriminative random fields: A discriminative framework for contextual interaction in classification,' *ICCV*, 2003.

[6] J. Lafferty, A. McCallum, and F. Pereira, 'Conditional random fields: Probabilistic models for segmenting and labeling sequence data,' *ICML*, 2001.

[7] M. Landgraf, T. Bossing, G.M. Technau, and M. Bate, 'The origin, location, and projections of the embryonic abdominal motorneurons of Drosophila,' *J. Neurosci.*, 17: 9642-9655, 1997.

[8] L. Landmesser, 'The distribution of motoneurones supplying chick hind limb muscles,' *J. Physiol.*, 284: 371-389, 1978.

[9] J. Larsen, C. Svarer, L. N. Andersen, and L. K. Hansen, 'Adaptive regularization in neural network modeling,' Neural Networks: Tricks of the Trade, G. B. Orr and K.-R. Müller (Eds.), *Lecture Notes in Comp Sci*, 1524:113-132, Springer, 1996.

[10] T. Lee and L. Luo, 'Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis,' *Neuron*, 22: 451-461, 1999.

[11] L. Liao, D. Fox, and H. A. Kautz, 'Hierarchical conditional random fields for gps-based activity recognition,' Springer Tracts in Advanced Robotics, S. Thrun, R. A. Brooks, and H. F. Durrant-Whyte (Eds.), pp. 487–506. Springer, 2005.

[12] A. Y. Ng., 'Feature selection, l1 vs. l2 regularization, and rotational invariance,' *ICML*, 2004.

[13] G. Passino and E. Izquierdo, 'Patch-based image classification through conditional random field model,' *Int'l Conf. on Mobile Multimedia Communications*, 2007.

[14] A. Quattoni, M. Collins, and T. Darrell, 'Conditional random fields for object recognition,' *NIPS*, 2004.

[15] Y.-H. Sung and D. Jurafsky, 'Hidden conditional random fields for phone recognition,' *IEEE Workshop on Automatic Speech Recognition & Understanding*, 2009.

[16] B. W. Sy, A. Quattoni, L.-P. Morency, D. Demirdjian, and T. Darrell, 'Hidden conditional random fields for gesture recognition,' *CVPR*, 2006.

[17] A. Torralba, K. P. Murphy, and W. T. Freeman, 'Contextual models for object detection using boosted random fields,' *NIPS*, 2004.

[18] D. L. Vail, M. M. Veloso, and J. D. Lafferty, 'Conditional random fields for activity recognition,' *Int'l JointConf. on Autonomous Agents and Multiagent Systems*, 2007.

[19] E. Vrieseling and S. Arber, 'Target-induced transcriptional control of dendritic patterning and connectivity in motor neurons by the ETS gene Pea3,' *Cell*, 127: 1439-1452, 2006.

[20] Y. Wang and S. Gong, 'Conditional random field for natural scene categorization,' *BMVC*, 2007.

[21] Y. Wang and G. Mori, 'Learning a discriminative hidden part model for human action recognition,' *NIPS*, 2008.