Center for Research in Computer Vision

UNIVERSITY OF CENTRAL FLORIDA

FINAL ORAL EXAMINATION

OF

Rodney LaLonde III

M.Sc., University of Central Florida, 2019
B.A., Saint Olaf College, 2012

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

(COMPUTER SCIENCE)

24 June, 2020, 2:00 P.M.
https://tinyurl.com/RLaLondePhDDefenseZoom

DISSERTATION COMMITTEE

Professor Ulas Bagci, *Chair*, *bagci@cs.ucf.edu*
Professor Mubarak Shah, *shah@crcv.ucf.edu*
Professor Abhijit Mahalanobis, *amahalan@crcv.ucf.edu*
Professor Michael B. Wallace, w*allace.michael@mayo.edu*

DISSERTATION RESEARCH IMPACT

Convolutional neural networks, despite their profound impact in countless domains, suffer from spatial ambiguities and a lack of robustness to pose variations. Capsule networks can likely alleviate these issues by storing and routing the pose information of features through their architectures, seeking agreement between the lower-level predictions of higher-level poses.

We make several contributions to advance the algorithms of capsule networks, with specific real-world applications in biomedical imaging data segmentation and classification: (1) The first ever capsule-based segmentation network in the literature, ***SegCaps***, provided five major novelties. (2) A capsule-based diagnosis network, ***D-Caps***, introduced a novel capsule-average pooling layer. (3) An explainable capsule network, ***X-Caps***, encodes high-level visual object attributes within its capsules by utilizing a multi-task framework and a novel routing sigmoid function, to provide predictions with human-level explanations and a meaningful confidence score.

SELECTED PUBLICATIONS & PATENTS

1. Rodney LaLonde, Dong Zhang and Mubarak Shah. "ClusterNet: Detecting Small Objects in Large Scenes by Exploiting Spatio-Temporal Information." *IEEE International Conference on Computer Vision and Pattern Recognition (**CVPR**), 2018.*
2. Rodney LaLonde and Ulas Bagci. "Capsules for Object Segmentation." *Medical Imaging with Deep Learning (**MIDL**), 2018.* **[Oral, CIFAR Award].**
3. Rodney LaLonde, Irene Tanner, Katerina Nikiforaski, Georgios Z. Papadakis, Candice W. Bolan, Michael B. Wallace, and Ulas Bagci. "INN: Inflated Neural Networks for IPMN Diagnosis." *International Conference on Medical Image Computing and Computer-Assisted Intervention (**MICCAI**), 2019.*
4. Rodney LaLonde and Ulas Bagci. "Capsules for Image Analysis." ***U.S. Patent*** *Application 16/431,387; 2019.*
5. Rodney LaLonde, Pujan Kandel, Concetto Spampinato, Michael B. Wallace, and Ulas Bagci. "Diagnosing Colorectal Polyps in the Wild with Capsule Networks." IEEE *International Symposium on Biomedical Imaging (**ISBI**), 2020.*
6. Rodney LaLonde, Drew Torigian, and Ulas Bagci. "Encoding Visual Attributes in Capsules for Explainable Medical Diagnoses." *International Conference on Medical Image Computing and Computer-Assisted Intervention (**MICCAI**), 2020.*
7. Rodney LaLonde, Candice W. Bolan, Michael B. Wallace, and Ulas Bagci. "Systems, Methods, and Media for Automatically Diagnosing Intraductal Papillary Mucinous Neoplasms Using Multi-Modal Magnetic Resonance Imaging Data." ***U.S. Patent*** *Application 16/459,437; 2020.*
8. Rodney LaLonde, Ziyue Xu, Sanjay Jain, and Ulas Bagci. "Capsules for Biomedical Image Segmentation." *Elsevier, **Medical Image Analysis**,* [Under Revision].

DISSERTATION

ALGORITHMS AND APPLICATIONS OF NOVEL CAPSULE NETWORKS

Deep learning methodologies, in particular convolutional neural networks (CNNs), have made a profound impact in countless domains across academia, government, and industry. Nonetheless, many have argued strongly against some of their core mechanisms. Capsule networks are a class of neural networks which aim to solve these shortcomings by storing both presence and pose information about extracted features, and route this information through the network seeking pose agreement between lower- and higher-level features. In this dissertation, we make several key contributions to advance the algorithms and applications of capsule neural networks. Specific focus is given to biomedical image applications for their significance in potentially life-saving technologies.

We introduce the first ever capsule network designed for the task of segmentation in the literature. This required several important advancements, including a novel locally-constrained dynamic routing algorithm, transformation matrix sharing, the concept of a "deconvolutional" capsule, extension of the reconstruction regularization to segmentation, and a new encoder-decoder capsule network structure. These advancements culminate in an architecture which we call ***SegCaps***. *SegCaps* outperforms state-of-the-art CNNs in pathological lung segmentation for both clinical and preclinical subjects, as well as retinal vessel segmentation, while using only a small fraction of the parameters as those CNNs. Further, we provide evidence that *SegCaps* can generalize to unseen poses of familiar objects far better than a state-of-the-art CNN.

Next, we design a capsule-based network for the task of diagnosis in the field of endoscopy. In order to classify real-world imaging data much larger in size than those in MNIST or CIFAR, we introduced the concept of *capsule-average pooling*. Our proposed architecture, which we call ***D-Caps***, combines this capsule-average pooling with the parameter saving techniques introduced in *SegCaps* to diagnose colorectal polyps from colonoscopy images. We hypothesize that D-Caps should be able to better handle the relatively limited training data and high intra-class variation present in our colorectal polyp dataset. We conducted a set of thorough experiments to validate our hypothesis, stratified across all polyp categories, imaging devices and modalities, and focus modes available. Our results show *D-Caps* can outperform the leading state-of-the-art CNN-based method by as much as 43% in the most difficult settings.

Lastly, we introduce algorithmic advances in capsule networks to improve the explainability of network predictions. CNN-based systems have largely not been adopted in many high-risk application areas, including healthcare, military, security, transportation, finance, and legal, due to their highly uninterpretable ``black-box" nature. Towards solving this deficiency, we teach a capsule network to explain its predictions using the same high-level language used by human-experts. Our explainable capsule network, ***X-Caps***, encodes high-level visual object attributes within the vectors of its capsules, then forms predictions based solely on these human-interpretable features. We implement a multi-task learning framework to learn the attribute and malignancy scores from a large multi-center dataset of lung cancer screening patients. *X-Caps* utilizes a routing sigmoid to independently route information from child capsules to parents for the visual attribute vectors. To estimate model confidence, we train our network on a distribution of expert labels, modeling inter-observer agreement and punishing over/under confidence during training supervised by human-experts' agreement.

Capsule networks show considerable promise for the future of deep learning-based applications and we hope the contributions of this dissertation provide a solid foundation for the further advancement of capsule network-based approaches. The source code for all of our algorithms has been made publicly available at https://github.com/lalonderodney.

Rodney LaLonde

1992	Born in Detroit, MI
2012-2016	B.A., Computer Science, St. Olaf College, Northfield, MN
2012-2016	B.A., Mathematics, St. Olaf College, Northfield, MN
2012-2016	B.A., Physics, St. Olaf College, Northfield, MN
2019	M.Sc., University of Central Florida, Orlando, FL
2016-2020	Ph.D., University of Central Florida, Orlando, FL
2017-2018	Machine Learning Intern, Lockheed Martin, Sunnyvale, CA
2018	Machine Learning Intern, Armored Things, Redmond, WA
2019	Machine Learning Intern, Aptiv, Santa Monica, CA