

深度学习第三次作业实验报告

2023 年 1 月 1 日

1 GNN

1.1 Task A, B

利用 pyg 中定义的 GINConv, MLP 等基本层实现了 GIN。

GCN, GAT, Node2Vec 三种模型以及实现的 GIN 模型的训练及测试曲线如下：

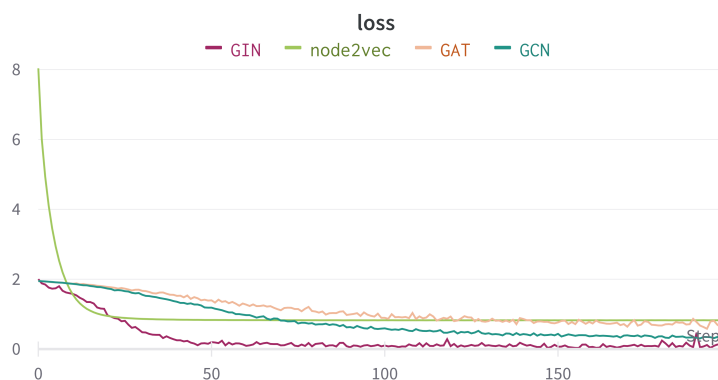


图 1: GNN 模型训练 loss 曲线

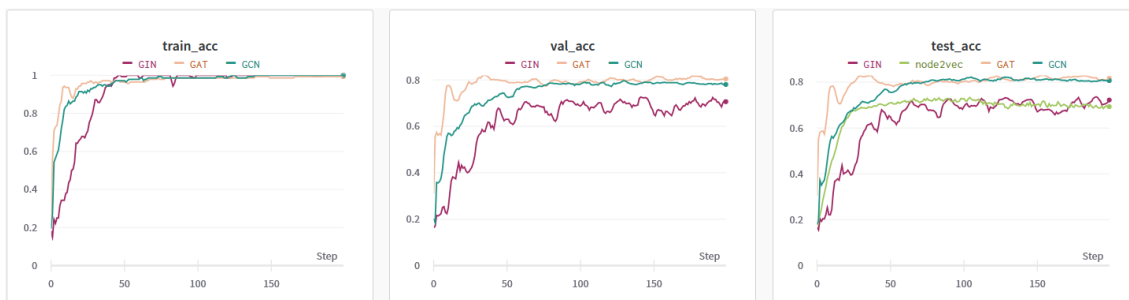


图 2: GNN 模型准确率曲线

2 Diffusion model

2.1 理论推导

证明前向过程 x_t 收敛到标准高斯分布：

由前向过程的真实分布可得 $x_t = \sqrt{1 - \beta_t}x_{t-1} + \sqrt{\beta_t}I = \sqrt{1 - \beta_t}x_{t-1} + \sqrt{c}I$

由高斯分布的可加性，即

$$\forall t, \epsilon_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, I), \quad a\epsilon_{t-1} + b\epsilon_{t-2} \stackrel{d}{=} \sqrt{a^2 + b^2}\bar{\epsilon}_{t-2} \sim \mathcal{N}(0, (a^2 + b^2)I)$$

令 $\alpha_t = 1 - \beta_t$ ，扩散核 $\bar{\alpha}_t = \alpha_t\alpha_{t-1}\dots\alpha_0$ ，由于 $\alpha_t < 1$ ，当 $t \rightarrow \infty$ 时有 $\bar{\alpha}_t \rightarrow 0$ 。

$$\begin{aligned} x_t &= \sqrt{\alpha_t}x_{t-1} + \sqrt{1 - \alpha_t}\epsilon_{t-1}, & \epsilon_{t-1} &\sim \mathcal{N}(0, I) \\ &= \sqrt{\alpha_t}(\sqrt{\alpha_{t-1}}x_{t-2} + \sqrt{1 - \alpha_{t-1}}\epsilon_{t-2}) + \sqrt{1 - \alpha_t}\epsilon_{t-1}, & \epsilon_{t-1}, \epsilon_{t-2} &\sim \mathcal{N}(0, I) \\ &= \sqrt{\alpha_t\alpha_{t-1}}x_{t-2} + \sqrt{1 - \alpha_t\alpha_{t-1}}\bar{\epsilon}_{t-2}, & \bar{\epsilon}_{t-2} &\sim \mathcal{N}(0, I) \\ &= \dots \\ &= \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1 - \bar{\alpha}_t}\epsilon, & \epsilon &\sim \mathcal{N}(0, I) \end{aligned}$$

