Bayes Optimal Unlabeled Graph Classification: Applications in Statistical Connectomics

Joshua T. Vogelstein, Mark Dredze, R. Jacob Vogelstein, Carey E. Priebe

Abstract—Because graphs can encode more information in their structure than vectors, they are becoming increasingly popular data structures for representing information. While the last century has witnessed the development of a plethora of statistical tools for the analysis of data, the vast majority of these tools natively operate in vector spaces, not graph spaces. Thus, algorithms for even relatively simple statistical inference tasks, such as two-class classification, are essentially absent for graph data. In this work, we propose a number of complementary algorithms to classify graphs, with special attention to the possibility of unknown vertex labels. Since exactly solving the graph-matching problem is currently computational intractable, we consider several approximate approaches. We introduce a multiple-restart Frank-Wolfe approach to solving the graph matching problem by formulating it as a quadratic assignment problem. Although this approach has superior performance than previous state-of-the-art approaches to the graph matching problem, even when it "should" do well in classification problems, it is outperformed by a graph invariant strategy. This is just the beginning.

Index Terms—statistical inference, graph theory, network theory, structural pattern recognition, connectome.

1 Introduction

This work addresses the graph classification exploitation task. Consider the following idealized scenario. Let $\mathbb{G} : \Omega \mapsto \mathcal{G}_n$ be a graph-valued random variable taking values $G \in \mathcal{G}_n$. Let Y be a categorical random variable, $Y : \Omega \mapsto \mathcal{Y} \subseteq \mathbb{Z}$, such that each graph has an associated class. A graph classifier $h : \mathcal{G}_n \mapsto \mathcal{Y}$ is any function that maps from graph space to class space. The *risk* of a classifier is the expected misclassification rate, $L_h = \mathbb{E}[h[G] \neq Y]$. The optimal classifier is the classifier that minimizes risk:

$$h^* = \operatorname*{argmin}_{h\in\mathcal{H}} \mathbb{E}[h(\mathbb{G}) \neq Y]. \tag{1}$$

where $\mathcal{H}$ is the set of possible classifiers. Let $L^* = L_{h^*}$ indicate minimal (optimal) risk.

Graph classification differs from classification of vector-valued random variables in several key aspects. First, the structure of a graph may encode information. Second, the vertex labels may or may not be observed. In unobserved scenarios, NP-hard problems rear their ugly heads [?].

2 Graph Shuffling

Let G and G' be isomorphic to another if and only if $\exists\, \sigma[G] = G'$, where $\sigma : \mathcal{G}_n \mapsto \mathcal{G}_n$ is a vertex permutation function. An *unlabeled graph* is actually a set: $\widetilde{G} = \{\sigma[G] : \forall \sigma \in \Sigma_n\}$, where Σ_n is the set of permutation functions on n vertices. Let $\widetilde{\mathcal{G}}_n$ be unlabeled graph space. A *shuffle channel*, $\mathcal{C} : \mathcal{G}_n \mapsto \widetilde{\mathcal{G}}_n$ is a channel that takes as input a graph, uniformly at random selects a permutation function $\sigma \in \Sigma_n$, and outputs an unlabeled graph. Let $\widetilde{h} : \widetilde{\mathcal{G}}_n \mapsto \mathcal{Y}$ be an unlabeled graph classifier, and $\widetilde{L}^*$ be the unlabeled optimal risk for the unlabeled optimal classifier, $\widetilde{h}^*$.

3 Model Based Graph Classifiers

Let $\mathbb{P} = \mathbb{P}_{\mathbb{G},Y}$ indicate a joint distribution of graphs and classes. This joint distribution may be decomposed into the product of a likelihood and prior term: $\mathbb{P}_{\mathbb{G},Y} = \mathbb{P}_{\mathbb{G}|Y}\pi_Y$. Let π_y denote class prior probabilities, $P[Y = y] \overset{\triangle}{=} P[\{\omega : Y(\omega) = y\}]$, and let $\mathbb{P}_{\mathbb{G}|Y=y} = \mathbb{P}_y$ denote the class-conditional distribution. Assuming the data were sampled from a distribution, $\mathbb{P}_{\mathbb{G},Y}$, a Bayes classifier which chooses the maximum a posteriori class is optimal [?]:

$$h^*[G] = \operatorname*{argmax}_{y\in\mathcal{Y}} \mathbb{P}_y \pi_y. \tag{2}$$

