How Data Augmentation Affects What Neural Networks Learn

Phuc Ngo[1], Dimitris Tsipras[2], Saachi Jain[3] and Aleksander Mądry[4]

[1]Department of Computer Science and Maths, Beloit College. [2, ,3, 4]Department of Electrical Engineering & Computer Science, Massachusetts Institute of Technology

Introduction

- Data augmentation is a simple and common technique that increases the model's robustness to class-preserving transformations
- Our understanding of this technique is still limited
- We try to understand the effect of data augmentation on neural network

Hypothesis

Invariance

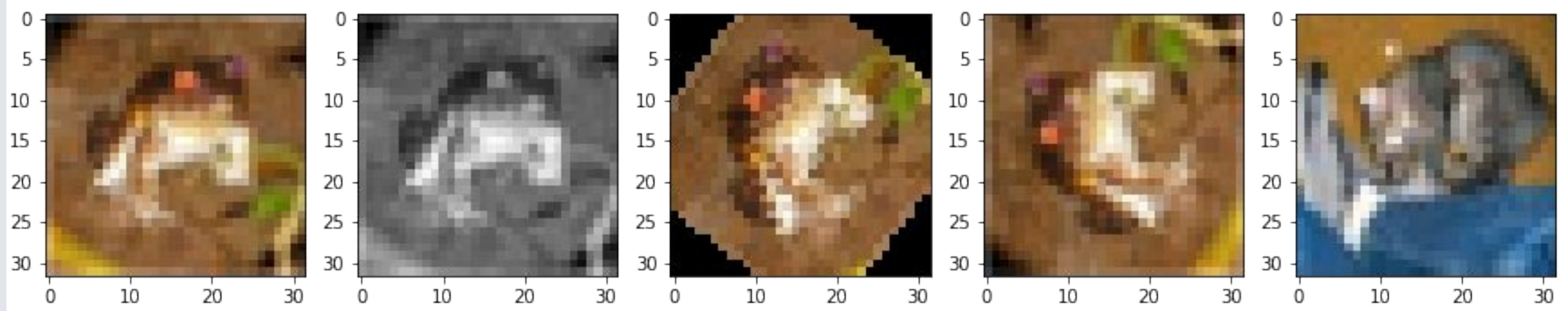
- Model maps augmented inputs to similar representations to the standard inputs

Subpopulation

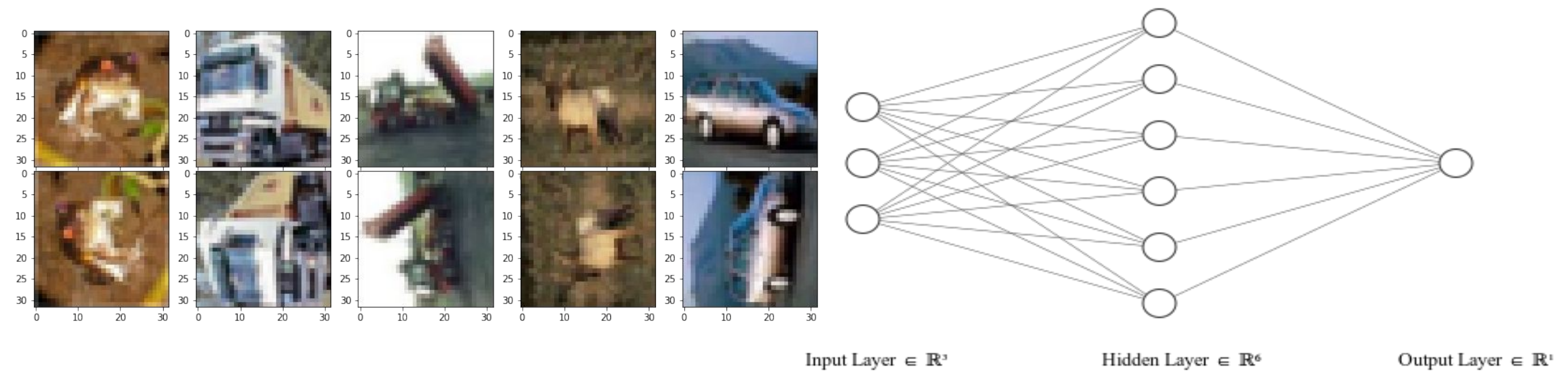
- Model uses a different set of prediction rules to classify augmented samples

Method

Using CIFAR10, we trained models using 2 types of augmentation: grayscale and 90-rotation.



Procedure of training an augmented model:



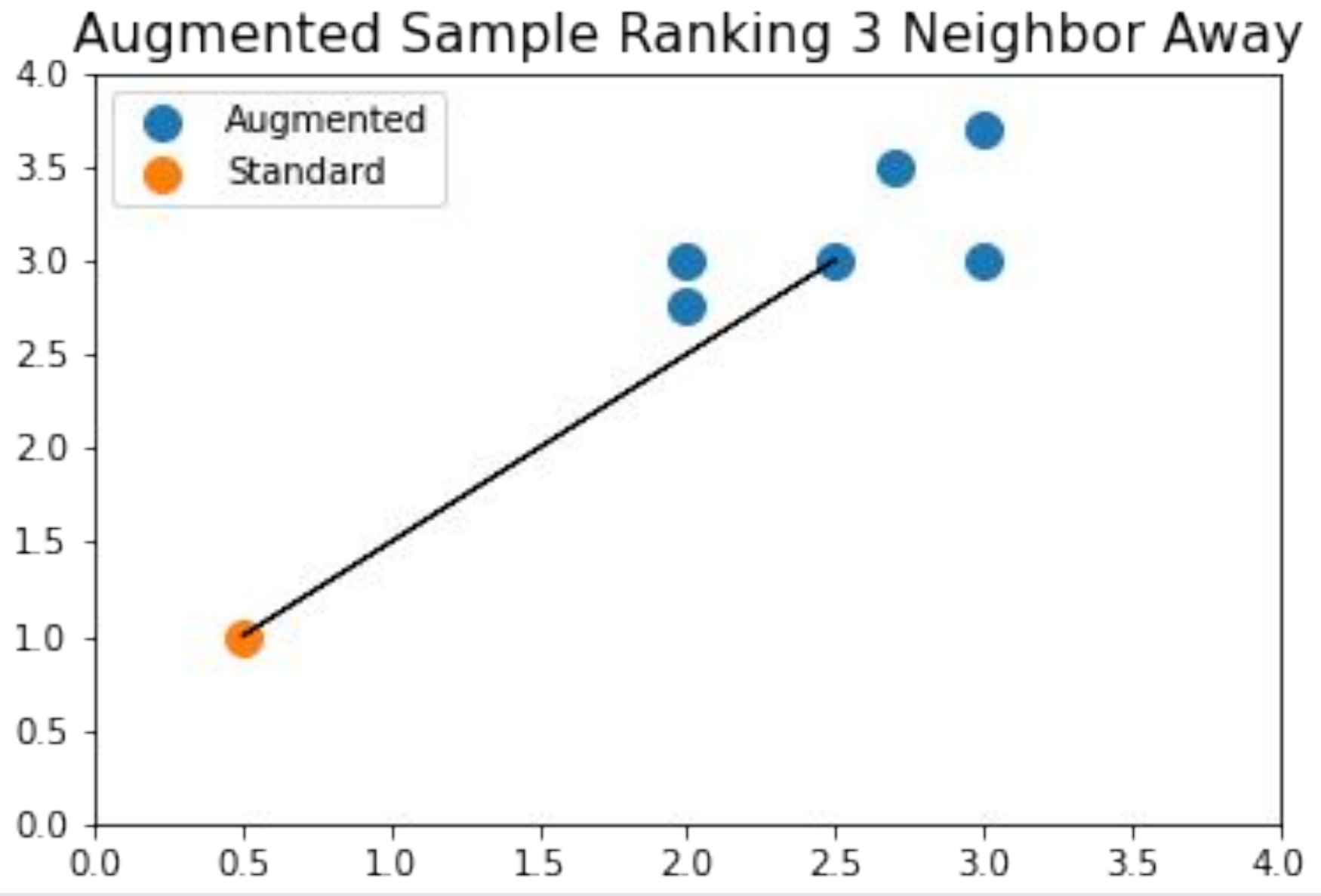
We then obtained models' prediction and final feature vector on standard and augmented testsets to calculate:

- Accuracy
- Correlation:

1	0	1	0	1	0
0	0	1	1	0	1

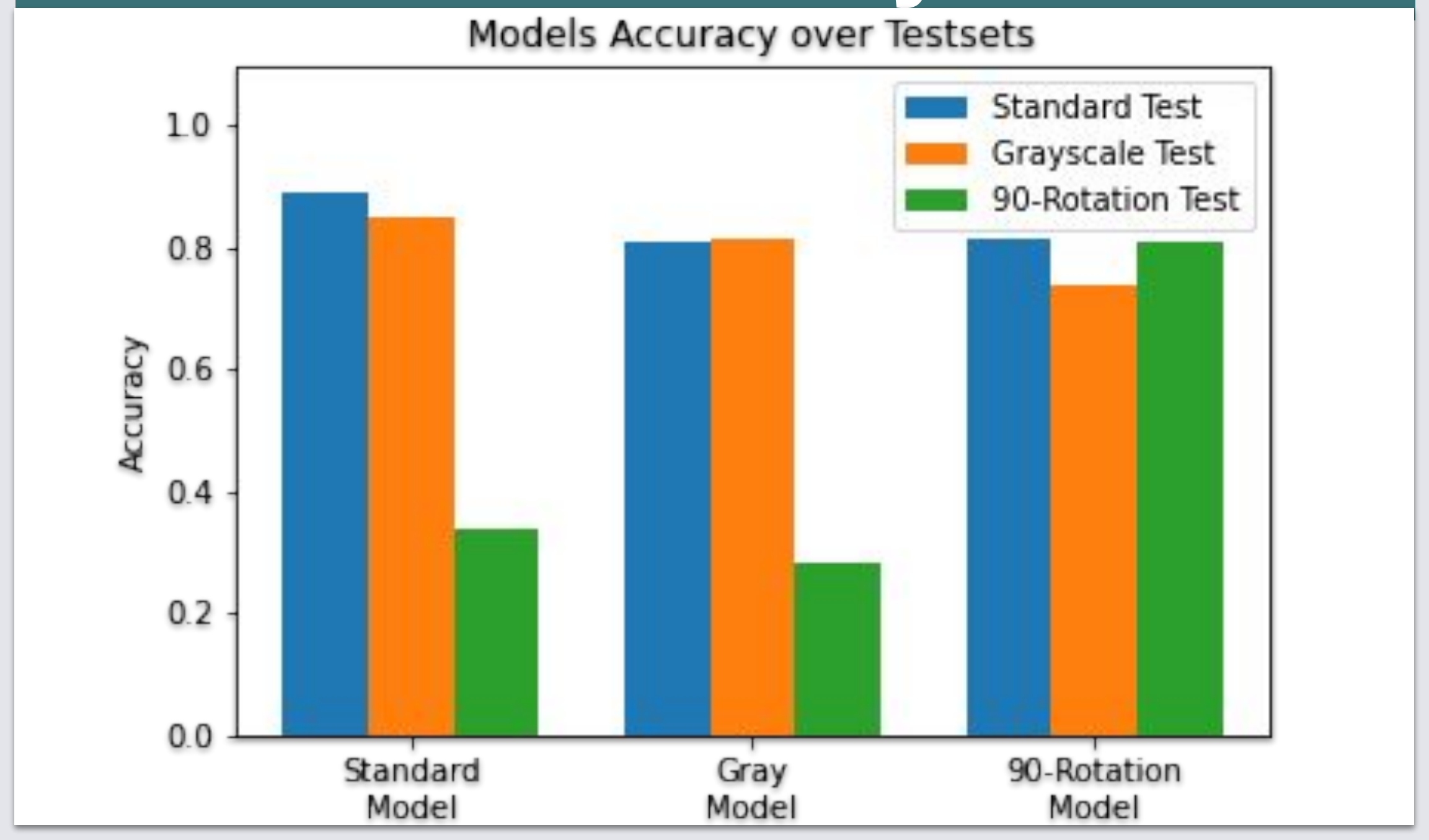
→ Negative Correlation

- Nearest neighbor diagram:



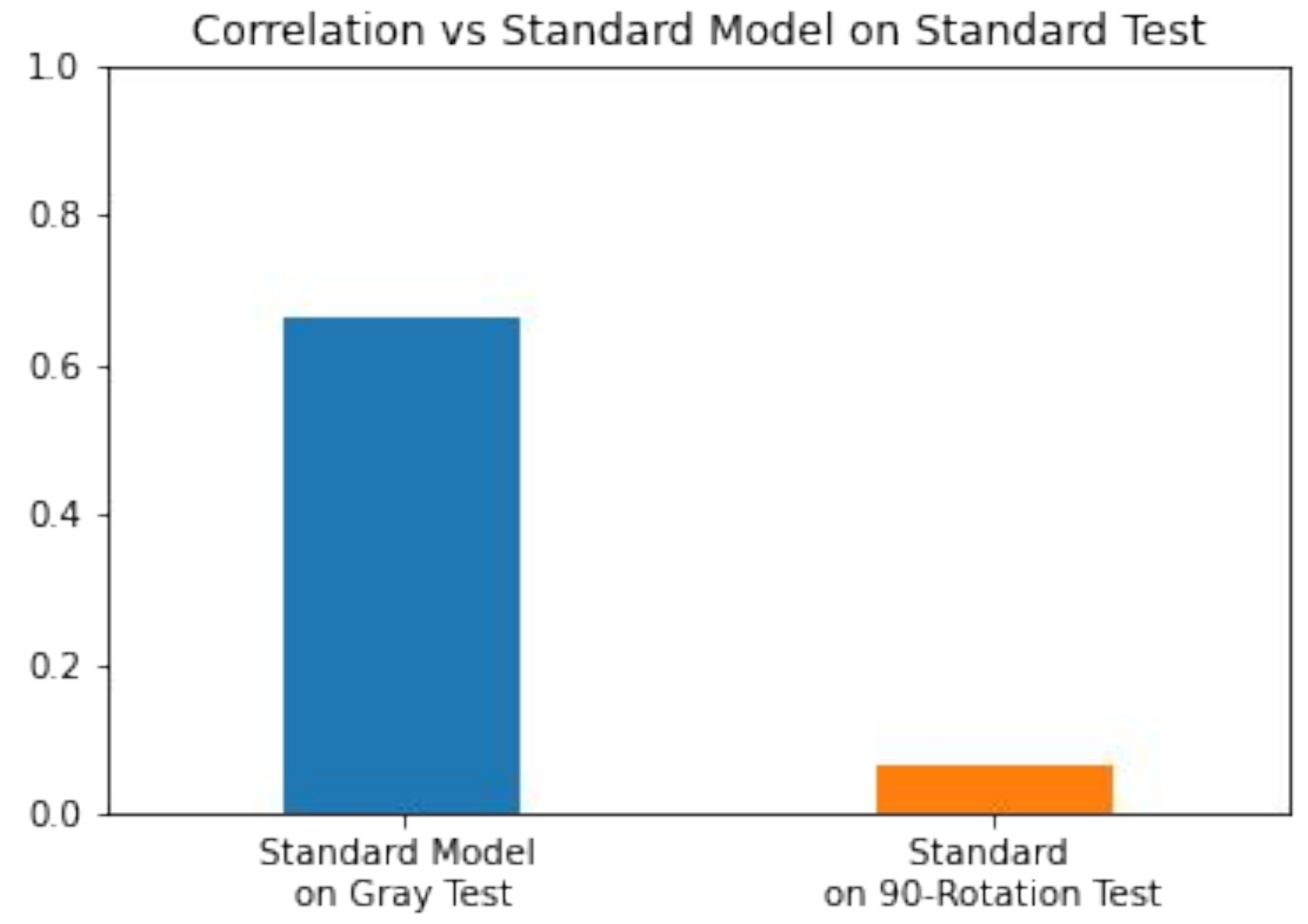
Results

Accuracy



Correlation (Standard)