故 $t \rightarrow \infty$ 时 $x_t \rightarrow \epsilon \sim \mathcal{N}(0, I)$ 。

证明后验分布是高斯分布并求其均值、方差：

$$q(x_{t-1}|x_t, x_0) = \frac{q(x_{t-1}, x_t, x_0)}{q(x_t, x_0)} = \frac{q(x_t|x_{t-1}, x_0)q(x_{t-1}|x_0)q(x_0)}{q(x_t|x_0)q(x_0)} = \frac{q(x_t|x_{t-1})q(x_{t-1}|x_0)}{q(x_t|x_0)}$$

前向过程中 $q(x_t|x_0) = \mathcal{N}(\sqrt{\bar{\alpha}_t}x_0, (1 - \bar{\alpha}_t)I)$ ， $q(x_t|x_{t-1}) = \mathcal{N}(\sqrt{\alpha_t}x_{t-1}, (1 - \alpha_t)I)$ ，故后验分布密度满足

$$\begin{aligned} q(x_{t-1}|x_t, x_0) &= \frac{q(x_t|x_{t-1})q(x_{t-1}|x_0)}{q(x_t|x_0)} \\ &\propto \exp \left\{ -\frac{(x_t - \sqrt{\alpha_t}x_{t-1})^2}{2(1 - \alpha_t)} - \frac{(x_{t-1} - \sqrt{\bar{\alpha}_{t-1}}x_0)^2}{2(1 - \bar{\alpha}_{t-1})} + \frac{(x_t - \sqrt{\bar{\alpha}_t}x_0)^2}{2(1 - \bar{\alpha}_t)} \right\} \\ &= \exp \left\{ -\frac{1}{2} \left[\left(\frac{\alpha_t}{1 - \alpha_t} + \frac{1}{1 - \bar{\alpha}_{t-1}} \right) x_{t-1}^2 + \left(-2\frac{\sqrt{\alpha_t}x_t}{1 - \alpha_t} - 2\frac{\sqrt{\bar{\alpha}_{t-1}}x_0}{1 - \bar{\alpha}_{t-1}} \right) x_{t-1} + C(x_t, x_0) \right] \right\} \end{aligned}$$

其中 $C(x_t, x_0)$ 不包含 x_{t-1} ，归一化后可知 $q(x_{t-1}|x_t, x_0)$ 满足高斯分布，其方差 $\tilde{\beta}_t$ 和均值 $\tilde{\mu}_t(x_t, x_0)$ 分别为

$$\begin{aligned} \tilde{\beta}_t &= 1 / \left(\frac{\alpha_t}{1 - \alpha_t} + \frac{1}{1 - \bar{\alpha}_{t-1}} \right) = \frac{(1 - \alpha_t)(1 - \bar{\alpha}_{t-1})}{1 - \alpha_t\bar{\alpha}_{t-1}} = \frac{(1 - \alpha_t)(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} \\ \tilde{\mu}_t(x_t, x_0) &= \frac{-2\frac{\sqrt{\alpha_t}x_t}{1 - \alpha_t} - 2\frac{\sqrt{\bar{\alpha}_{t-1}}x_0}{1 - \bar{\alpha}_{t-1}}}{-2\tilde{\beta}_t} = \frac{\sqrt{\bar{\alpha}_{t-1}}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t}x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1 - \alpha_t)}{1 - \bar{\alpha}_t}x_0 \end{aligned}$$

