Disentangling Visual Features for Novel Compositional Image Generation

Team Members:
- Ian Deal (Lead)
- Javier Pozo Miranda
- Sree Bhattacharyya
- Dev pratap Singh

Abstract

The project investigates the fundamental mathematical and practical limitations of modern generative models for image synthesis, focusing on their ability to generate compositional images –combinations of multiple features or objects. Current approaches face difficulties in generating novel combinations of learned visual features, often falling back to expressing existing patterns rather than producing unique outputs. Our goal is to establish if the disentanglement of latent representations enhances a model's generative capability for novel feature combinations not found in training data. Disentanglement refers to the process that separates different factors of variation present in the data, such that each factor can be manipulated independently without affecting the others. Secondly, we will explore what inductive biases are required to encourage disentanglement and guide the model in generating the new composition. Through the execution of a series of experiments grounded in established findings within the disentanglement literature, we aim to quantify the extent to which disentangled representations affect the capacity of generative models to produce novel and compositional outputs.