

Prostate cancer detection using photoacoustic imaging and deep learning

Arjun Raj Rajanna¹, Raymond Ptucha², Saugata Sinha³, Bhargava Chinni⁴, Vikram Dogra⁴, Navalgund A. Rao³;

¹Electrical Engineering, ²Computer Engineering, ³Imaging Science, Rochester Institute of Technology, Rochester, NY, USA

⁴Department of Imaging Sciences, University of Rochester, Rochester, NY, USA

Abstract

After skin cancer, prostate cancer is the most common cancer in American men. This paper introduces a new database which consists of a large sample size of patients gathered using multispectral photoacoustic imaging. As an alternate to the standard two class labeling (malignant, normal), our voxel based ground truth diagnosis consists of three classes (malignant, benign, normal). We explore deep neural nets, experiment with three popular activation functions, and perform different sub-feature group analysis. Our initial results serve as a benchmark on this database. Greedy based feature selection recognizes and eliminates noisy features. Ablation feature ranking at the feature and group level can simplify clinician effort and results are contrasted with medical literature. Our database is made freely available to the scientific community.

Keywords: heuristic, greedy, forward, backward, feature selection, deep learning, DNN, supervised, activation functions

1. Introduction

For the year 2013, the American Cancer Society estimates that about 238,590 new cases of prostate cancer will be diagnosed and about 29,720 deaths will result from prostate cancer in the United States [1]. Prostate cancer is usually detected based on an elevated prostate specific antigen test or abnormal digital rectal exam prompting transrectal ultrasound (TRUS) guided biopsy of the prostate for definitive diagnosis. The sensitivity, specificity and accuracy of TRUS in detecting prostate cancers is reported to be 41%, 81% and 67% respectively [2].

Photoacoustic Imaging (PAI) is an emerging, noninvasive, functional and molecular imaging modality that has not yet entered the clinic. It employs short laser light pulses to excite molecules in the tissue, producing localized thermal expansion to generate ultrasonic waves. A valuable feature of PAI is the ability to discriminate among tissue constituents on the basis of optical absorption properties, allowing for PAI spectroscopy, which can detect biological function [4]. PAI can map the concentration of deoxy and oxy-hemoglobin, as well as water and lipid in the tissue. Research based evidence pointing towards its usefulness in cancer diagnosis and disease management is growing day by day. Many believe that PAI is poised to become the next major clinical imaging modality after x-ray computed tomography (CT), Ultrasound (US), Magnetic resonance imaging (MRI) and Nuclear medicine. PAI is expected to significantly impact disease management for prostate, breast, thyroid and skin cancer. With a prototype fast PAI device we have collected image data on 42 human prostate tissue samples obtained after its surgical removal.

Histopathology was used as ground truth to determine which areas in the image are to be labeled as normal, cancer and benign prostate hyperplasia (BPH). From the raw image data, 29 different features were extracted on a pixel by pixel basis.

Feature ablation, is the process of choosing the best subset of features with an objective of better generalization on the validation set [7]. It is often a key preprocessing step for machine learning applications and can play an important role in medical diagnosis [9]. Smaller selection of features offer faster models [10] and lower cost clinical tests.

As the number of features increase, the computational cost required to find an optimal subset of features increases exponentially [12], often referred to as the *curse of dimensionality*. Disadvantages of the filter based, wrapper [8], or embedded models [10], [7] have either high computational costs or are model specific. To minimize compute cost, heuristic based feature selection approaches have recently become popular [11], mainly because they overcome the previous mentioned disadvantages. Greedy forward and greedy forward with backward removal [11] efficiently determine the best subset of features while minimizing classification error.

Ranking features [10] in medical diagnosis provides more information of underlying contributors of the disease within the cancer lesions themselves. These top ranked features can be considered as strongest indicators of the disease. Simultaneous feature selection & ranking, as we show in this paper, will prove to be a useful component in prostate cancer diagnosis as it helps to understand the contribution of these highly ranked features.

