

Classification of Neural Interferometric Imaging: Machine Learning for Non-invasive Neurosensing

Project Category: *Applied (Machine Learning for Experimental Neuroscience)*

George Sivulka (gsivulka), Maria Shcherbakova (mariash), Renee Reynolds (rreynol2)

June 9th, 2019

Submitted as coursework for CS 229A

1. Introduction

The frontiers of neuroscience, especially with recent advances in Brain Machine Interface (BMI) technology, have become increasingly focused on neurosensing—or technologies that deduce the presence and locations of firing neurons. Traditionally, methods for this electrophysiology require very invasive optical techniques—like electrode arrays or patch-clamp recordings—to return quantitative data on the firing of action potentials. However, recently, the Palanker Lab in the Hansen Experimental Physics Laboratory at Stanford has successfully proven the applicability of interferometric imaging of neurons as a noninvasive optical alternative to traditional electrophysiology [1]. These experiments can be considered the advent of optophysiology—allowing greater spatial and temporal resolution and lower barriers to implementation—which has the potential to revolutionize neurosensing in the future [2].

Currently humans fail to easily detect the changes in the interferometric images corresponding to action potential events in neurons—creating an excellent opportunity for the application of automated detection algorithms. This research proposes and compares a machine learning based solutions for the classification of this new type of neuron imaging data. The input to our algorithm is a set of 40 by 40 interferometric frames with values of the optical phase shift of one cell. We then compare the relative accuracies of different types of neural networks to output a prediction for whether or not an action potential is firing in any given frame—testing the viability of this new type of data classification.

2. Related Work

As this research focuses on neurosensing methodologies that are breaking new ground, to date there is not yet literature on this classification task with this interferogram data type. However, there does exist data processing research as well as a few machine learning papers especially in the field of interferograms tangentially related to this topic. Indeed, studies on optophysiology for the use of cell classification, especially Claire Chen’s 2016 paper in *Nature* on “Deep Learning in Label-free Cell Classification” show the applicability of machine learning for the classification of interferograms—achieving greatest classification success with Deep Neural Networks (DNNs) [3]. However, this existing optophysiology classification research serves to only classify different types of cells from each other across macro level cellular properties, whereas neurosensing optophysiology distinguishes between significantly smaller differences amongst cells themselves. Additionally, other classification tasks using one dimensional interferometric frames with values of the optical phase

shift--from lethal bioagent detection to aberrations and astigmatism in human eye tissue--indicate success with DNNs as well as Convolutional Neural Networks (CNNs) [4,5]. Still, no papers detect interferometric frames in time or classify data with a comparable signal-to-noise ratio.

The first techniques coming out of The Palanker Lab employ the binning of many frames over long spans of time (and cellular activity) to increase the legibility of action potential interferograms. Then, manually predetermined action potential templates are compared with the binned optical recording for manual detection. This sacrifices both temporal resolution and algorithmic generalizability.

3. Dataset

We used a dataset of 100,000 frames of interferograms containing over 3,000 recorded neural spiking events--with each action potential corresponding to approximately 15 subsequent frames. This data was kindly provided by Professor Daniel Palanker's lab with the help of P.h.D. candidate Kevin Boyle [6]. Each 'pixel' value of the interferogram contained a single angular displacement value measured in milliradians that measured the optical phase shift values of the cell under analysis.

Data processing involved first splitting videos of interferograms into individual frames, and reducing their resolution to a 40 by 40 matrix by bin averaging--allowing quicker training and iteration cycles to optimize models more efficiently. (Figure 1) These matrices were then unrolled into 1600 feature long vectors corresponding to each frame. Feature normalization was then conducted on each pixel by subtracting the mean from each value and then dividing by the standard deviation--allowing for faster convergence while training.

