



Center for Research in Computer Vision

UNIVERSITY OF CENTRAL FLORIDA

FINAL ORAL EXAMINATION

OF

Mamshad Nayeem Rizve

M.S., University of Central Florida, 2022

B.S., Bangladesh University of Engineering and Technology, 2016

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
(COMPUTER SCIENCE)

7 July 2023, 10:30 A.M.
Engineering II - Room 103

DISSERTATION COMMITTEE

Professor Mubarak Shah, *Chair*, shah@crcv.ucf.edu

Professor Yogesh Rawat, yogesh@crcv.ucf.edu

Professor Yanjie Fu, yanjie.fu@ucf.edu

Professor Qiyu Sun, qiyu.sun@ucf.edu

DISSERTATION RESEARCH IMPACT

Deep learning has advanced visual understanding tasks like classification, segmentation, object detection, tracking, and more. Fields such as autonomous vehicles, surveillance systems, medical imaging, robot manipulation, and augmented reality have greatly benefited from these advancements. Large-scale labeled datasets have played a crucial role in this success. However, acquiring labeled data can be impractical and expensive in many real-world scenarios because of the rarity of events or time and resource constraints. Additionally, the progress achieved so far has mostly been in controlled environments, which hinders the models' ability to handle new concepts in unconstrained environments.

This dissertation makes several contributions to address these fundamental challenges in deep learning models for visual recognition. We introduce an uncertainty-aware pseudo-label selection framework to improve semi-supervised learning and reduce annotation costs. We also propose two methods to tackle semi-supervised learning in open-world settings. Additionally, we focus on few-shot learning and propose a novel training mechanism to learn better generalizable features.

SELECTED PUBLICATIONS (h-index: 9, total citation: 620)

- 1. Gabriella: An online system for real-time activity detection in untrimmed security videos**, [MN Rizve](#), U Demir, P Tirupattur, AJ Rana, K Duarte, I Dave, YS Rawat and M Shah, in *International Conference on Pattern Recognition*, 2020
- 2. In Defense of Pseudo-labeling: An Uncertainty-aware Pseudo-label Selection Framework for Semi-Supervised Learning**, [MN Rizve](#), K Duarte, YS Rawat and M Shah, in *International Conference on Learning Representations*, 2021.
- 3. Exploring Complementary Strengths of Invariant and Equivariant Representations for Few-Shot Learning**, [MN Rizve](#), S Khan, FS Khan and M Shah, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.
- 4. TCLR: Temporal Contrastive Learning for Video Representation**, I Dave, R Gupta, [MN Rizve](#) and M Shah, in *Computer Vision and Image Understanding*, 2022.
- 5. Unicorn: Combating Label Noise through Uniform Selection and Contrastive Learning**, N Karim, [MN Rizve](#), N Rahnavard, A Mian and M Shah, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.
- 6. OpenLDN: Learning to Discover Novel Classes for Open-World Semi-Supervised Learning**, [MN Rizve](#), N Kardan, S Khan, FS Khan and M Shah, in *European Conference on Computer Vision*, 2022.
- 7. Towards Realistic Semi-Supervised Learning**, [MN Rizve](#), N Kardan and M Shah, in *European Conference on Computer Vision*, 2022.
- 8. TimeBalance: Temporally-Invariant and Temporally-Distinctive Video Representations for Semi-Supervised Action Recognition**, I Dave, [MN Rizve](#), C Chen and M Shah, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023.
- 9. PivoTAL: Prior-Driven Supervision for Weakly-Supervised Temporal Action Localization**, [MN Rizve](#), G Mittal, Y Yu, M Hall, S Sajeev, M Shah and M Chen, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023.

DISSERTATION

ANNOTATION EFFICIENT VISUAL RECOGNITION: FROM SEMI-SUPERVISED TO FEW-SHOT LEARNING

In recent years, supervised deep learning has achieved remarkable success in solving a wide range of visual recognition problems. Large-scale labeled datasets have been crucial for this success and the progress has primarily been limited to controlled environments. In this dissertation, we present methods to improve the annotation efficiency of deep visual recognition models and also propose methods to improve the performance of annotation-efficient models in unconstrained open-world settings.

To address the annotation bottleneck in supervised learning, we introduce a pseudo-labeling framework for semi-supervised learning. While consistency regularization methods dominate the field, they heavily rely on domain-specific data augmentations, limiting their applicability. We argue that even though pseudo-labeling is a general approach, it performs poorly due to high-confidence predictions from poorly calibrated models, leading to noisy training. To overcome this, we propose an uncertainty-aware pseudo-label selection method that greatly reduces the amount of noisy pseudo-labels. Furthermore, our proposed framework generalizes the pseudo-labeling process, allowing for the creation of negative pseudo-labels; these negative pseudo-labels can be used for multi-label classification as well as negative learning to improve the single-label classification.

Even though the previous semi-supervised learning method is very effective in reducing annotation costs, it is not suitable for real-world scenarios where there is limited control over the data collection process. Hence, our next focus is on open-world semi-supervised learning, which assumes that labeled and unlabeled data come from different distributions and unlabeled data may contain samples from unknown classes. We propose a method that utilizes a pairwise similarity loss to discover novel classes by implicitly clustering them while recognizing samples from known classes. Using a bi-level optimization rule this pairwise similarity loss exploits the information available in the labeled set. After discovering novel classes, our proposed method transforms the open-world semi-supervised learning problem into a standard semi-supervised learning problem to achieve additional performance gains using existing semi-supervised learning methods.

Despite being effective the previous solution relies on multiple objective functions, requires prior knowledge of the number of unknown classes, and only works on class-balanced data. To overcome these limitations, we propose a second, more practical and streamlined solution for open-world semi-supervised learning. Our proposed solution utilizes sample uncertainty and incorporate prior knowledge about class distribution to generate reliable class-distribution-aware pseudo-labels for unlabeled data belonging to both known and unknown classes. This method works with any distribution and does not require knowing the number of novel classes, making it more practical for deployment.

In the previous two works, we assume that samples from novel classes are available during training. However, this is difficult to satisfy in many real-world scenarios. Therefore, to study how models can achieve human-like performance in open-world settings, i.e., identifying new concepts with only a few examples, we shift our focus to few-shot learning. Learning good generalizable features is crucial for solving this problem. As a result, we propose a novel training mechanism that simultaneously enforces equivariance and invariance to a general set of geometric transformations. Simultaneous optimization for both of these contrasting objectives allows the model to jointly learn features that are not only independent of the input transformation but also the features that encode the structure of geometric transformations. These complementary sets of features help generalize well to novel classes with only a few labeled data samples. We achieve additional improvements by incorporating a novel self-supervised distillation objective.



Mamshad Nayeem Rizve

1992	Born in Kishoreganj, Bangladesh
2011-2016	B.S., Bangladesh University of Engineering and Technology
2016-2018	System Engineer, Grameenphone Limited
2021	Software Engineering Intern, Aurora Innovation
2022	Research Intern, Microsoft