Loosely inspired by the human brain, deep networks are one of the hottest topics in pattern recognition and computer vision. Deep neural networks are hierarchical artificial neural networks with more than two hidden layers arranged such that each higher level is an abstract generalization of lower layers. These hierarchical architectures enable the representation of complex concepts with fewer nodes than shallow architectures [13]. With regard to object classification, these networks have recently been shown to equal the performance of neurons in the primate inferior temporal cortex, even under difficult conditions such as pose, scale, and occlusions [21].

Recent success of deep neural networks (DNNs) applied to the task of prostate cancer classification [22], has shown that with an optimal compressed feature set, it is possible to enhance the performance of the model. In this paper, we show that a combination of an adaptive greedy forward with backward removal features selector along with a DNN classification engine offers clinical benefits to the medical community. Experimental results also suggest that different activation functions are optimal for dif-

ferent feature groups.

The rest of the sections in the paper is organized as follows: Section 2 gives a brief review of literature of previous work. Sections 3 & 4 provide a brief explanation of our dataset and methodologies respectively. Sections 5 shows the results. Sections 6 & 7, draw useful discussions and accomplishments respectively.

2. Literature review

Feature selection is a hot area for research in the field of cancer classification. Cancer growth has been attributed to genetic abnormalities [24]. There have been several approaches that consider the combination of genetic features [10], [23], [25] to choose the most non-redundant, informative subset of features and then use support vector machines to induce a strong classification model.

Previous works [26], [27] have used either greedy forward selection or backward elimination variants. The forward only, i.e. adding one feature at a time, cannot rectify the mistakes it might have made in choosing the earlier features. On the contrary, backward elimination i.e. removing the one feature at a time whose contribution minimizes the decrease the loss, risks the possibility of overfitting to noisy features. Hence it is ideal to make use of the strengths of both approaches in one algorithm [11].

Simultaneous selection and ranking of features using [11] has not been applied to different feature groups. We utilize feature selection at the individual and group level, and explain from a physics & medical standpoint as to why different arrangements make sense. There have been only a handful [22] of attempts, which have used a deep learning framework for prostate cancer classification. The results in [22], motivate us to use DNNs.

Previous work [5] introduced a subset of this dataset. Specifically, a group of chromophore features were investigated to see if they were strong indicators of cancer. This previous work used two classes (malignant and non-malignant) as opposed to our much larger dataset which uses three classes (malignant, benign, and normal), where the identification of benign from normal in an *ex-vivo* setting is quite high, and helps provide clarity for final diagnosis.

3. Dataset

Approximately 40 mm by 40 mm in size and 2-5 mm thick slices of human prostate tissue specimens (see Fig.1A) were imaged with our prototype device that uses an acoustic lens to focus the photoacoustically generated ultrasound signal [5]. Pixel size was 0.7 mm by 0.1 mm. The time signal was 2-3 microseconds long, also known as A-line signal was digitized at each pixel location. Every specimen was imaged (see Figure 1C, 1D) at 5 different wavelengths of laser light (760, 800, 850, 930 and 970 nanometers). Histopathology slide was prepared from the top 4 micron cut slice from sample (see Figure.1B) and then regions corresponding to three classes were marked by a genitourinary pathologist. By registering the marked histopathology slide to the photoacoustic (PA) image (Figure.1C and 1D), pixels belonging to the three classes were identified for further analysis. Combining all 42 specimens, the dataset for DNN classification consisted of total of 807 pixels out of which 398 were malignant cancer, 276 were normal and 133 were BPH. In all 29 different features at each pixel were made available for the analysis. These fea-

tures were further organized into 4 subgroups. Group-1 consists of feature 1 to 5, representing the peak amplitude of the signal at 5 laser wavelengths. Group-2 consists of feature 6 to 9, representing 4 major chromophore concentrations (oxy-hemoglobin, deoxy-hemoglobin, water and lipid) (REF). Group-3 features were derived from the Fourier transform (FFT) of the time domain A-line signal at each pixel location (REF). For each of the 5 laser wavelengths there were three parameters, slope, mid-band fit and intercept from a straight line fit to the FFT data. Thus this group has $3 \times 5 = 15$ features labeled as number 10 to 24. Group-4 consists of 5 centroid frequency values in the FFT of the A-line signal at each of 5 wavelengths. These 29 feature values at 807 pixel locations along with their class identification constitutes the basic dataset on which the classification in this paper was performed.

