

# Classification of Primary Cilia in Microscopy Images Using Convolutional Neural Random Forests

Sundaresh Ram\*, Mohammed S. Majdi<sup>†</sup>, Jeffrey J. Rodríguez<sup>†</sup>, Yang Gao<sup>§</sup>, and Heddwen L. Brooks<sup>‡</sup>

\*School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA

<sup>†</sup>Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, USA

<sup>§</sup>Department of Medicine, University of Utah, Salt Lake City, UT, USA

<sup>‡</sup>Department of Physiology, University of Arizona, Tucson, AZ, USA

Email: sr2255@cornell.edu, {mohammadsmajdi, jjrodrig}@email.arizona.edu

**Abstract**—Accurate detection and classification of primary cilia in microscopy images is an essential and fundamental task for many biological studies including diagnosis of primary ciliary dyskinesia. Manual detection and classification of individual primary cilia by visual inspection is time consuming, and prone to induce subjective bias. However, automation of this process is challenging as well, due to clutter, bleed-through, imaging noise, and the similar characteristics of the non-cilia candidates present within the image. We propose a convolutional neural random forest classifier that combines a convolutional neural network with random decision forests to classify the primary cilia in fluorescence microscopy images. We compare the performance of the proposed classifier with that of an unsupervised  $k$ -means classifier and a supervised multi-layer perceptron classifier on real data consisting of 8 representative cilia images, containing more than 2300 primary cilia using precision/recall rates, ROC curves, AUC, and  $F_\beta$ -score for classification accuracy. Results show that our proposed classifier achieves better classification accuracy.

**Index Terms**—Image classification, convolutional neural network, random forests, primary cilia, confocal microscopy.

## I. INTRODUCTION

Primary cilia are curvilinear non-motile sensory organelles protruding from the surface of many eukaryotic cells that are involved in many cell development and physiological processes. Recent research [1] has shown that primary cilia length in mammalian cells may change due to extracellular environment stimuli, such as renal injury [2] or external signaling modules. It has also been reported that renal primary cilia are involved in modulation of the mechanistic target of rapamycin (mTOR) pathway. Furthermore, it has been demonstrated that lithium treatment activates the mTOR pathway in renal collecting duct cells expressing aquaporin 2 (AQP2) [3]. The collecting duct cells in the kidney express primary cilia, as shown in Fig. 1 *a*. There is a significant interest developing an automated classifier to accurately and rapidly distinguish between primary cilia located in the AQP2 expressed region and primary cilia located elsewhere.

Manual detection and classification of primary cilia within tissues usually involves a large amount of time, especially for large-scale data processing, and is prone to multiple errors caused by background clutter, non-uniform illumination, imaging noise, and subjective bias. This serves as a motivation to develop an automatic algorithm capable of classifying

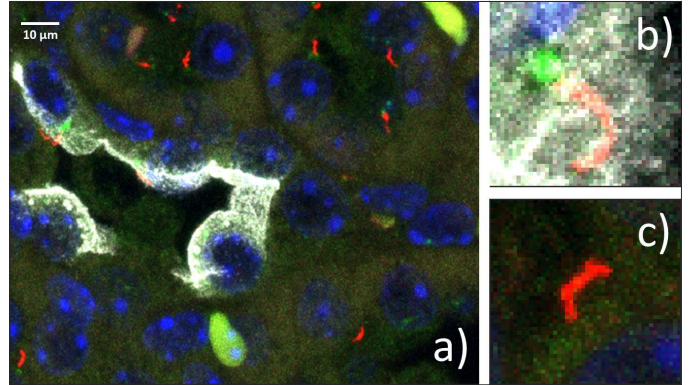


Fig. 1. Microscopy image showing primary cilia which are in red color. (a) Cilia (red color) near AQP2-expressing renal collecting duct (white color). (b) Primary cilium within an AQP2 expressed region that we are interested in. (c) Primary cilium elsewhere in the image.

the primary cilia of interest within the fluorescently labeled microscopy images. In order to classify the primary cilia, we need to first find all the cilia locations within the microscopy images. There are a number of curvilinear-structure detection methods such as the top-hat transform-based detector, ridge detector, steerable detector and the multi-scale variance stabilizing transform (MS-VST) detector that are capable of detecting all the cilia locations within the microscopy image [4]–[6]. Nevertheless, in this paper we analyze the already detected primary cilia, and we focus on classifying them as belonging to an AQP2 expressed region or elsewhere in the microscopy images.