2.2 模型实现

2.2.1 DDPM

如图3可以看出，在采样步数较少时，DDPM 生成的图像带有较多噪声。个人认为步数在 256 之后生成的图像效果最好。

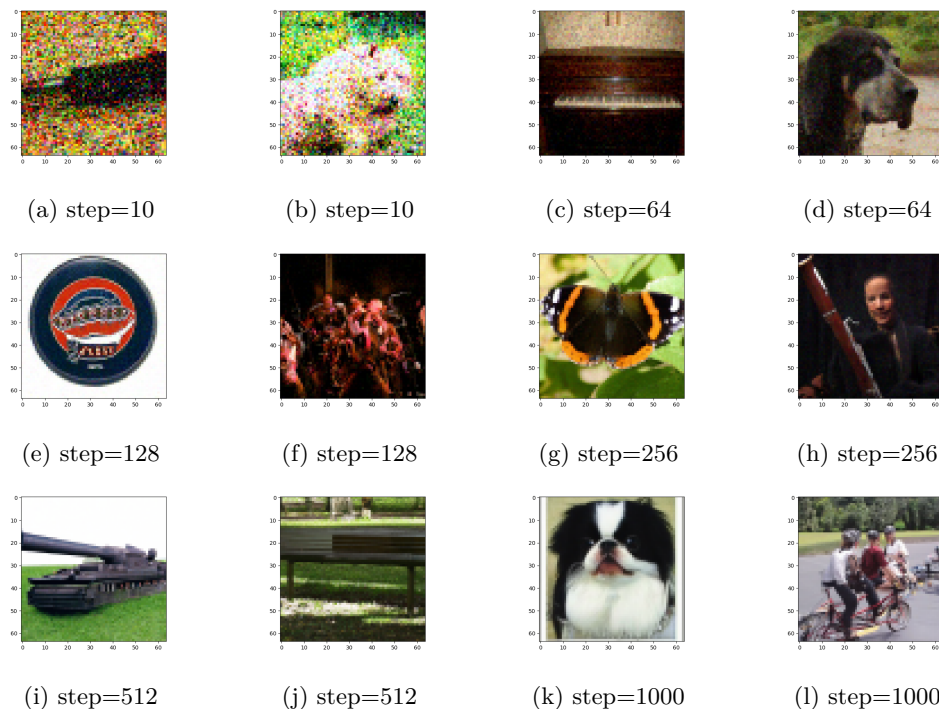


图 3: DDPM 不同采样步数对比

2.2.2 DDIM

实现方式: 根据 [1], 在父类 `GaussianDiffusion` 中新增了计算 `ddim_mean_variance` 的函数, 并在子类 `SpacedDiffusion` 中将其 `wrap` 计算 `rescaled timestep` 完成。

如图4可以看出, `step=10` 时 DDIM 生成的图像带噪情况明显好于 DDPM, `step=64` 时生成的图像效果已经基本无噪声, 而 DDPM 还会带有一些模糊的底噪; 但是, DDIM 生成的图像在生成纯色背景时偶尔会出现噪声。

2.2.3 CGDM

设置 `class_index=1`, 要求生成类别为金鱼的图片。改变 `classifier_scale`, 即控制 `cond_fn` 中类别贡献的梯度。可以看出,

- `scale` 较小时如图5, 生成的图像较为多样, 但类别不一定相符, 有 1-2 张图仅仅是金色占比大, 但不能辨认出是鱼;
- `scale` 适中时如图6, 保持了多样性, 同时基本可以辨认出是金鱼;
- `scale` 过大时如图7, 保证类别为金鱼, 但是多样性大大下降, 多数图片非常相似。