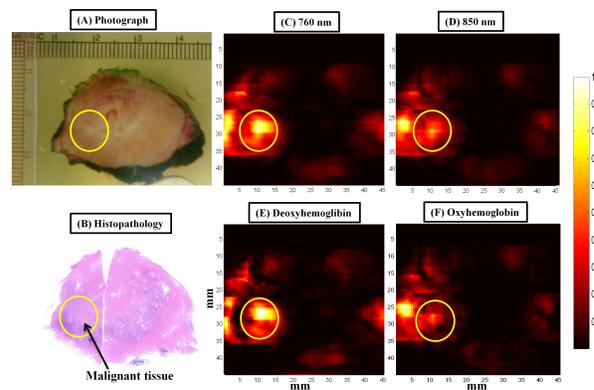


Figure 1. Multispectral PAI of prostate. PA images are acquired at multiple laser wavelengths. Each wavelength image is a composite image of all the tissue constituents such as deoxyhemoglobin (dHb), oxyhemoglobin (HbO_2), lipid and water. Chromophore analysis was performed to extract PA images showing absorption of individual constituents from the multiwavelength images. All the PA images are coregistered with histopathology and photograph of the gross specimen. (A) Photograph of gross prostate specimen; (B) Histopathology of prostate with malignant region encircled; (C) Composite PA image acquired at 760 nm wavelength; (D) Composite PA image acquired at 850 nm wavelength; (E) PA image showing absorption of dHb; and (F) PA image showing absorption of HbO_2 . Higher absorption of dHb was seen in the region of interest corresponding to malignant prostate tissue compared to HbO_2 .

4. Methodologies

4.1 Deep neural nets

Input data is normalized by mean subtraction and division by the standard deviation. Mathematically we can represent as, $I_{norm}(i) = \frac{X_i - \mu_i}{\sigma_i}$ where X_i is the i 'th input sample, $I_{norm}(i)$, is the normalized vector at a given input sample i , μ_i is the mean, σ_i is the standard deviation at sample i . We used two hidden layer feed forward networks [28]. The initial weights were randomly initialized using *Xavier initialization* [30] to overcome the *herd effect* [29]. The number of neurons per layer were chosen on experimental basis to reduce the overall cost (4). The input (I) \Rightarrow output (O) relationship was learned with a backpropagation algo-

rithm, where the weights W_{ij} were found to approximate a global minima. In addition to modifying the number of nodes per layer, we use three activation functions to non-linearly map the linear sum of inputs to each output at each node as listed below:

Sigmoid: Compresses wide range of inputs to the range of 0 to 1. Traditionally one of the most popular activation functions due to its mathematical properties. A negative of this function is that it only outputs positive values, whereby network weights are allowed to be both positive and negative. A result is slower convergence to a solution.

$$\phi(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

where $x = \sum_{i=0}^{layers} (W_i * I_{norm} + b_i)$; $W_i \in R^n$ are weights and b_i is the bias, for all future sections.

Hyperbolic tangent (Tanh): Compresses wide range of inputs to the range of -1 to +1. Like the sigmoid, the derivative is easily defined. Because weights are not constrained to be positive, it offers faster convergence.

$$\phi(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (2)$$

Rectified Linear units (ReLU): Both the sigmoid and tanh activation functions compress the tails of the distribution, making them invariant to small differences in very positive or negative values. The ReLU function clips negative values to zero, but allows the positive values to go unchanged. The resulting function is very fast to implement, and the non-saturating property on the positive side gives it good discriminating properties.

$$\phi(x) = \max(0, x) \quad (3)$$

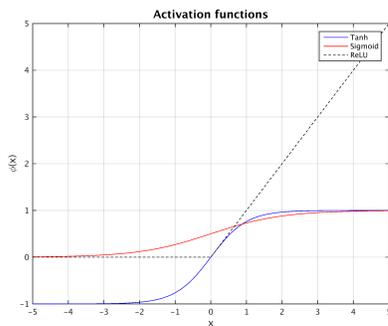


Figure 2. Activation functions: Hyperbolic tangent (Tanh), Sigmoidal, Rectifier linear unit (ReLU).