Let $\widetilde{\mathbb{P}} = \mathbb{P}_{\widetilde{\mathbb{G}},Y}$ indicate the joint distribution of unlabeled (or shuffled) graphs and classes. In other words, $\widetilde{\mathbb{P}}$ is the same as $\mathbb{P}$ except that the graphs have been passed through a shuffle channel, so $\mathbb{P}$ is a distribution over graphs, and $\widetilde{\mathbb{P}}$ is a distribution over sets of isomorphic graphs. A Bayes optimal unlabeled graph classifier is:

$$\widetilde{h}^*[\widetilde{G}] = \operatorname*{argmax}_{y\in\mathcal{Y}} \mathbb{P}_{\widetilde{\mathbb{G}}|Y=y}\mathbb{P}_{Y=y}. \tag{3}$$

Let L^* and $\widetilde{L}^*$ be the Bayes risk and unlabeled Bayes risk, under $\mathbb{P}$ and $\widetilde{\mathbb{P}}$, respectively.

It may be illustrative to consider the class-conditional distributions of graphs as a categorical distribution: $\mathbb{P}_y[G] = \text{Cat}(G; \boldsymbol{\theta}_y)$, where $\boldsymbol{\theta}_y = (\theta_1, \ldots, \theta_d)$, and $\theta_i \geq$

- *J.T. Vogelstein and C.E. Priebe are with the Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD 21218.*
E-mail: joshuav@jhu.edu
- *R.J. Vogelstein is with the Johns Hopkins University Applied Physics Laboratory, Laurel, MD, 20723.*

$0, \sum_{i\in[d]} \theta_i = 1$, $d = |\mathcal{G}_n|$ and $[d] = \{1, 2, \ldots, d\}$. After passing $\mathbb{G}$ through a shuffle channel, all graphs that were isomorphic to one another must have the same probability (because the shuffle channel samples a permutation uniformly at random from Σ_n). The distribution resulting from passing $\mathbb{G}$ through a shuffle channel can therefore also be thought of as a categorical distribution, but over sets of isomorphic graphs, $\widetilde{\mathbb{P}}_y[\widetilde{G}] = \mathbb{P}_{\widetilde{\mathbb{G}}|Y=y}[\widetilde{G}] = \text{Cat}(\widetilde{G}; \boldsymbol{\eta}_y)$, where $\boldsymbol{\eta}_y = (\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_{\widetilde{d}})$, with similar constraints on $\boldsymbol{\eta}_y$ as $\boldsymbol{\theta}_y$. Note that in general $\widetilde{d} \ll d$.

Thus, the effect of passing $\mathbb{G}$ through a shuffle channel is to change $\boldsymbol{\theta}$ such that graphs isomorphic to one another have identical probabilities, that is $\{\theta_i = \eta_j \, \forall i : G_i \in \widetilde{G}_j\}$. Prior to shuffling, there is no such constraint.

4 Shuffling Can Degrade Optimal Performance

The result of passing the graph through a shuffle channel can only degrade, but not improve, Bayes risk, as proven below.

Theorem 1. *Assuming $(\mathbb{G}, Y) \sim \mathbb{P}_{\mathbb{G},Y}$, $L^* \leq \widetilde{L}^*$*

Proof: Let $\mathbb{P}[\mathbb{G} = G | Y = y] = \mathbb{P}_y[\mathbb{G} = G] = \mathbb{P}_y[G]$, where $\mathbb{P}_y[G] \geq 0$ and $\sum_{G\in\mathcal{G}_n} \mathbb{P}_y[G] = 1$. Therefore:

$$\begin{aligned} L^* &= \mathbb{E}[h^*(\mathbb{G}) \neq Y] = \int_{\mathcal{G}_n} \mathbb{I}\{h^*(\mathbb{G}) \neq Y\} d\mathbb{P} \\ &= \sum_{G\in\mathcal{G}_n} \min_y \mathbb{P}_y[G] \end{aligned} \tag{4}$$

where $\mathbb{I}\{\cdot\}$ is the indicator function taking value unity whenever its argument is true and zero otherwise. When the graphs are shuffled, we have:

$$\begin{aligned} \widetilde{L}^* &= \sum_{G\in\mathcal{G}_n} \min_y \mathbb{P}_y[\sigma[G] = G] = \sum_{G\in\mathcal{G}_n} \min_y \mathbb{P}_y[\sigma[G] \in \widetilde{G}] \\ &= \sum_{\widetilde{G}\in\widetilde{\mathcal{G}}_n} \min_y \mathbb{P}_y[G \in \widetilde{G}] \end{aligned} \tag{5}$$

