

LEVERAGING MACHINE LEARNING
FOR IMPROVED DATA INTERPRETATION AND DECISION MAKING
UNDER UNCERTAINTY

By

Mohammad S. Majdi

Copyright © Mohammad S. Majdi 2023

A Dissertation Submitted to the Faculty of the
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY

In the Graduate College
THE UNIVERSITY OF ARIZONA

2023

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation
prepared by: Mohammad S. Majdi
titled:

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of
Doctor of Philosophy.



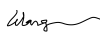
Jeffrey Rodriguez

Date: Aug 2, 2023



Carlos Alsua

Date: Aug 3, 2023



Janet Meiling Roveda

Date: Aug 8, 2023

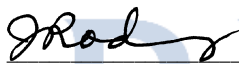


Abhijit Mahalanobis

Date: Aug 3, 2023

Final approval and acceptance of this dissertation is contingent upon the candidate's submission
of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend
that it be accepted as fulfilling the dissertation requirement.



Jeffrey Rodriguez

Dept. of Electrical and Computer Engineering

Date: Aug 2, 2023

STATEMENT by AUTHOR

This dissertation LEVERAGING MACHINE LEARNING FOR IMPROVED DATA INTERPRETATION AND DECISION MAKING UNDER UNCERTAINTY prepared by Mohammad S. Majdi has been submitted in partial fulfillment of requirements for a doctoral degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under the rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by

SIGNED: Mohammad S. Majdi

ACKNOWLEDGMENTS

As I reflect upon the conclusion of this significant journey, my heart swells with gratitude for those who have been unwavering pillars of support, wisdom, and guidance at every step.

Foremost, I extend my profound appreciation to my mentor, D. Jeffrey J. Rodriguez. His unyielding guidance and support throughout this venture have been unparalleled. The depth of his knowledge, relentless dedication, and ceaseless encouragement are beyond measure, and I remain eternally grateful. A special mention goes to my minor advisor in entrepreneurship, Dr. Carlos J. Alsua. His insights into the entrepreneurial aspects of my work have been pivotal in broadening my research horizons. His consistent support has been an integral part of my progress.

My heartfelt gratitude goes to Drs. Janet Roveda and Abhijit Mahalanobis, whose invaluable feedback and insights have immensely enhanced the quality of my work. Additionally, I owe a deep debt of thanks to Mr. Nirav Merchant, Ms. Maliaca Oxnam, and Drs. Manoj Saranathan, and Mahesh Keerthivasan. Their unwavering support has been crucial in this endeavor. The wisdom and mentorship provided by all have been fundamental to my growth both personally and academically.

Lastly, to my family, words fall short to convey my feelings. Your steadfast faith in me, patience in testing times, and endless motivation have been the bedrock of my achievements.

Contents

| | |
|---|-----------|
| List of Figures | 6 |
| List of Tables | 7 |
| List of Algorithms | 8 |
| 1 Introduction | 10 |
| 1.1 Introduction and Contextual Setting | 10 |
| 1.2 Underlying Motivation | 10 |
| 1.3 Problem Statement | 11 |
| 1.4 Research Objectives | 11 |
| 1.5 Research Queries | 12 |
| 1.6 Methodology Synopsis | 12 |
| 1.7 Scope and Limitations | 12 |
| 1.8 Dissertation's Organization | 13 |
| 2 Classification of Primary Cilia in Microscopy Images Using Convolutional Neural Random Forests | 15 |
| 2.1 Introduction | 16 |
| 2.2 Methods | 18 |
| 2.2.1 Preprocessing | 18 |
| 2.2.2 Convolutional Neural Random Forest | 18 |
| 2.3 Experiments and Results | 21 |
| 2.3.1 Algorithm Parameters | 21 |
| 2.3.2 Performance Evaluation | 21 |
| 2.4 Conclusion | 24 |