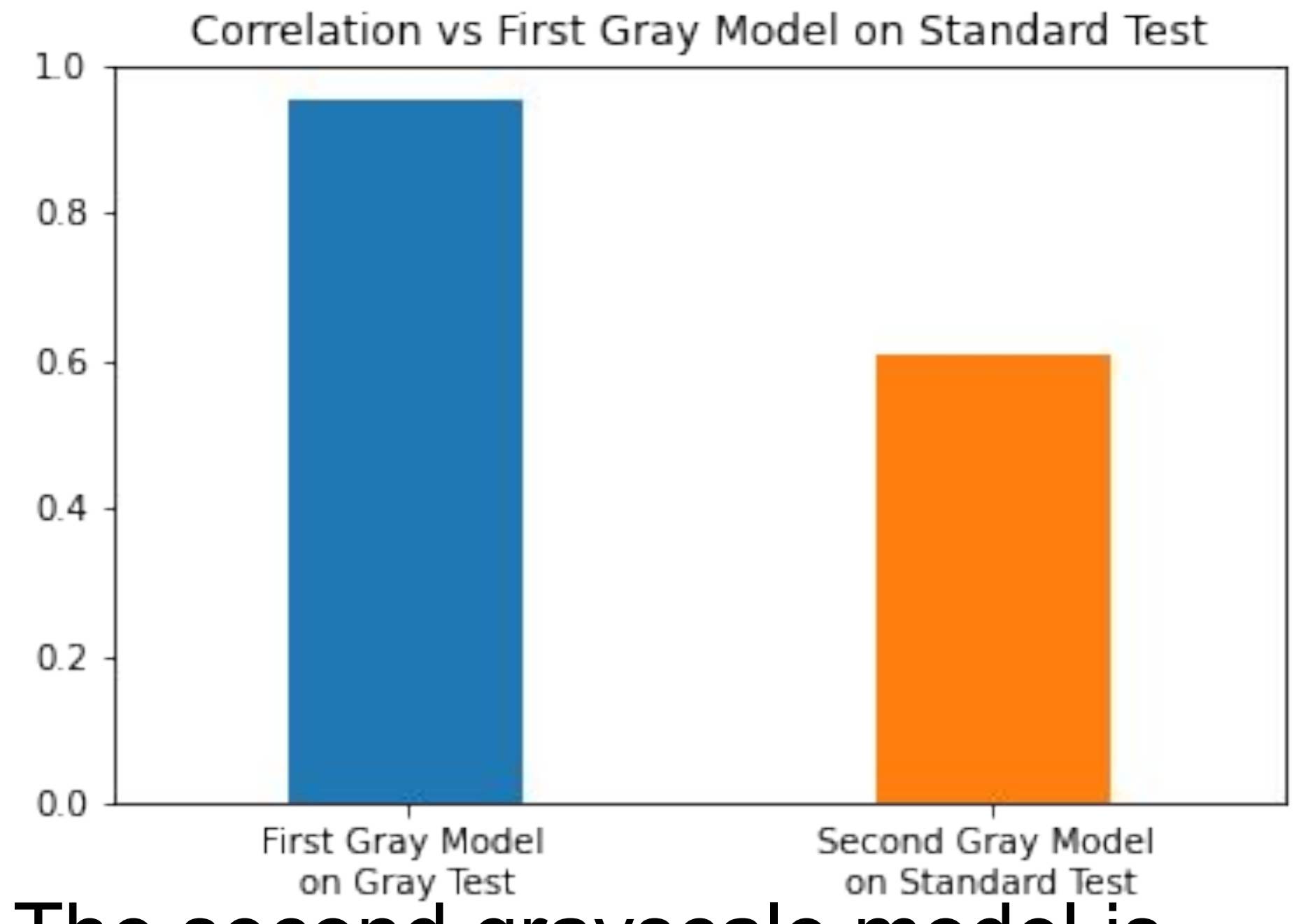
Standard Model



- The grayscale test relatively correlated to the standard test
- The 90-rotation test is not correlated
- Test the correlation again with augmented models

Correlation (Augmented)