Recently, the availability of large amounts of data and significant computational power have rapidly increased the popularity of machine learning (esp. deep learning) approaches. Convolutional neural networks (CNNs) [7] have outperformed the state-of-the-art in many computer vision applications [8]. Similarly, the applicability of CNNs has also been investigated for medical image analysis [9]. In particular, their capability to learn discriminative features when trained in a supervised fashion makes them useful for automated detection of structures in, e.g., microscopy images [9].

We propose a convolutional neural random forest classifier – a novel approach that unifies the appealing representation-learning properties of CNNs with the divide-and-conquer

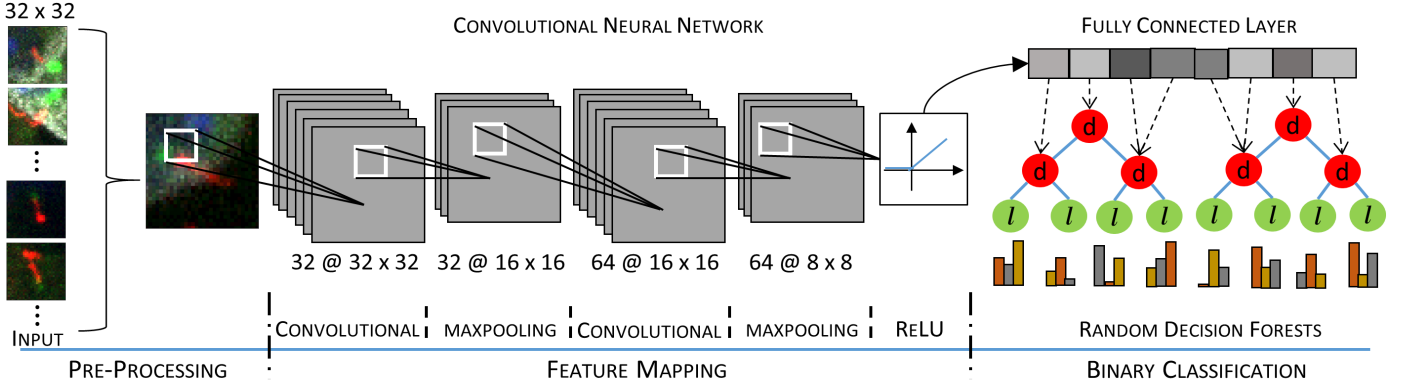


Fig. 2. An overview of the proposed convolutional neural random forest classifier. Our proposed CNN model where the feature mapping happens (shown on the left side) consists of two convolutional layers each followed by a maxpooling layer and a final ReLU activation layer, following which a dropout regularization is used to obtain the fully connected layer. The learned features are fed to our random forests classifier (shown on the right side), which have trees with decision nodes (d) (in red color) and leaf nodes (l) (in green color). At each leaf node we compute posterior probabilities belonging to each class.

principle of decision trees to classify the primary cilia of interest in fluorescence microscopy images of mouse kidney tissues labeled with AQP2 fluorescent antibodies. Our method differs from the conventional CNNs because we use a random decision forest to provide the final predictions. Our method also differs from traditional random forests as the inputs to these are the features from the CNN, which helps in reducing the uncertainty of the routing decisions of a sample taken at the split nodes of the decision trees. We describe our method in detail and present quantitative results comparing it to an unsupervised  $k$ -means classifier and a supervised multi-layer perceptron (MLP) classifier for cilia classification.