List of Figures

| | | |
|-----|--|----|
| 2.1 | Microscopy Images of Primary Cilia in Relation to AQP2-Expressing Renal Collecting Duct | 16 |
| 2.2 | Overview of the Convolutional Neural Random Forest Classifier: Integration of CNN Feature Mapping with Decision and Leaf Nodes | 19 |
| 2.3 | ROC Curves for Various Classification Methods | 23 |

List of Tables

| | | |
|-----|--|----|
| 2.1 | Performance of the Classification Algorithms | 23 |
|-----|--|----|

List of Algorithms

Abstract

Machine learning has emerged as a powerful tool for tackling complex problems across diverse domains. However, real-world applications of machine learning face inherent challenges due to complexity and variability in datasets. This dissertation explores machine learning techniques to address issues in crowdsourced labeling, medical diagnosis, neuroimaging, driver distraction detection, and biological image analysis. The objectives are to develop and evaluate innovative methods that enhance the accuracy, efficiency, and robustness of data interpretation. Proposed techniques include a weighted label aggregation scheme leveraging annotator consistency, hierarchical multi-label classification utilizing pathology taxonomy, cascaded neural networks for segmentation, and convolutional neural random forests for classification. By addressing challenges on improving reliability in crowdsourcing, leveraging taxonomic knowledge, transfer learning across modalities, and automating distraction and cilia detection, this dissertation aims to advance machine learning adoption in real life settings. The efficacy of introduced methods is assessed through comprehensive comparisons. By tackling these distinct challenges, this research provides broader insights into real-world applications of machine learning across data science, healthcare, transportation, and biology.

Chapter 1

Introduction

1.1 Introduction and Contextual Setting

Machine learning is a powerful technological tool that has progressively delivered cutting-edge solutions to complex problems in a variety of fields. Its application spans a wide array, from data interpretation and natural language processing to contemporary breakthroughs in domains such as computer vision, radiology, neuroimaging, and biological image analysis. This dissertation delves into distinct use cases, demonstrating the adaptability and capabilities of machine learning in solving varied challenges. Machine learning has signified a notable advancement in computational capacity, thereby enabling us to tackle issues once deemed insurmountable or too intricate. In sectors such as healthcare, machine learning algorithms help in disease diagnosis, predicting patient outcomes, and personalizing treatment plans. The automotive industry, with the emergence of self-driving vehicles, is undergoing a transformation propelled by machine learning. Likewise, in crowdsourcing, machine learning aids in aggregating information gathered from a vast populace to solve complex problems or drive insights.

1.2 Underlying Motivation

The rapid surge in digital data across all sectors necessitates innovative strategies for data analysis and interpretation. This escalating demand has been considerably addressed by machine learning, which leverages computational prowess to discover concealed patterns, make predictions, and aid in decision-making processes. However, despite progress, numerous challenges persist. These encompass issues such as the trustworthiness of crowdsourced labeling, efficient and reliable disease diagnosis,

and driver distraction detection. Additionally, practical applications of machine learning are plagued with instances where data can contain missing, incorrect, or uncertain values, which often leads to unpredictable model behavior and inconsistent performance. It is no longer sufficient to solely consider model accuracy; accounting for uncertainty in model predictions is equally crucial. Armed with a comprehensive understanding of model and data uncertainty, we can make informed decisions on whether to rely on model predictions or seek further information. This research endeavor aims to tackle these challenges, augmenting the capabilities of machine learning and deepening our understanding of its potential.

1.3 Problem Statement

The core issue this dissertation addresses lies in the inherent complexity and variability seen in real-world datasets, and the challenges these present for machine learning applications. Specifically, it explores five separate issues, each related to a distinct domain. Despite their differences, a common thread unites these issues - the requirement for advanced, resilient, and precise computational methods capable of effectively handling and analyzing large, complex datasets. Machine learning, while promising, presents its own set of challenges unique to each application. These include issues such as the varying reliability of crowd annotators, the overlap of radiographic indications of various thoracic diseases, and the similar attributes of non-cilia elements and imaging noise in biological image analysis.