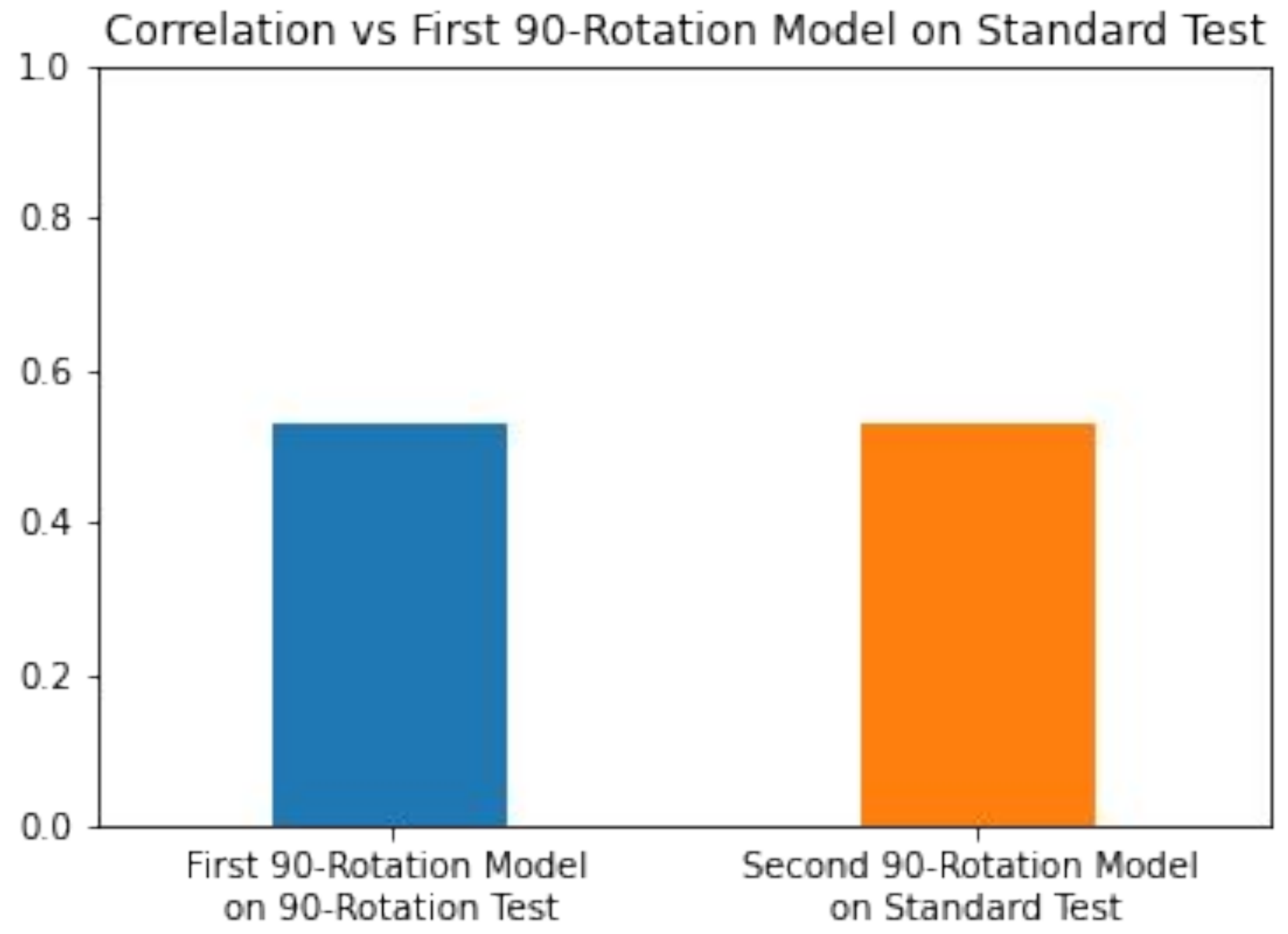
Grayscale Models



- The second grayscale model is trained as a benchmark
- The representations from two testsets are similar
- Shift focus to the 90-rotation testset

Correlation (Augmented)

90-Rotation Models

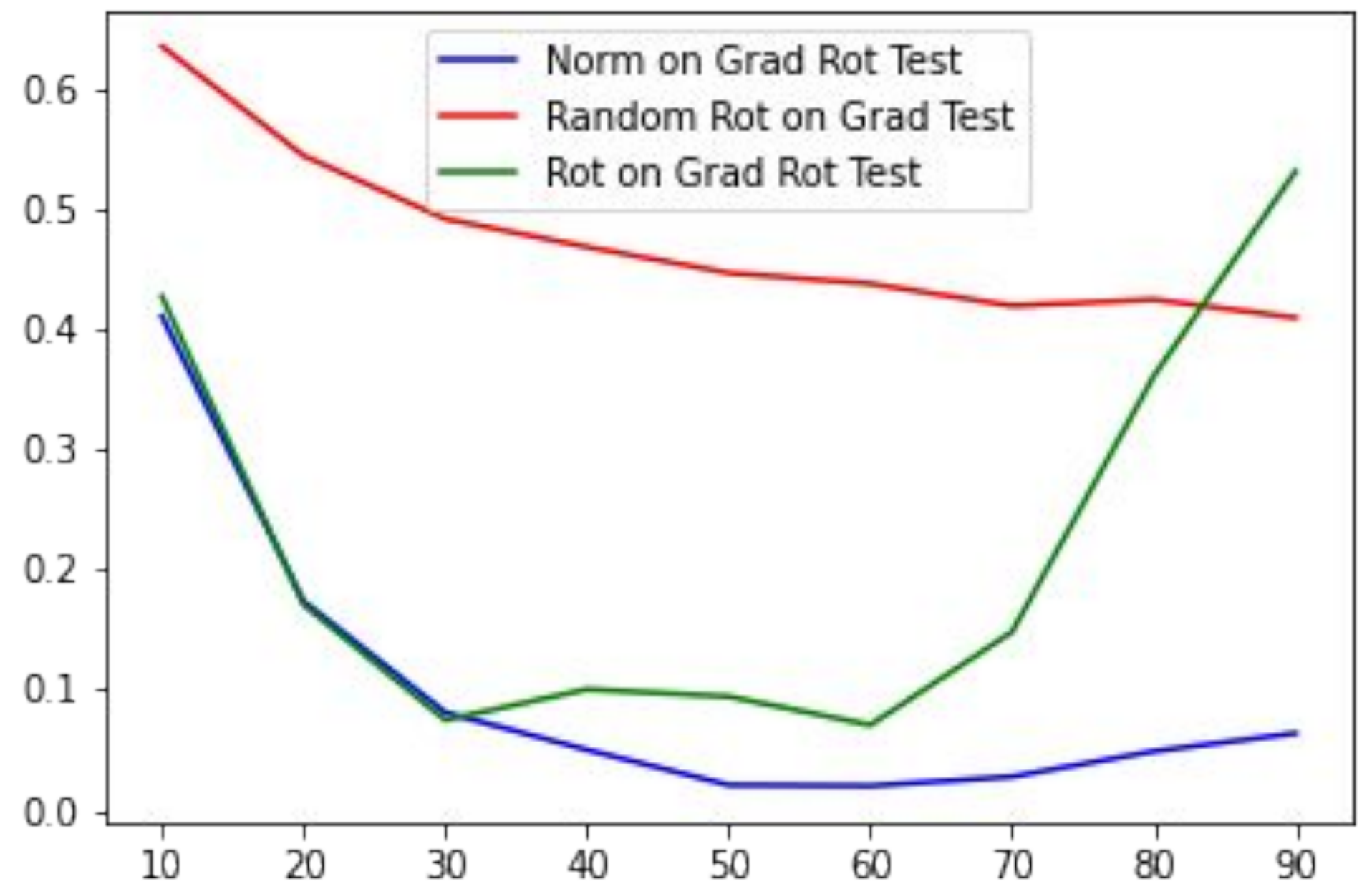


- The augmented test is more correlated than the standard model
- The correlation is as random as the correlation between 2 different models
- We attempt to leverage this using random rotation in training