The loss function we use is as given in (4). The first term being the cross entropy and the latter term being a regularization weight penalty. During learning, wrong guesses P_i are penalized by the cross entropy, the error of which is back propagated through the network based on the partial derivatives of the weights and activation functions. The weight values are adjusted during the learning process. To avoid overfitting to the test set as

well as multiple optimal solutions, the norm-1 or norm-2 regularization, referred to as L_1 or L_2 regularization are often used. In this work, the L_2 regularization is used.

$$Cost_{net} = - \sum_i \log\left(\frac{e^x}{\sum_i e^x}\right) * t_i + C\lambda \sum |w|^2 \quad (4)$$

C is a constant, λ is weight penalty and a value equal to $1e-5$ throughout the experiments, and t_i is the target output. We back propagate the output of (4) throughout the network to fine-tune the weights.

BFGS (Broyden Fletcher Goldfarb Shanno), an unconstrained numerical optimization algorithm [16], [19] is employed in our experiments. There are two parts in an optimization algorithm: **1)** Finding the direction of greatest descent; and **2)** the step size [18] in the direction of **1)** required to reach the global minima for convex functions or local minima for non-convex functions. Our implementation consists of initial gradients g_i computed using the objective function O_f of the network, which is then fed as the steepest descent direction to the optimization algorithm. After obtaining the first search direction, we then compute the BFGS search direction, i.e. the multiplication of the inverse hessian h^{-1} multiplied by the negative of gradient ($-g_i$). We use the default setting of 100 latest search direction results [17]. The function scalar value f_i and updated gradient vector g_i are then passed through an iterative *Wolfe line search* algorithm, required to calculate the number of steps to reach the global minima.

4.2 Greedy feature selection

Greedy feature selection is a family of search based algorithms which provide an optimal feature subset. The greedy forward feature selection (GF) and greedy forward feature selection with backward removal (GFBR) are explored on the dataset. The feature list to be investigated with GF & GFBR is $F_{orig} = \{f_1, f_2, \dots, f_n\} \in R^{29}$.

4.2.1 Greedy forward feature selection (GF)

With an empty chosen feature list F_c , start by greedily adding one feature at a time, for features which induce the least error calculated using O_f on DNN/baseline classifier. The model stops growing when $F_c \in R^{n-1}$.

Consider the scenario as shown from Figure 5,

- The top 2 features selected after the GF, were $\{f_6, f_8\}$.
- The top 3 features selected after the GF, were $\{f_6, f_8, f_9\}$.
- It may also happen, that given $\{f_6, f_9\}$, $\{f_8\}$ carries little value, and the triplet $\{f_6, f_9, f_{11}\}$ may yield better performance than $\{f_6, f_8, f_9\}$.

The drawback of sub-optimal initial set of features is overcome using GFBR.

4.2.2 Greedy forward feature selection with backward removal (GFBR)

Let us consider the same list of features F_{orig} . The GFBR [11] is a combination of the forward selection and backward elimination approaches. The features are chosen by greedily adding d best features, one at a time (measured by the least error on into

F_c), then removing the r least useful features. Hence from the features chosen thus far, we greedily eliminate the least contributing feature by going through various combinations of features in F_c .

5. Results

The results are obtained by averaging the performance metrics over 10 folds of cross validation. For GFBR models, $d = 2$, and $r = 1$.

5.1 Feature group analysis

The performance of these groups on a standalone basis holds a medical importance. PA signal at each pixel contained in group-1 carries local tissue information about the chromophore concentration and their spatial microstructure in a convoluted fashion. Features of group 2, 3 and 4 attempt to separate this tissue dependent information via additional pre-processing. Group analysis provides an insight into physiological factors responsible for tissue differentiation. For example group 2 features that include oxy and deoxy hemoglobin concentration are the leading performers. This is borne out by independent evidence on prostate cancer that supports the hypothesis that malignant cancer regions are generally hypoxic (hemoglobin in blood has lower oxygen content than normal regions [6]. A grid search for finding the best architectures and activation function was performed on these feature groups. The Table 1, shows the performance of different groups on the best architecture found by the grid search.

