ABSTRACT

Can we create a high quality visual representation without any real-world images? This paper explores this possibility by training contrastive learning models on computer generated random datasets. We first use computer graphics techniques to generate a random dataset from scratch, then train a contrastive learning model on this dataset. We evaluate the representation by training a linear classifier on top with the Imagenet dataset. Classifying on 50 classes, the synthetic dataset achieves 44.64% top-1 accuracy and 72.40% top-5 accuracy, compared with 48.56% and 76.64% top-1, top-5 accuracy for the real ImageNet dataset. As a result, strong visual representations can be created completely from scratch. Additionally, we notice that certain properties of a dataset lead to better contrastive learning performance. Future work can be done to create more efficient datasets for contrastive learning, such as screening and modifying existing datasets or even creating synthetic ones.