Computational discovery of physiomes in critically ill children using deep learning

David C. Kale, MS[1,2], Zhengping Che[1], Yan Liu, PhD[1]
[1] University of Southern California, Los Angeles, CA; [2] Children's Hospital, Los Angeles, CA

Introduction

Critical illness is dynamic and complex, but physicians often diagnose and treat based on parsimonious, static sets of symptoms and signs. For example, the Berlin definition of acute respiratory distress syndrome uses static criteria, including a threshold on the partial pressure of oxygen.[1] The Pediatric Risk of Mortality (PRISM III) score considers only the extreme values of a handful of variables during a 12-24 hour period.[2] However, the increasing volume of digital health data offers an opportunity to use computational methods to learn richer descriptors of illness (*physiomes*[3] or *phenomes*[4]) that incorporate temporal dynamics and more variables. In this work, we use deep neural networks to mine patterns from multivariate clinical time series. We apply this to a large pediatric intensive care unit (PICU) database from Children's Hospital Los Angeles (CHLA)[3] to learn patterns that are associated with known critical illnesses.

Methods

We want to train a deep neural network to recognize temporal patterns of length S in time series of P variables. This neural network has $D = PS$ input units and hidden layers with sizes of our choosing. We use fully connected layers, linear activations, and sigmoid nonlinearities. We pretrain each layer as a denoising autoencoder using stochastic gradient descent to minimize cross-entropy and finetune the weights with labeled data by adding an output layer and performing backpropagation.[5] From a data set of M time series with lengths $T \geq S$, we extract training set of at least $M(T - S + 1)$ overlapping patterns. Finally, we efficiently train a collection of neural networks that detect patterns of increasing lengths by using the weights of smaller neural networks to initialize the training of larger networks.[6]

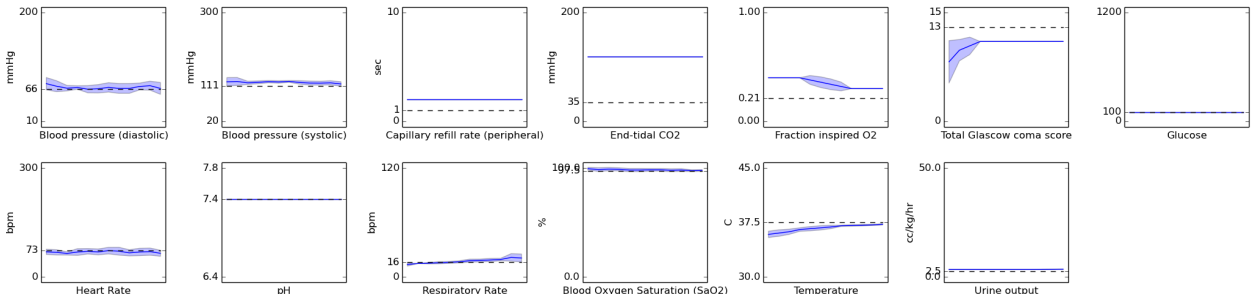
Experiments

We evaluated our framework on a preprocessed subset of the CHLA PICU database, which includes over 8,478 time series of 13 variables sampled at an hourly rate.[3] We trained a series of three-layer neural networks on all overlapping $4, 8, 12, 16, 20$, and 24 hour subsequences. We used a respiratory diagnosis label for supervised finetuning. **Figure 1** shows physiomes visualizations for two third-layer hidden units from a 12-hour neural network. For each, we found training examples with the highest activations and plotted the mean and standard deviation trajectories for each variable. In **Figure 1(a)**, we see elevated capillary refill rate (CRR), end-tidal CO_2 ($ETCO_2$), and fraction inspired oxygen (FIO_2). These indicate dehydration and breathing problems, consistent with complications due to chronic bronchopulmonary dysplasia (BPD). In **Figure 1(b)**, we see elevated CRR, $ETCO_2$, and FIO_2, plus elevated heart rate and respiratory rate (indicating respiratory distress) and declining cognitive function (GCS). This physiome is associated with pneumonia.

