dSprites VSA

1 Introduction

The goal of this work is to extract the features of the image and represent them as multidimensional vectors, which, when summed up, will allow you to reconstruct the original image. Having started with a single object in the scene, we will try to move to a set of several objects.

2 Dataset

The task consists of three consecutive steps for each of which a different dataset was created.

2.1 Disentanglement Dataset

The work is based on the dataset dsprites Figure 1 This dataset contains paired images from the dataset dSprites, which differ in one of the following features Table 1. An example of pairwise images in a dataset can be seen in the Figure 2

Table 1: List of features

Type	Description
Shape	square, ellipse, heart
Scale	6 values linearly spaced in [0.5, 1]
Orientation	40 values in [0, 2 pi]
Position X	32 values in [0, 1]
Position Y	32 values in [0, 1]

2.2 Multi-dsprites dataset

Dataset consists of 2 to 5 non-overlapping figures from dSprites dataset. An example of such images on Figure 3

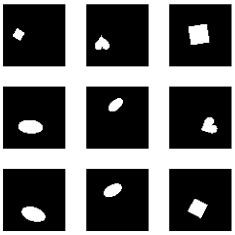

Figure 1: Example of images from dSprites dataset

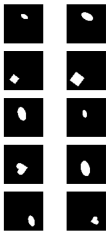

Figure 2: Example of paired images in Disentanglement Dataset

Figure 3: Example of collected scenes in the multidsprites dataset