## II. METHODS

### A. Preprocessing

For classification of primary cilia of interest within our images, we first need to identify all potential primary cilia. In this work, we use the MS-VST algorithm [10] to extract all potential locations of the primary cilia in the microscopy images. Once we obtain all the candidate primary cilia, for each primary cilium, we extract gray scale image patches of  $32 \times 32$  pixels centered at the candidate primary cilium centroid. Such a patch size is chosen to contain the primary cilium ( $\sim 20 - 25$  pixels in length), and some background around the cilium ( $\sim 7$  pixels) to include the context information. These image patches are then fed to a convolutional neural random forest classifier in order to classify whether the primary cilia are located within the AQP2 expressed region or elsewhere within the microscopy images.

### B. Convolutional Neural Random Forest

The convolutional neural random forest classifier consists of two stages: a CNN in the first stage whose output is cascaded and fed in as the input to a random decision forest to predict the final class label. We define in detail each stage below.

1) **CNN Configuration:** We adopt the U-Net architecture [11] as the basis of our CNN. The motivation behind this architecture is that the contracting path captures the context around the objects in order to provide a better representation of

the object as compared to architectures such as VGGnet [12]. Networks like VGGnet are very large networks that require learning a massive number of parameters and are very hard to train in general, needing significant computational time. Thus, we empirically modify the U-Net architecture to suit our application.

To construct our CNN, we discard U-Net's layers of up-convolution and the last two layers of down-sampling and replace them with a  $1 \times 1$  convolution instead to obtain a fully connected layer. We use the rectified linear units (ReLU) [11] as the activation function for our CNN as the constant gradient of ReLUs results in faster learning and reduces the problem of vanishing gradient compared to hyperbolic tangent (tanh). We implement the maxpooling layer instead of the average pooling as sub-sampling layer [8]. We observed that the performance is better when ReLU was configured with the maxpooling layer, resulting in higher classification accuracy after 50 epochs. We used the  $1 \times 1$  convolutional filters (as suggested in [9]) for the Adam [13] optimizer. All the other parameters such as number of layers, convolutional kernel size, training algorithm, and the number of neurons in the final dense layer were all experimentally determined for our application.

The inputs to our CNN are the  $32 \times 32$  image patches of primary cilia, extracted from both within AQP2 expressed regions and elsewhere within the microscopy images. We do not zero-pad the image patches, as we already select patch sizes that are much larger than a typical cilia length and thereby avoiding the additional computational cost. Next, two consecutive convolutional layers are used in the network. The first convolutional layer consists of 32 kernels of size  $5 \times 5 \times 1$ . The second convolutional layer consists of 64 kernels of size  $5 \times 5 \times 32$ . The sub-sampling layer is set as the maximum values in non-overlapping windows of size  $2 \times 2$  (stride of 2). This reduces the size of the output of each convolutional layer by half. After the two convolutional and sub-sampling layers, we use a ReLU, where the activation  $y$  for a given input  $x$  is obtained as

$$y = f(x) = \max(0, x) \quad (1)$$

A graphical representation of the architecture of the proposed CNN model is shown in Fig 2 (see the left side).

2) **Random Decision Forests:** A random forest classifier consists of a collection of decision tree classifiers combined together to predict the class label, where each tree is grown in some randomized fashion. Each decision tree classifier consists of decision (or split) nodes and prediction (or leaf) nodes. The prediction node of each tree in the random forest classifier are labeled by the posterior distribution over the image classes [14]. Each decision node contains a test that splits best the space of data to be classified. An image is classified by sending it down the decision tree and aggregating the reached leaf posterior distributions. Randomness is usually injected at two points during training: in sub-sampling the training data and in selecting node tests. Each tree within the random forest classifier is binary and grown in a top-down manner. We choose the binary test at each node by maximizing the information gain,

$$\Delta E = - \sum_i \frac{|Q_i|}{|Q|} E(Q_i) \quad (2)$$

obtained by partitioning the training set  $Q$  of image patches into two sets  $Q_i$  according to a given test. Here  $E(q)$  is the entropy of the set  $q$  and  $|\cdot|$  is the size of the set. We repeat this selection process for each decision node until it reaches a certain depth.