1.4 Research Objectives

This dissertation seeks to introduce and assess new machine learning methodologies for each of the problem areas, each aiming to enhance the accuracy, efficiency, and robustness of data interpretation in its corresponding field. By proposing these methods and evaluating their performance, it aims to enrich the understanding and development of machine learning applications. More specifically, the research objectives include creating an improved label aggregation technique that factors in uncertainty to provide a weighted aggregation scheme; devising accurate classification methods that leverage the taxonomic structure of different classes; developing robust techniques that use less common high contrast medical images to enhance the detection accuracy of more common lower contrast images;

and establishing automated processes for detecting driver distraction and classifying primary cilia within microscopy images.

1.5 Research Queries

This dissertation seeks to answer several pivotal questions:

- Can an innovative method increase the reliability of crowdsourced labeling and ensemble learning?
- How can machine learning tactics boost the precision of diagnosis by leveraging the taxonomic structure of medical pathologies?
- Is it possible to enhance the accuracy of less frequent data modalities by utilizing more common ones?
- How can transfer learning be utilized to effectively detect and classify driver distraction as well as primary cilia within microscopy images?

1.6 Methodology Synopsis

Each chapter of this dissertation proposes a unique methodology developed to tackle its respective problem, utilizing different machine learning techniques, including label aggregation, hierarchical multi-label classification, cascaded multi-planar schemes, and convolutional neural networks fused with random decision forests. The efficacy of each method is then assessed and compared with existing techniques, thereby providing a comprehensive understanding of their effectiveness.

1.7 Scope and Limitations

This dissertation delves into machine learning applications across a variety of domains. While the methods introduced can be generalized, their evaluations are carried out on specific datasets. Consequently, the findings and conclusions are subject to the constraints and properties of these datasets.

Nonetheless, the proposed methods are scalable and adaptable, hinting at their wider application in the future. We cover a broad spectrum of machine learning applications, each addressing unique problems and methodologies. Although the primary focus is on the advancement of machine learning techniques, the impact of these methods stretches across diverse fields, including data science, radiology, road safety, and biology.

1.8 Dissertation's Organization

The structure of this dissertation is as follows.

Chapter ??: Crowd-Certain: Towards Robust Label Aggregation in Crowdsourced and Ensemble Learning Classification

This chapter confronts the challenges of crowdsourced labeling by introducing a novel method for label aggregation termed as Crowd-Certain. This approach leverages the consistency of annotators versus a trained classifier to determine the reliability of each annotator, offering robustness and computational efficiency. The chapter delves into the workings of Crowd-Certain, discusses its core features, and evaluates its performance against ten other label aggregation techniques and for ten distinct datasets.

Chapter ??: Leveraging Disease Taxonomy for Enhanced Multi-Label Classification in Chest Radiography

Chapter ?? focuses on diagnosing thoracic diseases from chest radiographs using deep learning techniques. The chapter introduces two innovative hierarchical multi-label classification methods leveraging pathology taxonomy to enhance the accuracy and interpretability of disease classifications. The techniques cater to scenarios with and without available ground truth, broadening their adaptability. The chapter explores the methods, their evaluations on three large chest radiograph datasets (CheXpert [1], PadChest [2], and NIH ChestXRay8 [3]), and the proposed methods' potential benefits and limitations.

Chapter ??: Automated Thalamic Nuclei Segmentation Using Multi-Planar Cascaded Convolutional Neural Networks

This chapter develops a fast and accurate method for the segmentation of thalamic nuclei using a convolutional neural network (CNN) based approach. The proposed method uses a cascaded multi-planar scheme with a modified residual U-Net architecture. The novel approach delivers remarkable performance in speed, accuracy, and versatility. It also demonstrates its utility in studying thalamic nuclei atrophy in MS patients, providing potential for further advancements in neuroimaging.