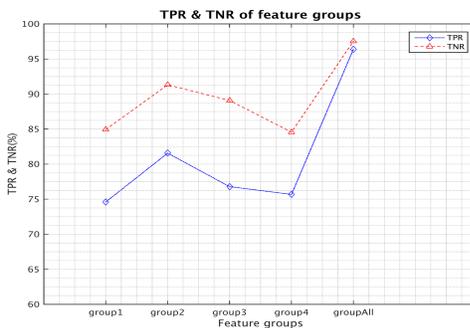


Figure 3. True positive rate, true negative rate of feature groups ; group 1= 1-5, group 2= 6-9, group 3= 10-24, group 4= 25-29, group all= 1-29.

Metrics used to gain insight into model performance are true positive rate (TPR), i.e. the number of times a sample is categorized by the model as cancer and is actually cancer, and true negative rate (TNR), i.e. the number of times a sample is categorized non-cancer and is actually non-cancer. These are also referred to as sensitivity and specificity respectively. As can be seen from Figure 3, that group 2 which consists of chromophores have the high-

Architecture	Activation function	Features	Average accuracy(%)
[50, 50]	Tanh	1-5	74.2
[50,50]	Simgoid	6-9	81.67
[70,70]	ReLU	10-24	79.55
[50,50]	Tanh	25-29	72.46
[160,160]	ReLU	All	95.04

Table 1: Performance of different feature groups with optimal activation functions.

est TPR and TNR on a standalone basis, followed by group 3 & 4. The least comes from group 1, which is a subset of group 2.

5.2 Greedy feature selection analysis

From Figure 4 we see the contribution (in terms of accuracy) of the top 11 features from the feature set F_c . The cumulative accuracy of 93.18% of top 11 features chosen initially from GFBR, in comparison to the cumulative accuracy of 85.76% of top 11 features chosen by GF, shows that the top 11 features chosen by GFBR, strongly influences the ability for the model to distinguish from one class to another. The rest of the features as shown in Figure 5, show saturation regions in both GF & GFBR after the first 11 features, clearly indicating that GFBR has the chosen non-noisy, relevant feature set F_c . Table 1 shows an accuracy of 95.04 %, the highest achieved when combining all features, meaning that the rest of the 18 features, only contribute to an approximate 1.5 to 2% difference.

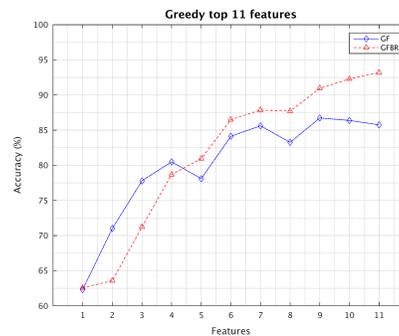


Figure 4. Greedy feature selections for first 11 features: GF Greedy forward selection, GFBR Greedy forward with backward removal. GFBR surpasses GF in terms of choosing features such that performance is enhanced within a very small subset.

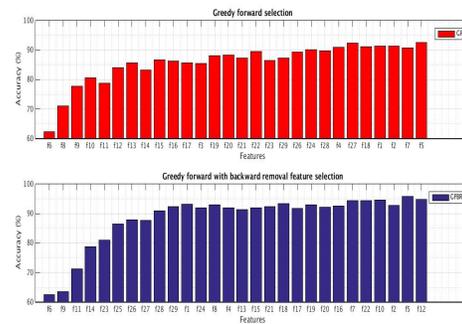


Figure 5. Greedy feature selections, for all features. It is feature ranking. Left most (ranked first), Right most (ranked last).