Suppose  $T$  is the set of all trees,  $C$  is the set of all classes, and  $L$  is the set of all leaves for a give tree. During training the posterior probabilities ( $P_{t,l}(Y(I) = c)$ ) for each class  $c \in C$  at each leaf node  $l \in L$ , are found for each tree  $t \in T$ . These probabilities are calculated as the ratio of the number of images  $I$  of class  $c$  that reach a leaf node  $l$  to the total number of images that reach that leaf node  $l$ .  $Y(I)$  is the class label  $c$  for image  $I$ . During test time, we pass a new image through every decision tree until it reaches a prediction (or leaf) node, average all the posterior probabilities and classify the image as

$$\hat{Y}(I) = \arg \max_c \left\{ \frac{1}{|T|} \sum_{t=1}^{|T|} P_{t,l}(Y(I) = c) \right\} \quad (3)$$

where  $l$  is the leaf node reached by the image  $I$  in tree  $t$ . A graphical representation of the proposed random forest classifier is shown in Fig. 2 (see the right side).

### III. EXPERIMENTS AND RESULTS

The Leica TCS SP5 II laser scanning confocal microscope (Leica Microsystems, Buffalo Grove, IL, USA) was used in this work to capture the images. A Plan-Neofluar lens with a magnification of 4x, numerical aperture = 0.9, and a pixel size of 0.4  $\mu\text{m}$  in the x- and y-directions with automatic focusing was used to acquire the images. The size of each image in our dataset is  $2048 \times 2048$ . We used a total of 8 images consisting of a total of 2357 primary cilia, with 406 primary cilia within AQP2 expressed regions in the images and 1951 primary cilia elsewhere within the image. A careful manual detection and classification of all the 2357 primary cilia was performed by an

expert and considered as ground truth for subsequent analysis. We compared our proposed classifier with an unsupervised  $k$ -means classifier [15] and a supervised multilayer perceptron (MLP) classifier [16].

#### A. Algorithm Parameters

The convolutional neural random forest classifier is implemented using TensorFlow [17], and runs on an NVIDIA GeForce GTX TITAN X GPU with 8GB of memory. We used 70% of the data for training and 30% of the data for testing. A total of 284 primary cilia from AQP2 expressed regions and 1365 cilia from elsewhere within the images were used for training the algorithm. The classifier was trained using the stochastic gradient descent (SCD) algorithm, Adam [18], to efficiently optimize the weights of the CNN. The weights were normalized using initialization as proposed in [9] and updated in a mini-batch scheme of 128 candidates. The biases were initialized with zero, and the learning rate was set to  $\alpha = 0.001$ . The exponential decay rates for the first and second moment estimates were set as  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$ , respectively. We used an  $\epsilon = 10^{-8}$  to prevent division by zero. A dropout rate of 0.2 was implemented as regularization, applied to the output of the last convolutional layer and the dense layer to avoid overfitting. Finally, we used an epoch size of 50. The softmax loss (cross-entropy error loss) was utilized to measure the error loss. We used 100 estimators and a keep rate of  $\gamma = 10^{-4}$  for the random forests algorithm. A 5-fold cross-validation was used during training.

#### B. Performance Evaluation

The  $k$ -means, MLP, and the proposed convolutional neural random forest classifier were tested on 30% of the whole data consisting of 122 primary cilia in AQP2 expressed regions and 586 cilia elsewhere within the images. We evaluated all the algorithms using the conventional metrics that have been used for evaluation of classification algorithms, namely precision  $P$ , recall  $R$ , receiver operating characteristic (ROC) curves, area under the curve (AUC), and coverage measure ( $F_\beta$ -score).

Precision  $P$  and recall  $R$  are given by

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad R = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

where TP is the number of true positive classifications, FP is the number of false positive classifications, and FN is the number of false negative classifications.