Chapter ??: Drive-Net: Convolutional Network for Driver Distraction Detection

This chapter presents Drive-Net, a novel supervised learning method that uses a CNN and a Random Decision Forest to classify images of drivers, effectively detecting driver distractions. Drive-Net's efficacy is validated through a comprehensive comparison with various other methods, achieving a higher detection accuracy rate. This chapter underlines the potential of automated methods in improving road safety measures.

Chapter 2: Classification of Primary Cilia in Microscopy Images Using Convolutional Neural Random Forests

The final chapter utilizes the technique proposed in Chapter ?? to accurately detect and classify primary cilia in microscopy images. The classifier combines the feature learning abilities of CNNs with the practicality of decision trees, offering a unique solution to the task of cilia classification. The chapter details the workings of the classifier and compares its performance with traditional classifiers, demonstrating superior classification accuracy. This novel method showcases the potential intersection of computer science and biology in diagnosing primary ciliary dyskinesia.

Chapter 2

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

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_β -score for classification accuracy. Results show that our proposed classifier achieves better classification accuracy.

KEYWORDS: Image classification, convolutional neural network, random forests, primary cilia, confocal microscopy.

2.1 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 [4] has shown that primary cilia length in mammalian cells may change due to extracellular environment stimuli, such as renal injury [5] 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) [6]. The collecting duct cells in the kidney express primary cilia, as shown in

Fig. 2.1 *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.

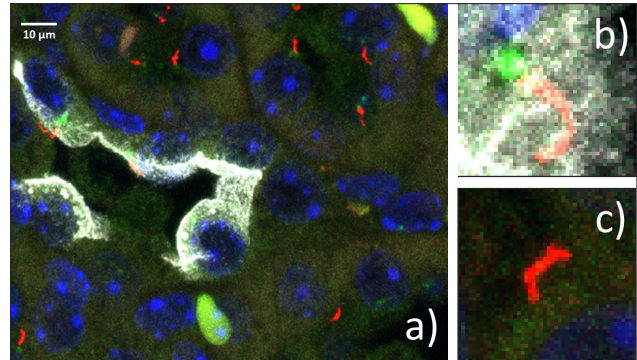


Figure 2.1: Microscopy image showing primary cilia which are in red color. (a) Cilia (red color) near AQP2-expressing renal collecting duct (white color). (b) Primary cilium within an AQP2 expressed region that we are interested in. (c) Primary cilium elsewhere in the image.

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 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 multiscale variance stabilizing transform (MS-VST) detector that are capable of detecting all the cilia locations within the microscopy image [7–9]. 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) [10] have outperformed the state-of-the-art in many computer vision applications [11]. Similarly, the applicability of CNNs has also been investigated for medical image analysis [12]. 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 [12].

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 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.

2.2 Methods

2.2.1 Preprocessing

For classification of primary cilia within our images, we first need to identify all potential primary cilia. In this work, we use the MS-VST algorithm [13] 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 32×32 pixels centered at the candidate primary cilium centroid. Such a patch size is chosen to contain the primary cilia ($\sim 20-25$ pixels in length), and some background around the cilium (~ 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.

2.2.2 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.

2.2.2.1 CNN Configuration

We adopt the U-Net architecture [14] 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 [15]. 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 1×1 convolution instead to obtain a fully connected layer. We use the rectified linear units (ReLU) [14] 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 [11]. 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 1×1 convolutional filters (as suggested in [12]) for the Adam [16] 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 32×32 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 2×2 (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

$$y = f(x) = \max(0, x) \quad (2.2.1)$$

A graphical representation of the architecture of the proposed CNN model is shown in Fig 2.2 (see the left side).

