A Technical Details

Textual Inversion in Stable Diffusion. In Sec. 4.2, we optimize a token embedding for data-scarce quantization. Specifically, this technique is based on the Textual Inversion [18] where a token embedding is optimized for accurate recreation of an object from 3-5 images. To optimize the token embedding, we freeze all other parts of the Stable Diffusion including all the parameters from the encoder, decoder, text encoder, and the denoising U-Net. The training objective is the same as the training of Stable Diffusion, *i.e.* diffuse the real images latent into noisy latent and learn how to denoise it except we only update the text embedding (*cf.* Eq. (11)). Note that the original textual inversion is designed to find the token of the exact *object* while we have multiple images from multiple classes. Hence, our objective is to find some token embedding that can universally describe the traits of the ImageNet data, aiming to learn distinct distribution characteristics.
Implementation of Token Embedding Learning. To optimize $\{\texttt{S}\}$, we optimize it for 50k iterations with a batch size of 32 and use Adam optimizer with a constant learning rate of $5e-4$, following the convention of implementation [18]. Note that for each batch, we construct 32 different prompts since each image may have a unique label name.

B Prompt Template

We use the following template to generate the prompt:

1. `photo of a {C}.`
2. `rendering of a {C}.`
3. `cropped photo of the {C}.`
4. `the photo of a {C}.`
5. `photo of a clean {C}.`
6. `photo of a dirty {C}.`
7. `dark photo of the {C}.`
8. `photo of my {C}.`
9. `photo of the cool {C}.`
10. `close-up photo of a {C}.`
11. `bright photo of the {C}.`
12. `cropped photo of a {C}.`
13. `photo of the {C}.`
14. `good photo of the {C}.`
15. `photo of one {C}.`
16. `close-up photo of the {C}.`
17. `rendition of the {C}.`
18. `photo of the clean {C}.`
19. `rendition of a {C}.`
20. `photo of a nice {C}.`
21. `good photo of a {C}.`

22. `photo of the nice {C}.`
23. `photo of the small {C}.`
24. `photo of the weird {C}.`
25. `photo of the large {C}.`
26. `photo of a cool {C}.`
27. `photo of a small {C}.`