An ROC curve is a plot between the true positive rate (a.k.a. sensitivity or recall ( $R$ )), which is defined by (4), and the false positive rate (a.k.a. complement of specificity), which is defined as  $\text{FP}/(\text{FP} + \text{FN})$ .

The coverage measure, also commonly known as the  $F_\beta$ -score is defined by

$$F_\beta = (1 + \beta^2) \frac{PR}{(\beta^2 P) + R} \quad (5)$$

We use  $F_1$  (i.e.,  $\beta = 1$ ) as this is the most common choice for this type of evaluation.

TABLE I  
PERFORMANCE OF THE CLASSIFICATION ALGORITHMS

Methods	Precision ( $P$ )	Recall ( $R$ )	AUC	$F_{\beta}$ -score
Our Method	0.9143	0.9062	0.8514	0.9102
MLP	0.8234	0.8239	0.8102	0.8237
$k$ -means	0.7961	0.8112	0.7891	0.8035

The AUC is the average of precision  $P(R)$  over the interval  $(0 \leq R \leq 1)$ , where  $P(R)$  is a function of recall  $R$ . It is given by

$$\text{AUC} = \int_0^1 P(R) dR. \quad (6)$$

The best classification algorithm among several alternatives is commonly defined as the one that maximizes either the AUC or the  $F_{\beta}$ -score.

Table I shows the average precision ( $P$ ), recall ( $R$ ), AUC, and  $F_{\beta}$ -score values for all the classification algorithms on the test data. From Table I we observe that the  $F_{\beta}$ -score of the proposed method is 10.67 percentage points greater than the  $k$ -means classifier, and is 8.65 percentage points greater than the MLP classifier. Table I also shows that the proposed classifier has the largest AUC among all the evaluated methods. Fig. 3 shows the ROC curves for all the methods under comparison. From Fig. 3, we observe that the proposed method has better classification accuracy compared to the other automated methods at all points along the curve.

#### IV. CONCLUSION

Accurate detection and classification of primary cilia in microscopy images is a challenging task. We propose a convolutional neural random forest classifier to classify primary cilia to determine whether they lie within an AQP2 expressed region or elsewhere within the microscopy images. We have shown how to model and train random forests, usable as alternative classifiers for batch learning in (deep) convolutional neural networks. Our approach combines the representation learning power of CNNs along with the divide-and-conquer principle of decision trees. We applied the proposed classifier to the problem of primary cilia classification in microscopy images and compared it with two methods, an unsupervised  $k$ -means classifier and a supervised MLP classifier. The results show that the proposed algorithm achieves better classification accuracy compared to the other two classifiers in terms of various figures of merit such as AUC, and  $F_{\beta}$ -score.

#### REFERENCES

- [1] K. Miyoshi, K. Kasahara, I. Miyazaki, and M. Asanuma, "Lithium Treatment Elongates Primary Cilia in the Mouse Brain and in Cultured Cells," *Biochemical and Biophysical Research Communications*, vol. 388, pp. 757–762, Oct. 2009.
- [2] E. Verghese, S. D. Ricardo, R. Weidenfeld, J. Zhuang, P. A. Hill, R. G. Langham, and J. A. Deane, "Renal Primary Cilia Lengthen after Acute Tubular Necrosis," *J. Am. Soc. Nephrol.*, vol. 20, pp. 2147–2153, Oct. 2009.