Correlation (Intermediate)

Random Rotation Model

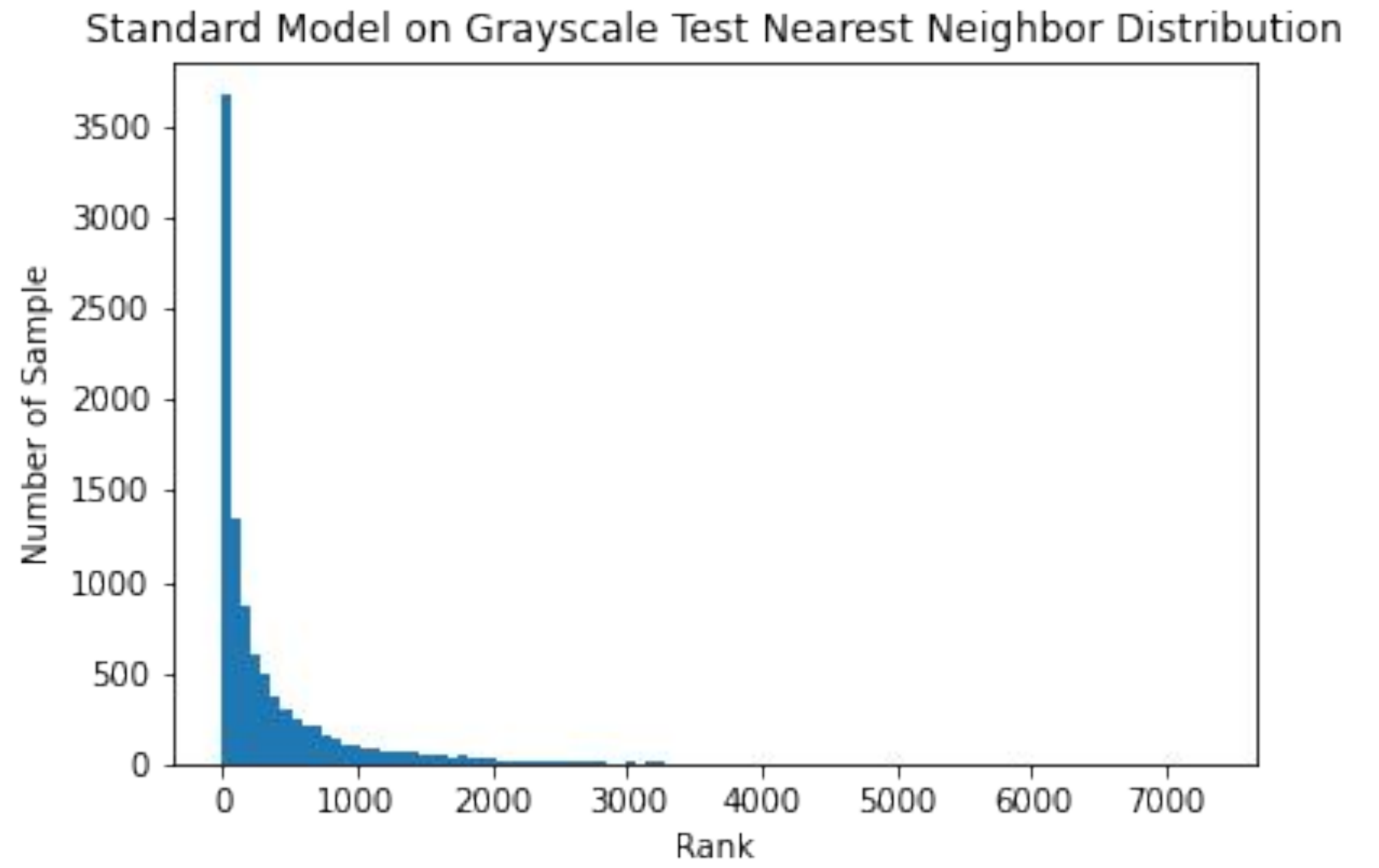
- For each image, we transform it with a random rotation (1-90)
- Unlike other models, we don't keep the original images