6. Discussions and Conclusions

Photoacoustic imaging is a potentially useful modality for cancer diagnosis and disease management. To the best of our knowledge this is the first ever multispectral PA image dataset on *ex-vivo* cancer bearing human prostate tissue samples that is large enough to be used for a DNN study described in this paper. This is also first ever 3 class classification effort on prostate samples to

differentiate between malignant cancer, normal tissue and benign prostate hyperplasia(BPH) which is non-malignant inflammation different from normal prostate tissue. Compared to TRUS, a state of the art clinical modality, improvement in accuracy from below 60% to nearly 95% is significant. We hypothesize that group 2 and group 3, contribute significantly because chromophore features inform us about the oxy de-oxy content in the blood vessels [6] and microstructural information gained by FFT play vital roles in enhancing separation between all the three classes. As evidenced from the ranked features by GFBR, we note that among the top 11 best features (Figure 5) the first five features which come from group 2 and 3 contribute to around 80% of the final accuracy. It is also worth noting that we are able to find the best features from a set of 19 features (group 2 & group 3). In future, this greedy feature analysis can be performed initially as a onetime effort on any given dataset to reduce the computational effort, cost and time during routine clinical implementation. Methodologies described in this paper can be used on any PA dataset, including *in-vivo* data if and when it becomes available. In an *in-vivo* setting it is quite likely that the accuracy will be lower due to calibration issues related to laser penetration in overlying tissue. Given that performance of TRUS as benchmark is so low, PAI along with DNNs seems to be a more promising alternative for *ex-vivo* tissue differentiation.

Acknowledgments

This work was supported by an National Institute of health (NIH) grant 1R15EB019726-01. The authors take this opportunity to thank the Radiological Society of North America (RSNA) for awarding the Research Seed Grant #RSD1118 to perform this *ex-vivo* human prostate tissue study. Authors would like to thank Research computing (RC) at RIT, for providing the computational resources required to run experiments.

References

- [1] <http://www.cancer.org/acs/groups/content/@epidemiologyandprevention/documents/document/acspc-036845.pdf>
- [2] Kuligowska E, Barish MA, Fenlon HM, Blake M. (2001) Predictors of prostate carcinoma: accuracy of gray-scale and color Doppler US and serum markers. *Radiology*; 220(3): 757-764. PMID: 11526279. <http://www.ncbi.nlm.nih.gov/pubmed/11526279>
- [3] Ehemann, Christie, et al. "Annual Report to the Nation on the status of cancer, 1975-2008, featuring cancers associated with excess weight and lack of sufficient physical activity." *Cancer* 118.9 (2012): 2338-2366.
- [4] Keerthi S. Valluru, Bhargava K Chinni, Navalgund A. Rao, Shweta Bhatt and Dogra V. (2009) "Basics and Clinical Applications of Photoacoustic Imaging". *Ultrasound Clinics*, 4(3): 403-429. Elsevier Inc. doi:10.1016/j.cult.2009.11.007. <http://www.sciencedirect.com/science/article/pii/S1556858X09000814>
- [5] Dogra, Vikram S., et al. "Multispectral photoacoustic imaging of prostate cancer: preliminary *ex-vivo* results." *Journal of clinical imaging science* 3 (2013).
- [6] Vaupel P, Kelleher DK. Blood flow and oxygenation status of prostate cancers. *Adv Exp Med Biol* 2013;765:299-305. <http://www.ncbi.nlm.nih.gov/pubmed/22879048>
- [7] Guyon, Isabelle, and Andr Elisseeff. "An introduction to variable and feature selection." *The Journal of Machine Learning Research* 3 (2003): 1157-1182.
- [8] Liu, Huan, and Hiroshi Motoda. *Feature selection for knowledge discovery and data mining*. Vol. 454. Springer Science & Business Media, 2012.
- [9] Fan, Ya-Ju, and Wanpracha Art Chaovaitwongse. "Optimizing feature selection to improve medical diagnosis." *Annals of Operations Research* 174.1 (2010): 169-183.
- [10] Saeys, Yvan, Thomas Abeel, and Yves Van de Peer. "Robust feature selection using ensemble feature selection techniques." *Machine learning and knowledge discovery in databases*. Springer Berlin Heidelberg, 2008. 313-325.
- [11] Zhang, Tong. "Adaptive forward-backward greedy algorithm for sparse learning with linear models." *Advances in Neural Information Processing Systems*. 2009.
- [12] Koller, Daphne, and Mehran Sahami. "Toward optimal feature selection." (1996).
- [13] Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh. "A fast learning algorithm for deep belief nets." *Neural computation* 18.7 (2006): 1527-1554.
- [14] Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *Science* 313.5786 (2006): 504-507.
- [15] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.
- [16] Nocedal, Jorge, and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.
- [17] Gill, Philip E., Walter Murray, and Margaret H. Wright. "Practical optimization." (1981).
- [18] Kelley, Carl T. *Iterative methods for optimization*. Vol. 18. Siam, 1999.
- [19] M. Schmidt. minFunc: unconstrained differentiable multivariate optimization in Matlab. <http://www.cs.ubc.ca/~schmidtm/Software/minFunc.html>, 2005
- [20] Ng, A.Y. Unsupervised feature learning and deep learning @ONLINE. URL <http://ufldl.stanford.edu/>
- [21] Yamins, Daniel L., et al. "Hierarchical modular optimization of convolutional networks achieves representations similar to macaque IT and human ventral stream." *Advances in Neural Information Processing Systems*. 2013.
- [22] Fakoor, Rasool, et al. "Using deep learning to enhance cancer diagnosis and classification." *Proceedings of the International Conference on Machine Learning*. 2013.
- [23] Guyon, Isabelle, et al. "Gene selection for cancer classification using support vector machines." *Machine learning* 46.1-3 (2002): 389-422.
- [24] Stratton, Michael R., Peter J. Campbell, and P. Andrew Futreal. "The cancer genome." *Nature* 458.7239 (2009): 719-724.
- [25] Mishra, Debahuti, and Barnali Sahu. "Feature selection for cancer classification: a signal-to-noise ratio approach." *International Journal of Scientific & Engineering Research* 2.4 (2011): 1-7.
- [26] Merging microarray data, robust feature selection, and predicting prognosis in prostate cancer
- [27] Song, Le, et al. "Feature selection via dependence maximization." *The Journal of Machine Learning Research* 13.1 (2012): 1393-1434.
- [28] Lapedes, Alan, and Robert Farber. *How neural nets work*. Los Alamos National Laboratory, 1988.
- [29] Van Otterlo, Martijn. *The logic of adaptive behavior: knowledge representation and algorithms for adaptive sequential decision making under uncertainty in first-order and relational domains*. Vol. 192. Ios Press, 2009.