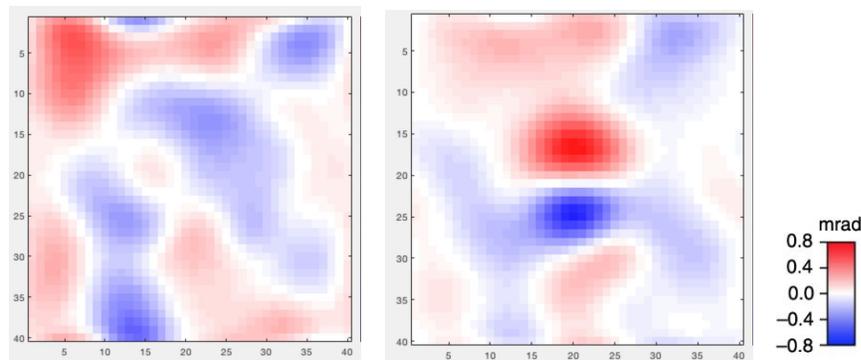


Figure 1. Two interferogram frames from our dataset visualizing the readout over the single cell. *Left:* No action potential event. *Right:* Action potential event--emphasized to an almost ~600 nm deformation for human legibility

All frames in the original dataset were labeled with electrical data indicative of the depolarisation characteristic of the electrical ion influx during an action potential. Thus, in order to label each feature vector with a ground truth for our classification task, electrical activity exceeding the action potential threshold was mapped to binary values classifying the presence of spiking event. Data was verified to be well-balanced, with exactly 47.98% positively labeled frames--allowing accuracy on the test set to be a meaningful evaluation metric.

After equally processing all data, resulting in a 100,000 by 1,601 matrix, individual data points were randomly assigned using a predetermined random seed into train, validation, and test sets with an 80-10-10 split.

4. Methods

4.1. Baseline (Logistic Regression)

To first establish our baseline we started by implementing an unregularized Logistic Regression model. Logistic Regression is a classification algorithm that works by learning a function that approximates $P(Y = 1 | X = x)$. It makes the central assumption that $P(Y = 1 | X = x)$ can be approximated as a sigmoid function applied to a linear combination of input features:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad \text{where } z = \theta_0 + \sum_{i=1}^m \theta_i x_i$$

Because the range of the sigmoid function is 0 to 1, it is especially useful for models that predict probability as an output.

In order to improve the algorithm's generalizability, we added regularization to the Logistic Regression model to compare performance.

4.2. Traditional Neural Networks

Neural Networks compute systems with multiple layers similar to logistic regression that learn features through model optimization to more effectively perform classification. In order to ultimately select the optimal architecture, we implemented three neural networks with different numbers of hidden layers. The impact of varying the number of hidden layers (two, three, and nine) on an NN's performance is discussed in detail in Section 5.

Throughout the process of training our models, we consistently used ReLU activation function in hidden layers (100 neurons each), and sigmoid function in the output layer (binary). ReLU is linear for all positive values, and zero for all negative values, which both induces sparsity and prevents the vanishing gradient issue, allowing the model to converge faster.

4.3. CNN

Convolutional Neural Networks are deep neural networks with the same methodology as above that include intermittent "convolutional" and "pooling layers." They have proven to perform consistently better on image classification tasks as compared to traditional neural networks. The practical benefit is that a CNN looks only at a small patch of the image, and thus has to learn fewer parameters which in turn significantly reduces the learning and training time required by a model--providing a deep learning comparison point.

5. Experiments/Results/Discussion

Our primary metrics included accuracy, precision, recall, F1-score, as well as confusion matrix - a categorical representation of the accuracy of a model with two or more classes.

Classification accuracy is the number of correct predictions out of all predictions made. Precision is the number of correctly predicted positive values out of all values predicted to be positive. Recall refers to the percentage of total relevant results correctly classified by the algorithm. F1 score is the weighted average of precision and recall, which thus takes both false positives and false negatives into account.

5.1. Baseline (Logistic Regression)

The baseline unregularized logistic regression model achieved training set accuracy of 0.739, validation set accuracy of 0.731, precision of 0.748, recall of 0.645, and F-Score of 0.693.

Due to the large number of inputs to our model, we decided to regularize the logistic regression model. We chose the value of the parameter, C, from a grid of values ranging between 1.0 and 10.0 using cross-validation with 3 folds. The optimal regularization parameter (C = 2.5) was calculated using grid search (.GridSearchCV in scikit-learn library), yielding training set accuracy of 0.750, validation set accuracy of 0.744, precision of 0.761, recall of 0.663, and F-Score of 0.708. We notice that the regularized logistic regression model is a slight improvement upon its unregularized version.

5.2. Traditional Neural Networks

For the sake of consistency, we chose to hold the activation functions, number of epochs and batch size constant for all neural network models. We chose 30 epochs – rather than hundreds or thousands, which are typically recommended – due to our limited computational power available and the amount of time required to train neural networks with such large numbers of epochs. We chose a batch size of 5000 to be approximately 10% of our training data size (the optimal batch size was calculated using a grid search). As mentioned earlier, we used the ReLU activation function for input and all hidden layers, and the sigmoid activation function for the output layers.

