# Cones2

- 【题目】Cones 2: Customizable Image Synthesis with Multiple Subjects
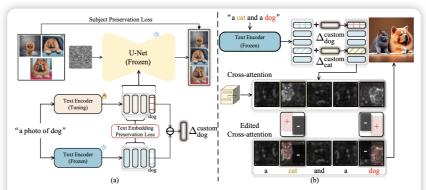- 【链接】https://arxiv.org/abs/2305.19327
- NIPS 2023



Figure 2: **Illustration of the proposed approach.** (a) We first learn a residual token embedding (*e.g.,* $\Delta_{\text{dog}}^{\text{custom}}$) on top of the base embedding to register a user-specified subject, which allows composing various subjects arbitrarily without further model tuning. (b) Given a layout as the spatial guidance, we then arrange the subjects by rectifying the activations in cross-attention maps, which enables the control of the location of each subject and reduces the interference between them.

- 【主要贡献、核心方法】

　　我们首先学习一个残差的词嵌入（embedding），在基础嵌入的基础上注册一个用户指定的主题，这允许通过残差的 embedding 来任意组合多个主题，而且不需要对模型进行微调；然后再通过布局作为空间引导，来控制每个主题的位置，这样可以减少主题之间的干扰。

　　主要就是学习一个残差的词嵌入，来注册一个用户指定的主题，然后再通过布局作为空间引导来控制每个主题的位置。



Figure 1: **Customizable image generation** with the subjects listed on the left. **Cones 2** is highlighted from three aspects. (1) Using a simple yet effective representation to register a subject, we can compose various subjects arbitrarily *without any model tuning*. (2) Employing spatial layout, which is very easy to obtain in practice, as a guidance, we can *control the specific location* of each subject and meanwhile *alleviate the interference* across subjects. (3) Our method achieves appealing performance even under some *challenging settings*, such as customizing the synthesis with six or more subjects and exchanging the sunglasses on the two dogs.

- 【期刊、会议】
- 【数据集、评价指标、任务】多主题定制生成，主题位置控制
- 【算力】1 A100
- 【代码】https://github.com/ali-vilab/Cones-V2

---

| 2024-10-22 | 2024-10-21 |
| --- | --- |
| Face2Diffusion | CAFE |

Archive  RSS feed  GitHub  Email  QR Code

Made with Montaigne and bigmission 🇺🇦