The result follows from the fact that $\min_y \mathbb{P}_y[G \in \widetilde{G}] \geq \min_y \mathbb{P}_y[\mathbb{G} = G]$ for all $G \in \mathcal{G}_n$. □

5 When Shuffling Necessarily Degrades Bayes Optimal Performance

The above proof demonstrates that shuffling can degrade Bayes risk, but not necessarily. A natural follow-up question is: "under what circumstances does shuffling degrade Bayes risk?" Below we prove that if the labels contain any class-conditional signal, then shuffling necessarily degrades Bayes risk. We write $\mathbb{P}_{\widetilde{\mathbb{G}}|Y} = \mathbb{P}_{\mathbb{G}|Y}$ if and only if $\mathbb{P}[G' \in \widetilde{G}|Y] = \mathbb{P}[\mathbb{G} = G'|Y]$ for all $G' \in \mathcal{G}_n$. In other words, $\mathbb{P}_{\widetilde{\mathbb{G}}|Y} = \mathbb{P}_{\mathbb{G}|Y}$ if and only if $\{\theta_i = \eta_j \, \forall i : G_i \in \widetilde{G}_j\}$.

When shuffling changes the distribution of graphs, such that it was the case that $\{\theta_i \neq \eta_j \, \forall i : G_i \in \widetilde{G}_j\}$ prior to shuffling, then shuffling necessarily degrades misclassification performance.

Theorem 2. *Assume $\pi_y = 1/|\mathcal{Y}|$ for all $y \in \mathcal{Y}$ without loss of generality. If $\mathbb{P}_{\widetilde{\mathbb{G}}|Y} \neq \mathbb{P}_{\mathbb{G}|Y}$ then $L^* < \widetilde{L}^*$ (note the strictly less than).*

Proof: If $\mathbb{P}_{\widetilde{\mathbb{G}}|Y} \neq \mathbb{P}_{\mathbb{G}|Y}$, then it must be the case that for at least one $G' \in \mathcal{G}_n$ such that $\mathbb{P}_y[\mathbb{G} = G'] \neq \mathbb{P}_y[G' \in \widetilde{\mathbb{G}}]$. Thus, for all $G' \in \widetilde{G}'$, $\min_y \mathbb{P}_y[G' \in \widetilde{G}] > \min_y \mathbb{P}_y[\mathbb{G} = G']$, which implies that:

$$L^* = \sum_{G\in\mathcal{G}_n} \min_y \mathbb{P}_y[G] < \sum_{\widetilde{G}\in\widetilde{\mathcal{G}}_n} \min_y \mathbb{P}_y[G \in \widetilde{G}] = \widetilde{L}^*. \tag{6}$$

□

6 Bayes Optimal Graph Classification After Shuffling

If the graph G has been passed through a shuffle channel, and one still desires to classify it, one might consider two complementary approaches. First, one might try to "unpermute" the graph, to recover the vertex labels, and then use a Bayes optimal graph classifier. Second, one might try to use a graph-invariant based classifier. A graph invariant is any function: $\psi : \mathcal{G}_n \mapsto \mathbb{R}^d$ such that $\psi(G) = \psi(\sigma(G))$ for all $\sigma \in \Sigma_n$ and $G \in \mathcal{G}_n$. A graph invariant based classifier first projects a graph into an invariant space and then classifies. The Bayes optimal graph invariant classifier minimizes risk over all invariants:

$$h^*_\psi = \operatorname*{argmin}_{h_\psi \in \mathcal{H}_\psi} \mathbb{E}[h_\psi(\mathbb{G}) \neq Y], \tag{7}$$

and L^*_ψ is the Bayes invariant risk.

Below we show two seemingly contradictory results. First, trying to recover the vertex labels is futile: it cannot be done. Second, it is optimal.