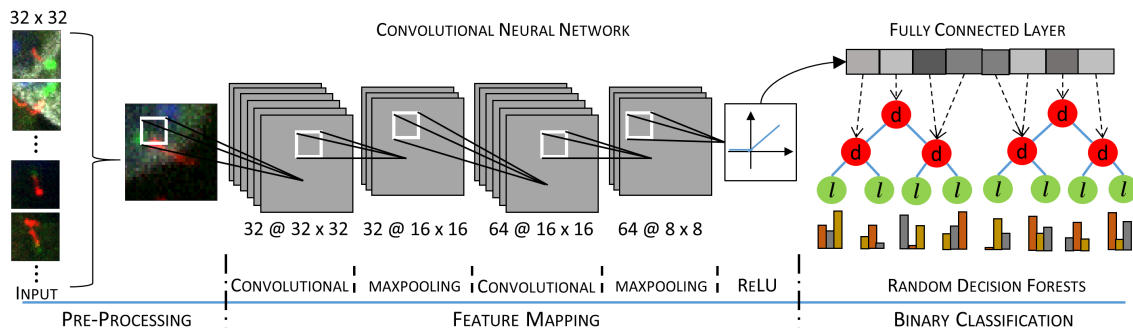


Figure 2.2: An overview of the proposed convolutional neural random forest classifier. Our proposed CNN model where the feature mapping happens (shown on the left side) consists of two convolutional layers each followed by a maxpooling layer and a final ReLU activation layer, following which a dropout 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.

2.2.2.2 Random Decision Forests

A random forest classifier consists of a collection of decision tree classifiers combined 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 [17]. 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,

$$\Delta E = - \sum_i \frac{|Q_i|}{|Q|} E(Q_i) \quad (2.2.2)$$

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

$$\hat{Y}(I) = \arg \max_c \left\{ \frac{1}{|T|} \sum_{t=1}^{|T|} P_{t,l}(Y(I) = c) \right\} \quad (2.2.3)$$

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. 2.2 (see the right side).

2.3 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 2048×2048 . 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 [18] and a supervised multilayer perceptron (MLP) classifier [19].

2.3.1 Algorithm Parameters

The convolutional neural random forest classifier is implemented using TensorFlow [20], 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 (SGD) algorithm, Adam [21], to efficiently optimize the weights of the CNN. The weights were normalized using initialization as proposed in [12] 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.

2.3.2 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_β -score).

Precision P and recall R are given by

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN} \quad (2.3.1)$$

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 (2.3.1), and the false positive rate (a.k.a. complement of specificity), which is defined as $FP/(FP + FN)$.

The coverage measure, also commonly known as the F_β -score is defined by

$$F_\beta = \left(1 + \beta^2\right) \frac{PR}{(\beta^2 P) + R} \quad (2.3.2)$$

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

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

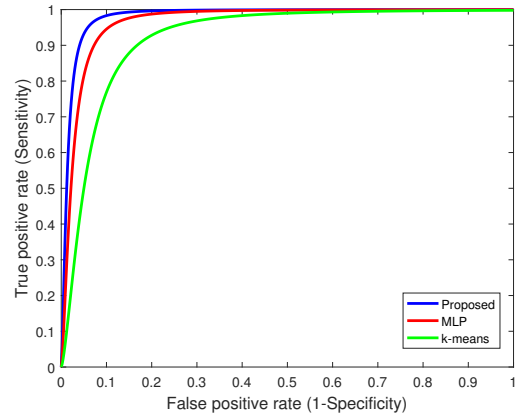
$$AUC = \int_0^1 P(R) dR. \quad (2.3.3)$$

The best classification algorithm among several alternatives is commonly defined as the one that maximizes either the AUC or the F_β -score.

Table 2.1: Performance of the Classification Algorithms

| Methods | Precision (P) | Recall (R) | AUC | F_β - 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 |

Table 2.1 shows the average precision (P), recall (R), AUC, and F_β -score values for all the classification algorithms on the test data. From Table 2.1 we observe that the F_β -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 2.1 also shows that the proposed classifier has the largest AUC among all the evaluated methods. Fig. 2.3 shows the ROC curves for all the methods under comparison. From Fig. 2.3, we observe that the proposed method has better classification accuracy compared to the

**Figure 2.3:** ROC curves for various classification methods.

2.4 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_β -score.

