

UNSUPERVISED LEARNING WITH SYNTHETIC DATA AND ATTENTION MASKS

ABSTRACT

Embodiments of the present disclosure relate to unsupervised learning with synthetic data and attention masks. According to example embodiments of the present disclosure, a plurality of sample images are generated by providing a plurality of text prompts into a trained generative model, respectively. For a sample image of the plurality of sample images, at least one attention map is obtained from a generative model, the at least one attention map being determined by the generative model for generating the sample image, an attention map indicating visual elements of an object within the sample image. Training of a target model is performed according to unsupervised learning at least based on the plurality of sample images and attention maps for the plurality of sample images, the target model being configured to perform an image processing task.

FIG. 2