Theorem 3. *After passing a graph through a shuffle channel, $\mathcal{C}(G) = \widetilde{G}$, all graphs in $\mathcal{G}$ are equally likely to have given rise to $\widetilde{G}$.*

Proof: The number of different permutation functions $|\Sigma_n|$ is also the number of graphs within an isomorphism class, $|\widetilde{G}|$. A shuffle channel samples a permutation function uniformly at random from Σ_n. Thus, once G is shuffled, the probability of any graph G giving rise the the unlabeled graph $\mathcal{G}$ is $1/|\widetilde{G}|$. □

Thus, there is no hope to recover the original graph. Yet, once the graph has been shuffled, a Bayes optimal unlabeled graph classifier cannot be beat.

Theorem 4. *After passing a graph through a shuffle channel, $L \geq \widetilde{L}^*$.*

Proof: After passing a graph through a shuffle channel, the effect on the graph distribution is a normalizing of likelihoods. Specifically, $\mathbb{P}_y[\sigma(G)] = \mathbb{P}_y[G]$ for all

$\sigma \in \Sigma_n$ and $G \in \mathcal{G}_n$. Thus, $\widetilde{\mathbb{P}}_y[\widetilde{G}] \propto \mathbb{P}_y[G]$ and

$$\operatorname*{argmax}_{y \in \mathcal{Y}} \mathbb{P}[\widetilde{\mathbb{G}} = \widetilde{G} | Y = y] = \operatorname*{argmax}_{y \in \mathcal{Y}} \mathbb{P}[\mathbb{G} = G | Y = y] \quad (8)$$

□

In other words, $\widetilde{L}^* = L^*_\psi$, that is, the Bayes unlabeled graph classifier is the optimal invariant-based classifier.

7 A Consistent and Efficient Unshuffling-Based Classifier

The above results show that after passing a graph through a shuffle channel, although nothing can be gained trying to unpermute the vertices, classifying the unlabeled graphs is optimal. Here, we show that we can induce a consistent and efficient classifier from training data, that is, a classifier can be estimated that is asymptotically guaranteed to be optimal.

Assume that a collection of n graph/class pairs are sampled independently and identically from some true but unknown distribution, $(\mathbb{G}_i, Y_i) \overset{iid}{\sim} \mathbb{P}_{\mathbb{G},Y} = \mathbb{P}_Y \pi_Y$, and that each graph has been passed through a shuffle channel. The *training data* is therefore $\mathcal{T}_n = \{(\widetilde{\mathbb{G}}_i, Y_i)_{i \in [n]}$. In a slight abuse of notation, an *induced unlabeled graph classifier*, $h : \widetilde{\mathcal{G}}_n \times (\widetilde{\mathcal{G}}_n \times \mathcal{Y})^n \mapsto \mathcal{Y}$ aims to estimate h^* from the training data. A Bayes plugin unlabeled graph-classifier estimates the likelihood and prior terms and plugs them in to Eq. (3)

$$\hat{h}_{BPI}[G] = \operatorname*{argmax}_{y \in \mathcal{Y}} \hat{\mathbb{P}}_{\widetilde{\mathbb{G}}|Y=y}[G] \hat{\pi}_y \quad (9)$$

Given consistent estimators for $\mathbb{P}_y[G]$ and π_Y, the Bayes plugin classifier is also consistent [**?**]. Formally, if $\hat{\mathbb{P}}_y[G] \to \mathbb{P}_y[G]$ and $\hat{\pi}_Y \to \pi_Y$ as $n \to \infty$, then $\hat{h}_{BPI} \to h^*$ as $n \to \infty$.

As described above, the class-conditional distributions of unlabeled graphs can be characterized as categorical distributions, $\mathbb{P}_{\widetilde{\mathbb{G}}|Y=y} = \text{Cat}(\boldsymbol{\eta}_y)$. Because a categorical distribution is in the exponential family, the maximum likelihood estimate for its parameters are exist, are unique, consistent, and efficient. Moreover, the class prior can also be represented as a categorical random-variable, and therefore has the same properties. Taken together, these results demonstrate that the Bayes plugin unlabeled graph classifier is consistent and efficient.