- [30] He, Kaiming, et al. "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification." arXiv preprint arXiv:1502.01852 (2015).

Author Biography

Arjun Raj Rajanna obtained his undergraduate degree in Electronics & communication from Dr. MVSIT, VTU, and is currently pursuing his Master's in Electrical Engineering in Rochester Institute of Technology. His research focus is in machine learning, with an emphasis on deep learning with audio and visual signals.

Raymond Ptucha is an Assistant Professor in Computer Engineering at Rochester Institute of Technology specializing in machine learning, computer vision, robotics, and embedded control. Ray was a research scientist with Eastman Kodak Company for 20 years where he worked on computational imaging algorithms and was awarded 27 U.S. patents with another 25 applications on file. He graduated from SUNY/Buffalo with a B.S. in Computer Science (1988) and a B.S. in Electrical Engineering (1989). He earned a M.S. in Image Science from RIT in 2002. He earned a Ph.D. in Computer Science from RIT in 2013. Ray was awarded an NSF Graduate Research Fellowship in 2010 and his Ph.D. research earned the 2014 Best RIT Doctoral Dissertation Award.

Saugata Singha received his PhD from Rochester Institute of Technology in 2014. His research involves photoacoustic imaging and image data analysis. Saugata won the Best Student Paper award at the 2013 IEEE Western New York Image Processing Workshop for his work on detecting thyroid cancer. Presently he is an assistant professor at NIT Nagpur, India.

Dr. Navalgund Rao is a professor at Rochester Institute of Technology, Rochester, NY with an adjunct appointment at the University of Rochester medical center. For the past 30 years he has been involved in teaching and research in medical imaging, digital image and signal processing with emphasis on ultrasound and photoacoustic imaging.