Comparison of mean downstream classification accuracy across augmentations