Summary 16

Shaun Pritchard

Florida Atlantic University

CAP 6778

November 10, 2021

M. Khoshgoftaa

**Investigating the Generalization of Image Classifiers with Augmented Test Sets**

The purpose of this study is to implement experiments that introduce prior knowledge into deep learning computer vision models through the augmentation of data. The aim of this study is to investigate inductive bias in architecture and augmentation through evaluating generalization to test sets using data augmentation policies such as RandAugemnt on ResNet and Vision Transformer models.

For the experiment, the WILDS benchmark was used to simulate real-world data distribution shifts. Study objectives include understanding how data augmentation policies optimized for convolutional networks would translate to the Vision Transformer, as well as understanding the performance of the new models.

Data Augmentation in general consists of a set of techniques to enhance the size and quality of training datasets in order to construct more accurate Deep Learning models. The inductive bias that labels transformation invariants arises from enhancements such as rotations, horizontal flips, and crops. An inductive bias is a set of assumptions that a learner makes to predict the output of unknown inputs. Using machine learning, one aims to construct algorithms that can predict a specific outcome.

Through this evaluation of priors in augmentations, they aimed to show the relationships with priors in architecture design through the testing of ResNet and Vision transformer models. In the study, it was also a goal to curate functional diversity for ensemble learning. This will also require a better understanding of the performance of different architectural biases.

In the end, the results showed that the ResNet and Vision Transformer models did not generalize to test sets built with when trained with the RandAugment augmentation policy, data augmentation still occurs. However, not enough evidence was seen to isolate the effect of the priors in terms of architecture and augmentation in experiments with Vision Transformers and ResNets.