

Diverse Image-to-Image Translation via Disentangled Representations

Hsin-Ying Lee¹, Hung-Yu Tseng¹, Jia-Bin Huang², Maneesh Singh³,
Ming-Hsuan Yang^{1,4}

¹University of California, Merced ²Virginia Tech ³Verisk Analytics ⁴Google Cloud

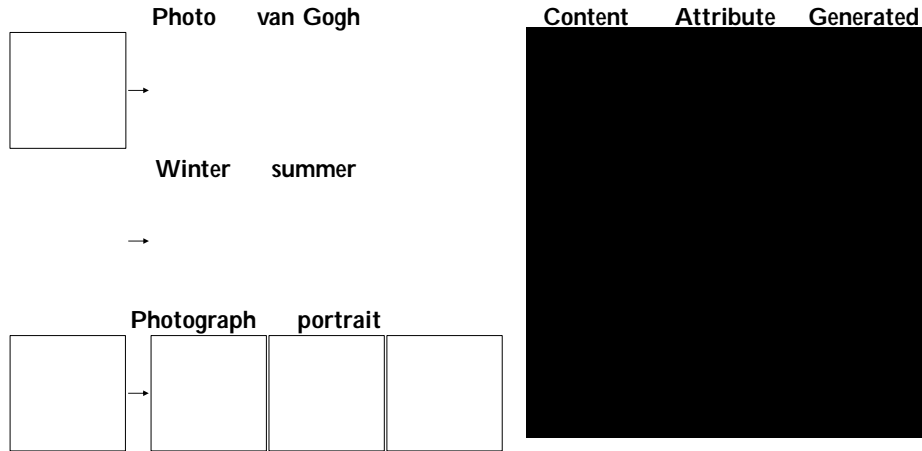


Fig. 1: Unpaired diverse image-to-image translation. (Left) Our model performs diverse translation between two collections of images without aligned training pairs. (Right) Example-guided translation.

Abstract. Image-to-image translation aims to learn the mapping between two visual domains. There are two main challenges for many applications: 1) the lack of aligned training pairs and 2) multiple possible outputs from a single input image. In this work, we present an approach based on disentangled representation for producing diverse outputs without paired training images. To achieve diversity, we propose to embed images onto two spaces: a domain-invariant content space capturing shared information across domains and a domain-specific attribute space. Using the disentangled features as inputs greatly reduces mode collapse. To handle unpaired training data, we introduce a novel cross-cycle consistency loss. Qualitative results show that our model can generate diverse and realistic images on a wide range of tasks. We validate the effectiveness of our approach through extensive evaluation.

equal contribution