Models trained for 25 epochs	Training Accuracy	Test Set Accuracy	Test Set Precision	Test Set Recall	Test Set F-Score
<i>Logistic Regression</i>	0.739	0.731	0.748	0.645	0.693
<i>Regularized Logistic Reg. (C = 2.5)</i>	0.750	0.744	0.761	0.663	0.708
<i>Neural Net (2 layers)</i>	0.787	0.783	0.823	0.682	0.746
<i>Neural Net (3 layers)</i>	0.788	0.783	0.801	0.715	0.756

<i>Neural Net (9 layers)</i>	0.815	0.822	0.882	0.718	0.791
<i>CNN</i>	0.777	0.777	0.859	0.661	0.726

Table 1. Results of primary metrics for all models.

As anticipated, the NN with nine hidden layers exceeded all other models tested on every evaluation metrics (see *Table 2*), while the difference in performance between the models with two and three hidden layers was negligible. This being the case, we expect deeper neural nets to yield greater success for this classification task in the future.

TN: 2427	FP: 225
FN: 663	TP: 1685

Table 2. Confusion matrix for NN with nine hidden layers.

5.3. CNN

Through the use of maximum pooling as our pooling operation, we down sampled our feature map in such a way that it highlighted only the most prominent features in each patch. After extensive experimentation with different architectures and configurations of pooling layers, we found that the use of max pooling layers consistently resulted in the CNN's better overall performance.

6. Conclusion/Future Work

All of our models produced results that outperformed our baseline unregularized logistic regression model; however, the neural network model with 9 hidden layers performed the best overall, even better than the CNN. In fact, all three of the neural network models – with 2, 3, and 9 hidden layers – performed better than the CNN. This signals that further work could be done to tune the parameters of the CNN model, given more time and resources.

Additionally, while we're currently fitting our network to a quite uniform data type – with only one firing cell present – experimenting with different configurations of cell organization, action potential events with mitigated deformation, and increased interferogram resolution are areas we're excited to explore with additional datasets from the Palanker Lab in the hopes that our research can create a generalized model for different neurosensing scenarios. Given more time and computational resources, we would be interested in exploring regularized neural networks/CNN and further fine-tuning our hyperparameters through testing different combinations of activation functions, varying the number and configuration of hidden layers, and experimenting with greater numbers of epochs (i.e. hundreds or thousands) as well as smaller batch sizes.

Acknowledgements and Contributions:

All three project members contributed equally to the project across data collection and formatting, machine learning experiments, and writing of this report.

This research would not be possible without the assistance of Kevin Boyle and Prof. Daniel Palanker of The Palanker Lab in the Hansen Experimental Physics Laboratory at Stanford who were invaluable in providing the data.

Additionally, this research is indebted to the teaching staff of CS 299A: “Applied Machine Learning” for which this project was submitted as coursework.

References:

- [1] Ling, T., Boyle, K. C., Goetz, G., Zhou, P., Quan, Y., Alfonso, F. S., ... & Palanker, D. (2018). Full-field interferometric imaging of propagating action potentials. *Light: Science & Applications*, 7(1), 107.
- [2] Zhang, P. C., Keleshian, A. M., & Sachs, F. (2001). Voltage-induced membrane movement. *Nature*, 413(6854), 428.
- [3] Chen, C. L., Mahjoubfar, A., Tai, L. C., Blaby, I. K., Huang, A., Niazi, K. R., & Jalali, B. (2016). Deep learning in label-free cell classification. *Scientific reports*, 6, 21471.
- [4] Bandyopadhyay, A., Sengupta, A., Sinyukov, A. M., Barat, R. B., Gary, D. E., Michalopoulou, Z. H., & Federici, J. F. (2006). Artificial neural network analysis in interferometric THz imaging for detection of lethal agents. *International journal of infrared and millimeter waves*, 27(8), 1145-1158.
- [5] Yan, K., Yu, Y., & Jiaying, L. I. (2018, July). Neural networks for interferograms recognition. In *Sixth International Conference on Optical and Photonic Engineering (icOPEN 2018)* (Vol. 10827, p. 108273Q). International Society for Optics and Photonics.