8 A Practical Approach to Unlabeled Graph Classification

While the likelihood parameters are available, estimating them requires first solving a computationally hard problem. Specifically, to estimate each η_j one must first determine whether each additional graph is isomorphic to a previously observed graph. Graph isomorphism is not known to be in P or NP, which means there is no known algorithm for solving it in polynomial time (in the worst case). Thus, the above consistent result depends on solving an infinite number of NP problems! While this sounds bad, in practice, many graph isomorphism algorithms are available, including approximate ones [**?**].

Figure **??** shows the performance of a Bayes plugin unlabeled graph classifier on simulated data....

9 Connectome Classification

A "connectome" is a graph in which vertices correspond to biological neural units, and edges correspond to connections between the units. Diffusion Magnetic Resonance (MR) Imaging and related technologies are making the acquisition of MR connectomes routine [14]. We use 49 subjects from the Baltimore Longitudinal Study on Aging, with acquisition and connectome inference details as reported in [15]. For each connectome, we obtain a 70×70 element adjacency matrix, where each element of the matrix encodes the number of streamlines between a pair of regions, ranging between 0 and about 65,000. Associated with each graph is class label based on the gender of the individual (24 males, 25 females). Because the vertices are labeled, we can compare the results of having the labels and not having the labels. As such, we implement the following classification strategies. In each case, we use a leave-one-out strategy to evaluate performance:

- `N/A-QAP` Using the vertex labels, implement a standard 1NN classifier, where distance is the norm of the difference between any pair of adjacency matrices.
- `1-QAP` Permute only the vertex labels of the test graph, and then implement `1NN`$\circ$`QAP`$_1$.
- `48-QAP` Permuting the vertex labels, then implement `1NN`$\circ$`QAP`$_1$.
- `AVG-QAP` Permuting the vertex labels, `QAP`$_1$ each of the 48 training graphs to the test graph. Then, given those permuted adjacency matrices, compute the average, and then implement a standard 1NN classifier.
- `1NN-GI` Use the graph invariant approach as described above. We provide the normalized graph invariants as inputs into a number of standard classifiers, including kNN, linear classifiers, support vector machines, random forests, and CW. On this data, the CW classifier performed best; we therefore only report its results.

Table 1 shows leave-one-out misclassification rates for the various strategies.

TABLE 1
MR Connectome Leave-One-Out Misclassification Rates

N/A-QAP	1-QAP	48-QAP	AVG-QAP	1NN-GI
20%	31%	45%	??	25%

10 Discussion

In this work, we have presented a number of approaches one could take to classifier graphs. Importantly, when the vertex labeling function is unavailable, one must deal with this uncertainty somehow. We compare a number of approaches on both simulated and connectome data. A multiple-restart Frank-Wolfe approach to approximating QAP outperforms previous state-of-the-art approaches in terms of approximating the graph matching problem. Simulations demonstrate that only the first iteration of such an iterative algorithm, starting from the identity matrix, yields classification performance better than chance. Moreover, the first iteration is identical to LAP, which is a linear problem with linear and non-negativity constraints, and therefore can be solved quite easily.

On a connectome dataset, we compare the performance of various `QAP` classification algorithms with several graph invariant (GI) based strategies. Of the algorithms that we tried, a graph invariant approach was most effective, even though, in theory, a QAP based approach could have done better (compare the first and last columns of Table 1).

These analyses leave many open questions. Perhaps most interestingly, when might one expect a QAP-based approach to outperform a GI-based approach? Resorting to a generative model, it should be clear that if the class conditional difference is independent of the vertex labels, then there is no reason to even try to implement graph matching. However, if one believes that the labeling function might convey some class-conditional signal (as in the connectome data), then QAP-based approaches could outperform any approach that ignores the labeling function. Which QAP-based approach to use in such a scenario, however, will depend on many factors, including the assumed model and computational resources.

Acknowledgments

References

[1] H. Bunke and K. Riesen, "Towards the Unification of Structural and Statistical Pattern Recognition," *Pattern Recognition Letters*, vol. in press, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167865511001309

[2] D. Conte, P. Foggia, C. Sansone, and M. Vento, "THIRTY YEARS OF GRAPH MATCHING," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 18, no. 3, pp. 265–298, 2004.