Bibliography

- [1] Jeremy Irvin et al. “CheXpert: A Large Chest Radiograph Dataset With Uncertainty Labels and Expert Comparison”. In: *Proc. AAAI Conf. Artif. Intell.* Vol. 33. July 17, 2019, pp. 590–597. doi: [10.1609/aaai.v33i01.3301590](https://doi.org/10.1609/aaai.v33i01.3301590). (Visited on 11/21/2022).
- [2] Aurelia Bustos et al. “Padchest: A Large Chest X-Ray Image Dataset With Multi-Label Annotated Reports”. In: *Medical Image Analysis* 66 (Dec. 2020), p. 101797. issn: 13618415. doi: [10.1016/j.media.2020.101797](https://doi.org/10.1016/j.media.2020.101797). (Visited on 01/07/2023).
- [3] Xiaosong Wang et al. “ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases”. In: *2017 IEEE Conf. Comput. Vis. Pattern Recognit. CVPR*. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI: IEEE, July 2017, pp. 3462–3471. isbn: 978-1-5386-0457-1. doi: [10.1109/CVPR.2017.369](https://doi.org/10.1109/CVPR.2017.369). (Visited on 04/02/2023).
- [4] Ko Miyoshi et al. “Lithium Treatment Elongates Primary Cilia in the Mouse Brain and in Cultured Cells”. In: *Biochemical and Biophysical Research Communications* 388.4 (Oct. 2009), pp. 757–762. issn: 0006291X. doi: [10.1016/j.bbrc.2009.08.099](https://doi.org/10.1016/j.bbrc.2009.08.099). (Visited on 06/03/2023).
- [5] Elizabeth Verghese et al. “Renal Primary Cilia Lengthen after Acute Tubular Necrosis”. In: *J. Am. Soc. Nephrol.* 20.10 (Oct. 2009), pp. 2147–2153. issn: 1046-6673. doi: [10.1681/ASN.2008101105](https://doi.org/10.1681/ASN.2008101105). (Visited on 06/03/2023).
- [6] Yang Gao et al. “Rapamycin Inhibition of Mtorc1 Reverses Lithium-Induced Proliferation of Renal Collecting Duct Cells”. In: *American Journal of Physiology-Renal Physiology* 305.8 (Oct. 15, 2013), F1201–F1208. issn: 1931-857X, 1522-1466. doi: [10.1152/ajprenal.00153.2013](https://doi.org/10.1152/ajprenal.00153.2013). (Visited on 06/03/2023).
- [7] Sundaresh Ram. “Sparse Representations and Nonlinear Image Processing for Inverse Imaging Solutions”. PhD thesis. Tucson, AZ: The University of Arizona., 2017. url: <http://hdl.handle.net/10150/626164>.