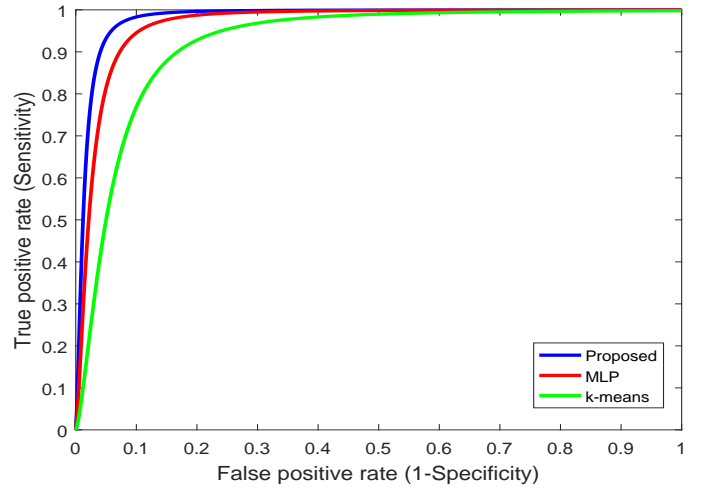


Fig. 3. ROC curves for all the classification methods.

- [3] Y. Gao, M. J. Romero-Aleshire, Q. Cai, T. J. Price, and H. L. Brooks, "Rapamycin Inhibition of Mtorc1 Reverses Lithium-Induced Proliferation of Renal Collecting Duct Cells," *American Journal of Physiology-Renal Physiology*, vol. 305, pp. F1201–F1208, Oct. 2013.
- [4] S. Ram, *Sparse Representations and Nonlinear Image Processing for Inverse Imaging Solutions*. PhD thesis, Dept. Elect. Comput. Engg., Univ. Arizona, Tucson, AZ, 2017.
- [5] S. Ram and J. J. Rodriguez, "Vehicle Detection in Aerial Images Using Multiscale Structure Enhancement and Symmetry," in *2016 IEEE Int. Conf. Image Process. ICIP*, (Phoenix, AZ, USA), pp. 3817–3821, IEEE, Sept. 2016.
- [6] S. Ram, F. Danford, S. Howerton, J. J. Rodriguez, and J. P. V. Geest, "Three-Dimensional Segmentation of the Ex-Vivo Anterior Lamina Cribrosa From Second-Harmonic Imaging Microscopy," *IEEE Trans. Biomed. Eng.*, vol. 65, pp. 1617–1629, July 2018.
- [7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, pp. 436–444, May 2015.
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Commun. ACM*, vol. 60, pp. 84–90, May 2017.
- [9] A. Gupta, A. Suveer, J. Lindblad, A. Dragomir, I.-M. Sintorn, and N. Sladoje, "Convolutional Neural Networks for False Positive Reduction of Automatically Detected Cilia in Low Magnification TEM Images," in *Image Anal.* (P. Sharma and F. M. Bianchi, eds.), vol. 10269, (Cham), pp. 407–418, Springer International Publishing, 2017.
- [10] Bo Zhang, J. Fadili, and J. Starck, "Wavelets, Ridgelets, and Curvelets for Poisson Noise Removal," *IEEE Trans. on Image Process.*, vol. 17, pp. 1093–1108, July 2008.
- [11] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Med. Image Comput. Comput.-Assist. Interv. – MICCAI 2015* (N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds.), vol. 9351, (Cham), pp. 234–241, Springer International Publishing, 2015.
- [12] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," 2014.
- [13] D. P. Kingma, T. Salimans, and M. Welling, "Variational Dropout and the Local Reparameterization Trick," in *Adv. Neural Inf. Process. Syst.*, 2015.
- [14] A. Bosch, A. Zisserman, and X. Munoz, "Image Classification Using Random Forests and Ferns," in *2007 IEEE 11th Int. Conf. Comput. Vis.*, (Rio de Janeiro, Brazil), pp. 1–8, IEEE, 2007.
- [15] M. Dundar, Q. Kou, B. Zhang, Y. He, and B. Rajwa, "Simplicity of Kmeans Versus Deepness of Deep Learning: A Case of Unsupervised Feature Learning with Limited Data," in *2015 IEEE 14th Int. Conf. Mach. Learn. Appl. ICMLA*, (Miami, FL, USA), pp. 883–888, IEEE, Dec. 2015.
- [16] S. S. Haykin, *Neural Networks and Learning Machines*. New York: Prentice Hall, third ed., 2009.
- [17] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow,

- A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "Tensorflow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," 2016.
- [18] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," 2014.