Introduction

Our Contributions

- We propose a framework to add products as advertisement in an arbitrary scene image. The framework consists of two parts: bounding box predictor and advertisement generator.
- We propose local-global discriminators in order to generate the products more realistic in the whole scene and detail information such as characters and logo.
- To demonstrate our method more convincing, we not only train on existing public dataset, but also make a dataset with some specific brands such as RIO cocktail, Coca Cola, HP laptop, Audi A6 and so on in a wide range products.

Methods

Bounding Box Predictor

- Firstly, we extract the semantic label map of the given scene image. We use Upernet to do semantic segmentation.
- Secondly, our bounding box predictor takes the semantic label map as input and infers a proper bounding box, containing the top left corner and bottom right corner. It provides both proper location and size information of the product in the scene image.

Advertisement Generator

- Given the scene image domain X, bounding box domain B, our goal is to synthesize the product image Y. Then, the problem of adding advertisement aims to learn the mapping \( G: X \times B \rightarrow Y \).
- Now, let \( M \) denotes the mask domain, so mask generator \( G_M \) can be defined as a mapping \( G_M: X \times B \rightarrow M \) and product generator can defined as a mapping \( G_Y: X \times M \rightarrow Y \).

Results

Products Synthesized by Our Framework

Our Method VS Patch GAN

Conclusions

In this paper, we have proposed a novel framework for context-aware natural integration. A bounding box predictor and an adversarial network are used to make sure where and what the object should be.

Experiments on datasets, including public dataset and our own dataset, confirm that our method obtains better performance, especially Chinese logo objects.

Reference