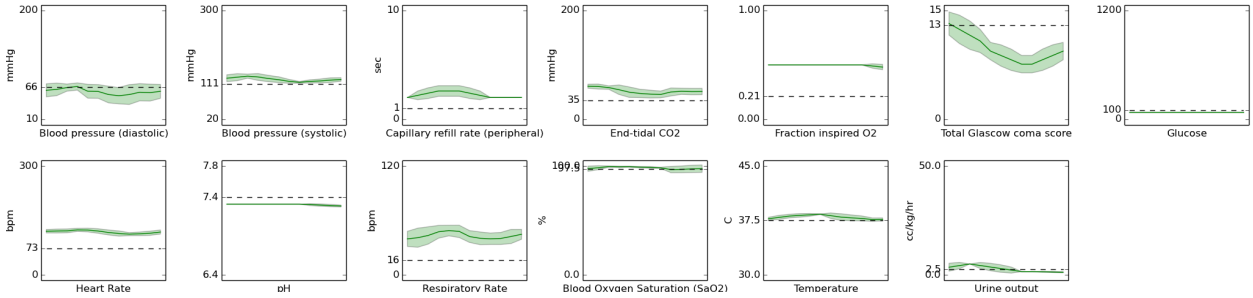
We also performed a pair of classification experiments in which the goal was to detect patients with respiratory conditions based on only their first 12 or 24 hours post-admission. We used a linear support vector machine with squared hinge loss and l_2 penalty as our classifier and five-fold cross validation to estimate the mean and standard deviation of the area under the receiver operator characteristic curve (AUC). We use the physiomic representation of each patient's first 12 or 24 hours of data (i.e., the third-layer activations in the corresponding neural network) as features. We compare the performance of these features with several baselines, including the raw time series and PRISM III-style extreme value features. **Figure 1(c)** shows the AUC for the baselines, as well as several neural network variants, in both tasks. Most of the neural network-based features are superior or equal to the PRISM III-style features.

Discussion

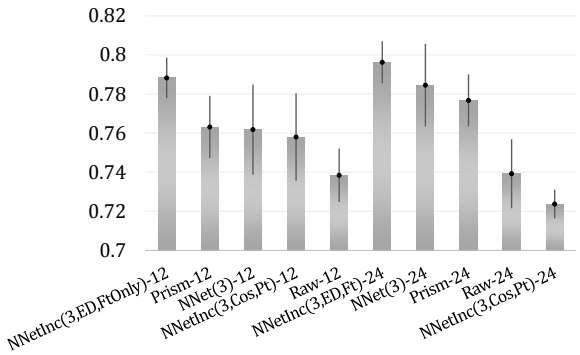
Our results are preliminary but very promising. The visualized physiomes, along with our classification results, demonstrate that we can learn trajectory-like physiomes that are physiologically plausible, associated with critical illness, and useful for classification. Consistent with the findings of Lasko, et al., our learned features did not dramatically beat the hand-engineered features (e.g., PRISM III-style) in classification.[4] This, however, does not diminish their utility or interest. The ability to learn useful features from data via automated means, rather than a painstaking manual process, could be a boon for future clinical research. Also, we believe that such physiomes may prove an intuitive supplement to traditional signs and symptoms. One question that arises is when patients were diagnosed with respiratory conditions; unfortunately, this information is not available in our data set. Also, in some cases we may be detecting the results of treatment, but we cannot verify this as treatments are not available in the current version of the CHLA PICU data set. Both of these questions point to the broader challenges of working with observational medical data.



(a) Physiome associated with bronchopulmonary dysplasia.



(b) Physiome associated with pneumonia and respiratory distress.



(c) Classification AUC for different features.

Figure 1: Experimental results on the CHLA PICU data subset.

References

[1] The ARDS Definition Task Force. *Acute respiratory distress syndrome: The Berlin definition*. JAMA. 2012;307(23):2526--2533.

[2] Pollack MM, Patel KM, Ruttimann UE. *PRISM III: an updated Pediatric Risk of Mortality score*. Critical Care Medicine. 1996;24(5):743--752.

[3] Marlin BM, Kale DC, Khemani RG, Wetzel RC. *Unsupervised Pattern Discovery in Electronic Health Care Data Using Probabilistic Clustering Models*. In: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium; 2012. p. 389--398.

[4] Lasko TA, Denny JC, Levy MA. *Computational Phenotype Discovery Using Unsupervised Feature Learning over Noisy, Sparse, and Irregular Clinical Data*. PLoS ONE. 2013 Jun;8(6):e66341.

[5] Bengio Y, Courville A, Vincent P. *Representation Learning: A Review and New Perspectives*. IEEE Trans Pattern Anal Mach Intell. 2013;35(8):1798--1828.

[6] Kale D, Che Z, Liu Y. *Fast Incremental Deep Network Training for Time Series Feature Learning*.; Under review.