

Ani-GIFs: A Benchmark Dataset for Domain Generalization of Action Recognition from GIFs

Shoumik Majumdar¹, Shubhangi Jain¹, Isidora Chara Tourni², Arsenii Mustafin¹, Diala Lteif¹, Stan Sclaroff¹, Kate Saenko¹, and Sarah Bargal¹

¹Boston University

²Boston University College and Graduate School of Arts and Sciences

June 25, 2021

Abstract

Deep learning models perform remarkably well for the same task under the assumption that data is always coming from the same distribution. However, this is generally violated in practice, mainly due to the differences in the data acquisition techniques and the lack of information about the underlying source of new data. Domain Generalization targets the ability to generalize to test data of an unseen domain; while this problem is well-studied for images, such studies are significantly lacking in spatiotemporal visual content – videos and GIFs. This is due to (1) the challenging nature of misalignment of temporal features and the varying appearance/motion of actors and actions in different domains, and (2) spatiotemporal datasets being laborious to collect and annotate for multiple domains. We collect and present the first synthetic video dataset of Animated GIFs for domain generalization, *Ani-GIFs*, that is used to study domain gap of videos *vs.* GIFs, and animated *vs.* real GIFs, for the task of action recognition. We provide a training and testing setting for *Ani-GIFs*, and extend two domain generalization baseline approaches, based on data augmentation and explainability, to the spatiotemporal domain to catalyze research in this direction.

Hosted file

Ani_GIFs__WileyNJD_ (2).pdf available at <https://authorea.com/users/422056/articles/527832-ani-gifs-a-benchmark-dataset-for-domain-generalization-of-action-recognition-from-gifs>