|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **组织** | **架构** | **数据集** | **学习步长** | **分辨率** | **其他** |
| V1.1 | CompVis | CLIP text\_encoder +LDM | laion2B-en  laion-high-resolution | 237000  194000 | 256\*256  512\*512 | 从零训练 |
| V1.2 | laion-improved-aesthetics | 515,000 | 512\*512 | 基于V1.1权重训练 |
| V1.3 | 195000 | 基于V1.2权重训练  10% drop text-conditioning |
| V1.4 | 195000 |
| V1.5 | runwayml |  | 595000 | 基于V1.2权重训练  10% drop text-conditioning  EMA/剪枝 |
| V2-base | Stability-AI | OpenCLIP text\_encoder(重训) +LDM | laion5B(NSFW) | 550000  850000 | 256\*256  512\*512 | 从零训练 |
| V2 | 150000  140000 | 512\*512  768\*768 | 基于V2-base权重训练 |
| V2.1 | 55000(punsafe=0.1)  155000(punsafe=0.9) | 768\*768 | 基于V2权重训练 |
| V3 | DiT+ Flow Match | - | - | - | - |