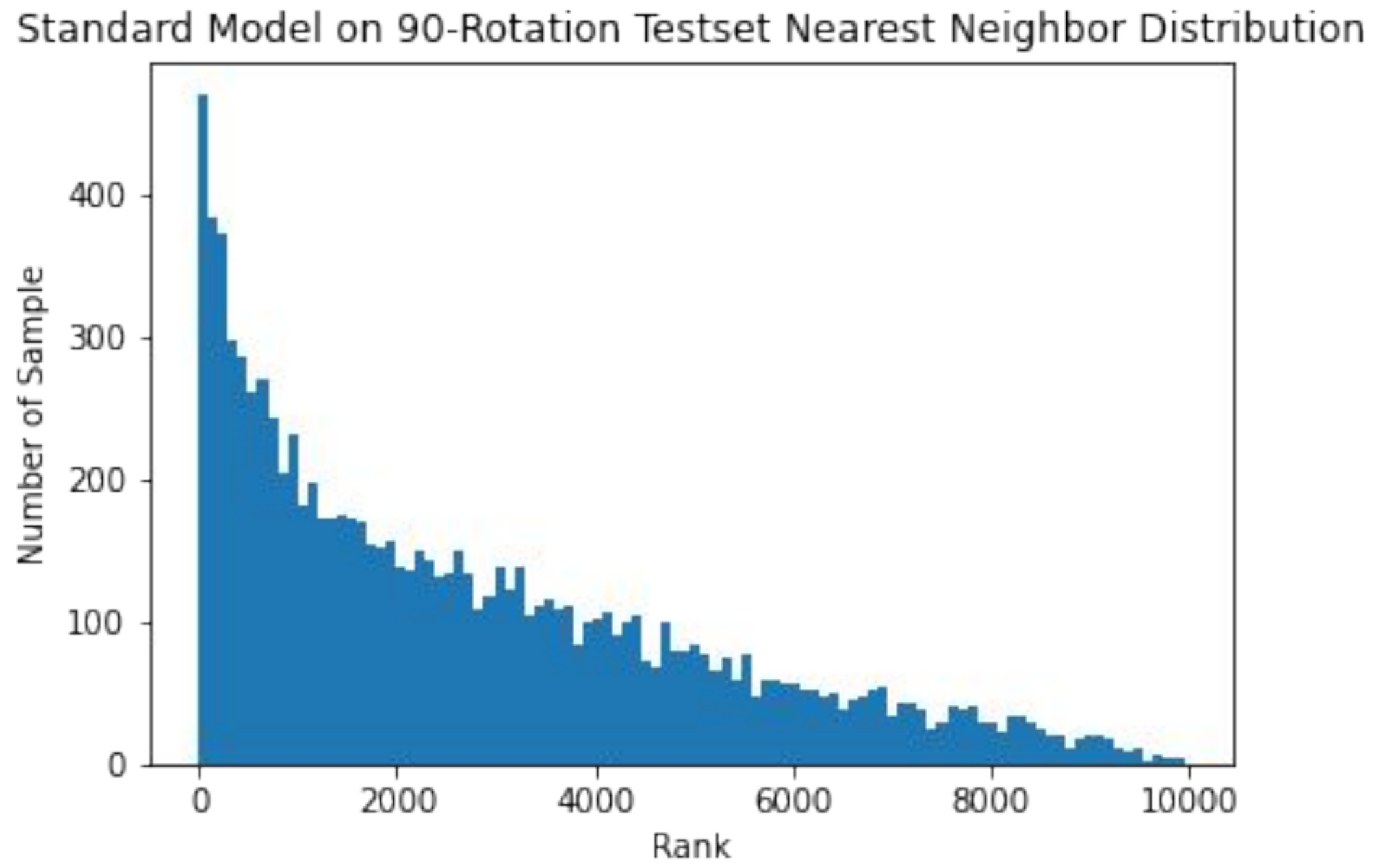


- Random rotation doesn't improve the correlation for the 90-rotation testset

Nearest Neighbor (Std)