- [8] Sundaresh Ram and Jeffrey J. Rodriguez. "Vehicle Detection in Aerial Images Using Multiscale Structure Enhancement and Symmetry". In: *2016 IEEE Int. Conf. Image Process. ICIP*. 2016 IEEE International Conference on Image Processing (ICIP). Phoenix, AZ, USA: IEEE, Sept. 2016, pp. 3817–3821. isbn: 978-1-4673-9961-6. doi: [10.1109/ICIP.2016.7533074](https://doi.org/10.1109/ICIP.2016.7533074). (Visited on 06/03/2023).
- [9] Sundaresh Ram et al. "Three-Dimensional Segmentation of the Ex-Vivo Anterior Lamina Cribrosa From Second-Harmonic Imaging Microscopy". In: *IEEE Trans. Biomed. Eng.* 65.7 (July 2018), pp. 1617–1629. issn: 0018-9294, 1558-2531. doi: [10.1109/TBME.2017.2674521](https://doi.org/10.1109/TBME.2017.2674521). (Visited on 06/03/2023).
- [10] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep Learning". In: *Nature* 521.7553 (May 28, 2015), pp. 436–444. issn: 0028-0836, 1476-4687. doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539). (Visited on 06/03/2023).
- [11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Commun. ACM* 60.6 (May 24, 2017), pp. 84–90. issn: 0001-0782, 1557-7317. doi: [10.1145/3065386](https://doi.org/10.1145/3065386). (Visited on 01/11/2023).
- [12] Anindya Gupta et al. "Convolutional Neural Networks for False Positive Reduction of Automatically Detected Cilia in Low Magnification TEM Images". In: *Image Anal.* Ed. by Puneet Sharma and Filippo Maria Bianchi. Vol. 10269. Cham: Springer International Publishing, 2017, pp. 407–418. isbn: 978-3-319-59125-4 978-3-319-59126-1. doi: [10.1007/978-3-319-59126-1_34](https://doi.org/10.1007/978-3-319-59126-1_34). (Visited on 06/03/2023).
- [13] Bo Zhang, J.M. Fadili, and J.L. Starck. "Wavelets, Ridgelets, and Curvelets for Poisson Noise Removal". In: *IEEE Trans. on Image Process.* 17.7 (July 2008), pp. 1093–1108. issn: 1057-7149, 1941-0042. doi: [10.1109/TIP.2008.924386](https://doi.org/10.1109/TIP.2008.924386). (Visited on 06/03/2023).
- [14] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Med. Image Comput. Comput.-Assist. Interv. – MICCAI 2015*. Ed. by Nassir Navab et al. Vol. 9351. Cham: Springer International Publishing, 2015, pp. 234–241. isbn: 978-3-319-24573-7 978-3-319-24574-4. doi: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28). (Visited on 01/11/2023).
- [15] Karen Simonyan and Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. Version 6. 2014. doi: [10.48550/ARXIV.1409.1556](https://doi.org/10.48550/ARXIV.1409.1556). arXiv: [1409.1556](https://arxiv.org/abs/1409.1556). (Visited on 06/03/2023). preprint.

- [16] Diederik P. Kingma, Tim Salimans, and Max Welling. “Variational Dropout and the Local Reparameterization Trick”. In: *Adv. Neural Inf. Process. Syst.* 2015.
- [17] Anna Bosch, Andrew Zisserman, and Xavier Munoz. “Image Classification Using Random Forests and Ferns”. In: *2007 IEEE 11th Int. Conf. Comput. Vis.* 2007 IEEE 11th International Conference on Computer Vision. Rio de Janeiro, Brazil: IEEE, 2007, pp. 1–8. isbn: 978-1-4244-1630-1. doi: [10.1109/ICCV.2007.4409066](https://doi.org/10.1109/ICCV.2007.4409066). (Visited on 01/11/2023).
- [18] Murat Dundar et al. “Simplicity of Kmeans Versus Deepness of Deep Learning: A Case of Unsupervised Feature Learning with Limited Data”. In: *Proc. 14th Int. Conf. Mach. Learn. Appl. ICMLA*. 14th International Conference on Machine Learning and Applications (ICMLA). Miami, FL, USA: IEEE, Dec. 2015, pp. 883–888. isbn: 978-1-5090-0287-0. doi: [10.1109/ICMLA.2015.78](https://doi.org/10.1109/ICMLA.2015.78).
- [19] Simon S. Haykin. *Neural Networks and Learning Machines*. 3rd ed. New York: Prentice Hall, 2009. 906 pp. isbn: 978-0-13-147139-9.
- [20] Martín Abadi et al. *Tensorflow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. Version 2. 2016. doi: [10.48550/ARXIV.1603.04467](https://doi.org/10.48550/ARXIV.1603.04467). arXiv: [1603.04467](https://arxiv.org/abs/1603.04467). (Visited on 06/03/2023). preprint.
- [21] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. Version 9. 2014. doi: [10.48550/ARXIV.1412.6980](https://doi.org/10.48550/ARXIV.1412.6980). arXiv: [1412.6980](https://arxiv.org/abs/1412.6980) [cs.LG]. (Visited on 01/11/2023). preprint.