[3] M. R. Garey and D. S. Johnson, *Computer and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[4] R. E. Burkard, M. Dell'Amico, and S. Martello, *Assignment Problems*. SIAM, 2009. [Online]. Available: http://books.google.com/books?id=nHIzbApLOr0C\&pgis=1

[5] R. Sinkhorn, "A relationship between arbitrary positive matrices and doubly stochastic matrices," *The Annals of Mathematical Statistics*, vol. 35, no. 2, pp. 876–879, 1964. [Online]. Available: http://www.jstor.org/stable/2238545

[6] V. N. Vapnik, *Statistical Learning Theory*, ser. Wiley Series on Adaptive and Learning Systems for Signal Processing, Communications and Control, S. Haykin, Ed. Wiley, 1998, vol. 2, no. 4. [Online]. Available: http://www.amazon.com/Statistical-Learning-Theory-Vladimir-Vapnik/dp/0471030031

[7] L. Devroye, L. Györfi, and G. Lugosi, *A Probabilistic Theory of Pattern Recognition (Stochastic Modelling and Applied Probability)*. Springer, 1997. [Online]. Available: http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20\&amp;path=ASIN/0387946187

[8] R. P. Duin and E. Pkalskab, "The dissimilarity space: bridging structural and statistical pattern recognition," *Pattern Recognition Letters*, vol. in press, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167865511001322

[9] A. Rukhin and C. E. Priebe, "A Comparative Power Analysis of the Maximum Degree and Size Invariants for Random Graph Inference," *Journal of Statistical Planning and Inference*, vol. 141, no. 2, pp. 1041–1046, 2011.

[10] K. Crammer, M. Dredze, and F. Pereira, "Exact Convex Confidence-Weighted Learning," *Journal of Multivariate Analysis*, vol. 99, no. 5, pp. 1–8, 2008. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0047259X07000784

[11] R. E. Burkard, S. E. Karisch, and F. Rendl, "QAPLIB A Quadratic Assignment Problem Library," *Journal of Global Optimization*, vol. 10, no. 4, pp. 391–403, 1997. [Online]. Available: http://www.springerlink.com/content/n5468q3g33184443/fulltext.pdf

[12] M. Zaslavskiy, F. Bach, and J.-P. Vert, "A path following algorithm for the graph matching problem." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 12, pp. 2227–2242, 2009. [Online]. Available: http://eprints.pascal-network.org/archive/00004435/

[13] C. Schellewald, S. Roth, and C. Schnörr, "Evaluation of Convex Optimization Techniques for the Weighted Graph-Matching Problem in Computer Vision," in *Proceedings of the 23rd DAGMSymposium on Pattern Recognition*. Springer-Verlag, 2001, pp. 361–368.

[14] P. Hagmann, L. Cammoun, X. Gigandet, S. Gerhard, P. Ellen Grant, V. Wedeen, R. Meuli, J. P. Thiran, C. J. Honey, and O. Sporns, "MR connectomics: Principles and challenges," *J Neurosci Methods*, vol. 194, no. 1, pp. 34–45, 2010. [Online]. Available: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve\&db=PubMed\&dopt=Citation\&list_uids=20096730

[15] J. T. Vogelstein, W. R. Gray, J. L. Prince, L. Ferrucci, S. M. Resnick, C. E. Priebe, and R. J. Vogelstein, "Graph-Theoretical Methods for Statistical Inference on MR Connectome Data," *Organization Human Brain Mapping*, 2010.

PLACE
PHOTO
HERE

Joshua T. Vogelstein Joshua T. Vogelstein is a spritely young man, engorphed in a novel post-buddhist metaphor.

R. Jacob Vogelstein R. Jacob Vogelstein received the Sc.B. degree in neuroengineering from Brown University, Providence, RI, and the Ph.D. degree in biomedical engineering from the Johns Hopkins University School of Medicine, Baltimore, MD. He currently oversees the Applied Neuroscience programs at the Johns Hopkins University (JHU) Applied Physics Laboratory as an Assistant Program Manager, and has an appointment as an Assistant Research Professor at the JHU Whiting School of Engineerings Department of Electrical and Computer Engineering. He has worked on neuroscience technology for over a decade, focusing primarily on neuromorphic systems and closed-loop brainmachine interfaces. His research has been featured in a number of prominent scientific and engineering journals including the IEEE Transactions on Neural Systems and Rehabilitation Engineering, the IEEE Transactions on Biomedical Circuits and Systems, and the IEEE Transactions on Neural Networks.

PLACE
PHOTO
HERE

Carey E. Priebe Buddha in training.