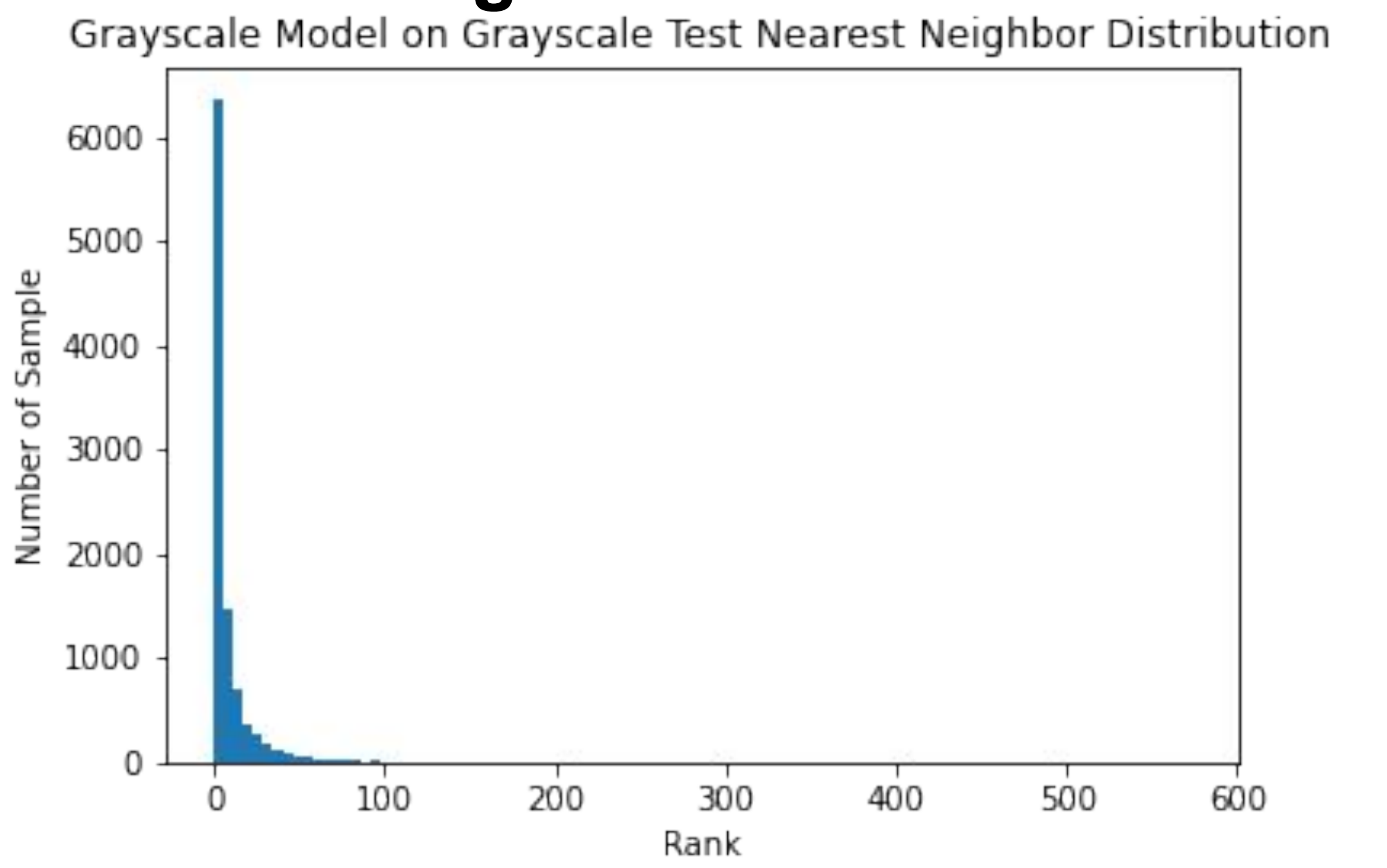
Standard Model

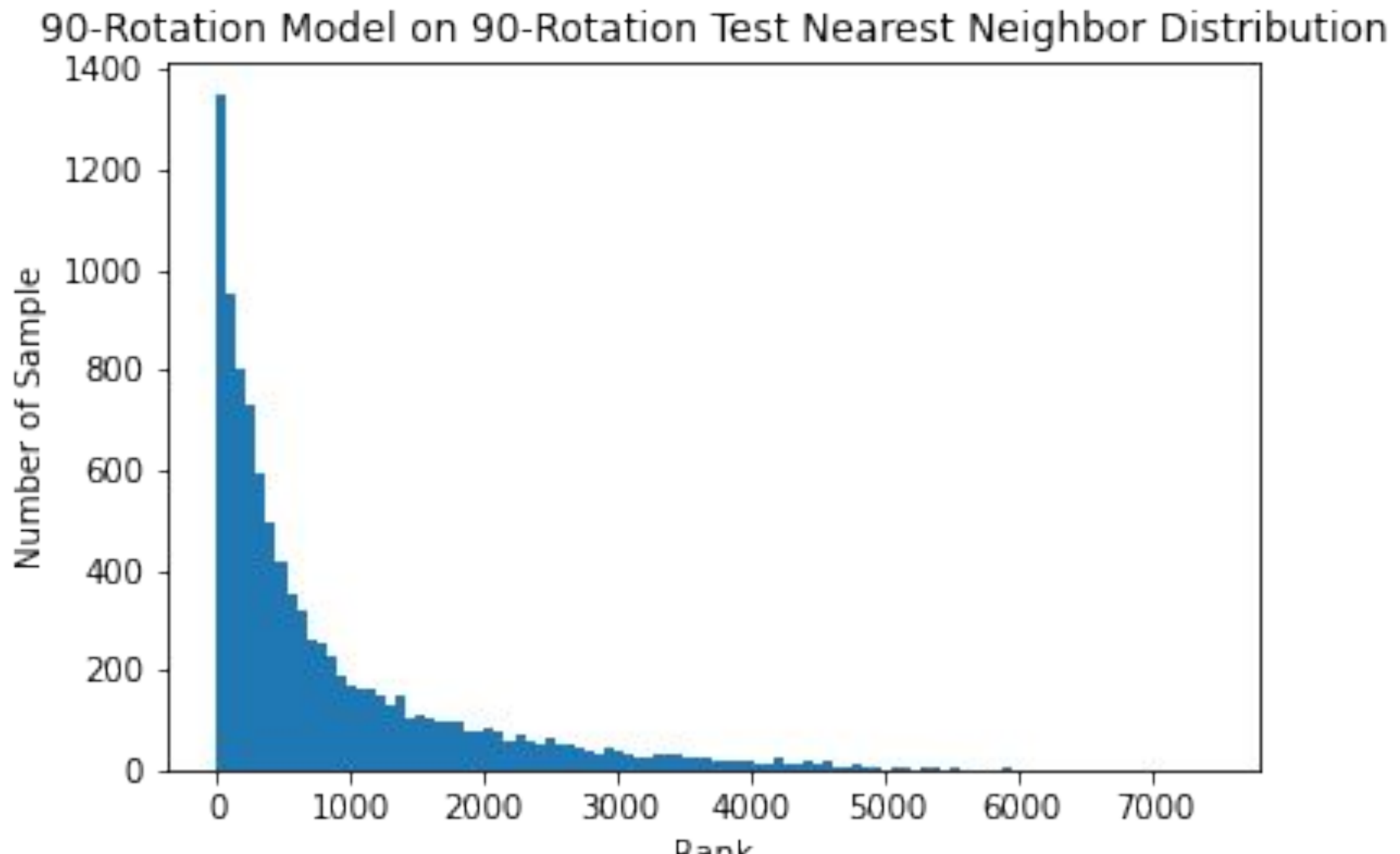




Correlation (Aug)

Augmented Models





- There is a shift in trend from standard to augmented: taller head, smaller tail