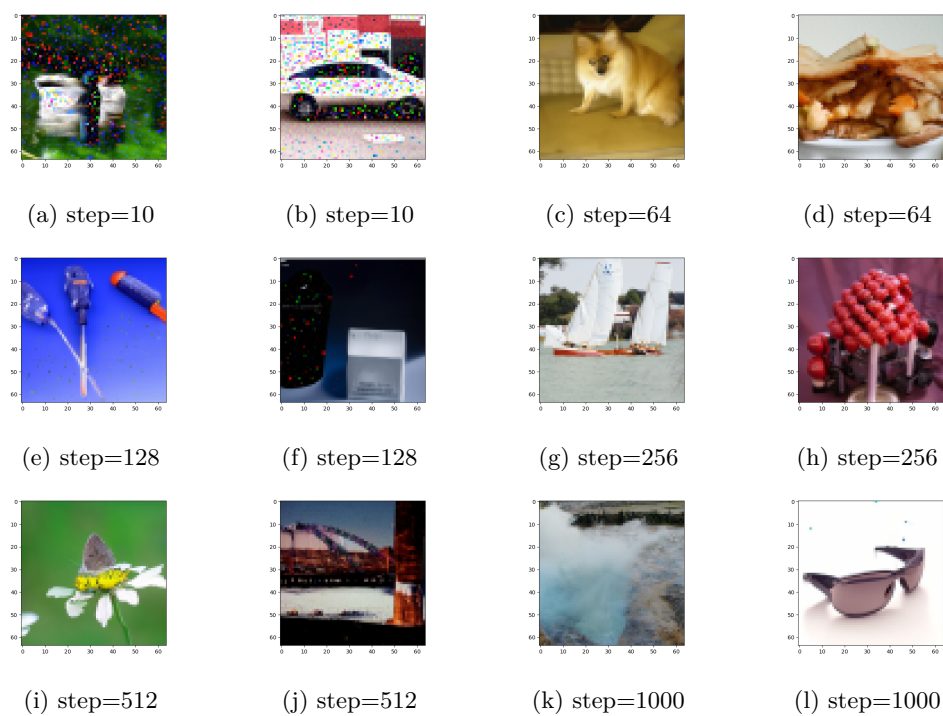


图 4: DDIM 不同采样步数对比

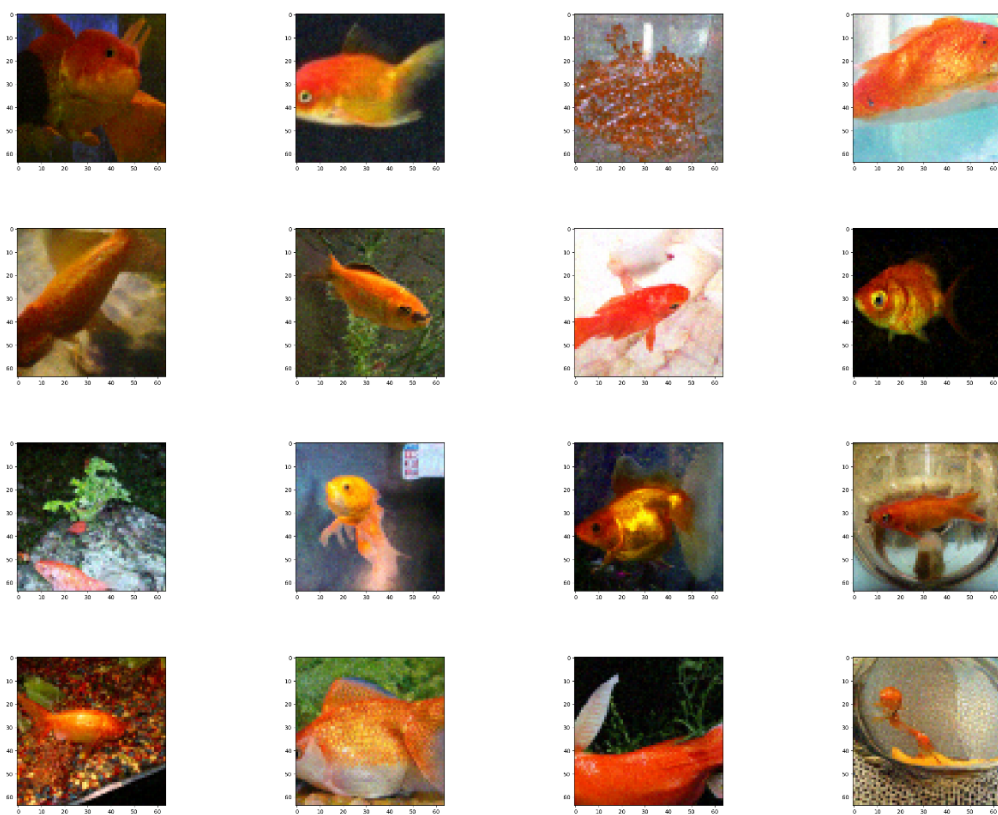


图 5: Classifier Guided Diffusion Model (CGDM) classifier scale=0.1

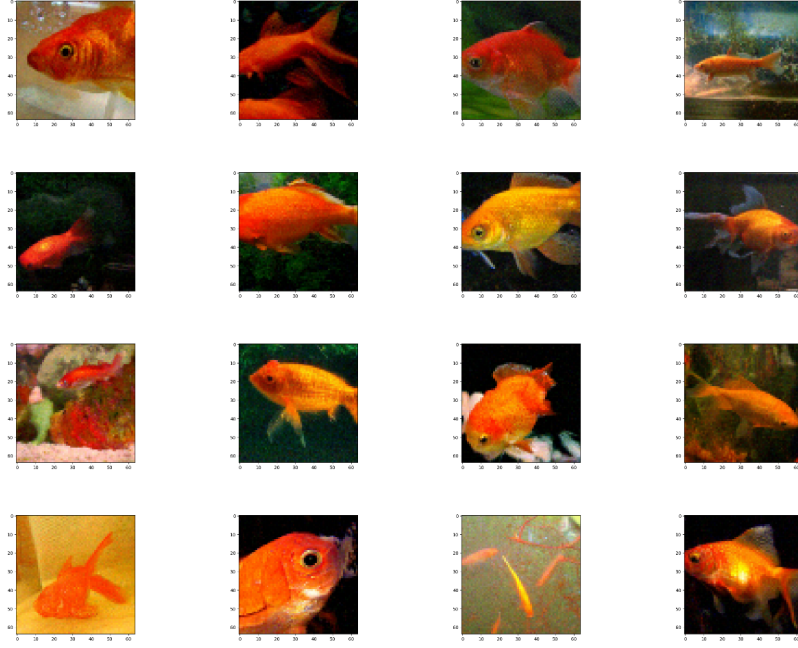


图 6: CGDM classifier scale=1

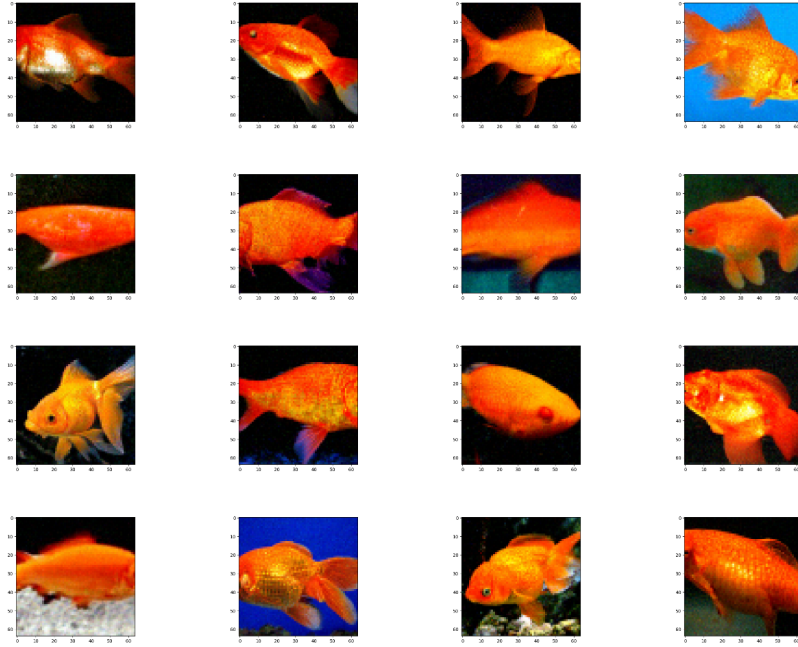


图 7: CGDM classifier scale=10

2.3 Text-to-image failure case study

使用 Stable Diffusion API, 找到五个 failed text prompts 如图8:

1. 位置错误, 蓝色的背包应该在白猫前面而不是旁边;
2. 数量和物品错误, 出现了两只绿鹦鹉, 没有棕色的碗;
3. 特征错误混合, 红色小轿车上长了羊毛, 车内有两个羊头;
4. 常识错误, 卡塔尔的圣诞节在夏季, 不应该穿厚衣服;
5. 数量错误, 6 个苹果图中只出现了至多 4 个。



(a) a blue backpack in front of (b) a green parrot and a brown (c) a red car and a white sheep
a white cat bowl



(d) in the yard mom and dad (e) six apples
celebrating Christmas in Qatar

图 8: 5 Stable Diffusion Failure Cases

参考文献

- [1] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *CoRR*, abs/2010.02502, 2020.