# Supplementary Material for DISCO: DISCrete nOise for Conditional Control in Text-to-Image Diffusion Models

## 1 Proof

**Theorem 1.** *Let $n \geq 2$ denote the ambient dimension, and let $\theta = \angle(\boldsymbol{\mu}, \boldsymbol{\epsilon})$ be the angle between the mean vector $\boldsymbol{\mu}$ and a random vector $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{I})$. Then, $\upsilon = \cos^2\theta = \left(\frac{\langle \boldsymbol{\mu}, \boldsymbol{\epsilon} \rangle}{\|\boldsymbol{\mu}\| \cdot \|\boldsymbol{\epsilon}\|}\right)^2$ follows the non-central beta distribution:*

$$\upsilon \sim \textit{Beta}\left(\alpha,\ \beta;\ \lambda\right). \tag{1}$$

*where $\alpha = \frac{1}{2}$, $\beta = \frac{n-1}{2}$, and $\lambda = \|\boldsymbol{\mu}\|^2$. Let ${}_2F_1\left(a,\ b;\ c;\ z\right)$ be a Gaussian hypergeometric function. The expectation of $\cos^2\theta$ is given by*

$$\mathbb{E}[\upsilon] = \frac{\alpha + \lambda}{\alpha + \beta + \lambda} \cdot {}_2F_1\left(1, \alpha + 1; \alpha + \beta + 1; \frac{1}{1 + \lambda}\right) \cdot \frac{1}{\alpha + \beta}. \tag{2}$$

*Proof.* Let $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{I}_n)$ be a multivariate Gaussian random vector in $\mathbb{R}^n$, where $\boldsymbol{\mu} \in \mathbb{R}^n$ is a fixed non-zero mean vector, and $\boldsymbol{I}_n$ is the identity covariance matrix. We are interested in the squared cosine similarity between the random vector $\boldsymbol{\epsilon}$ and the mean direction $\boldsymbol{\mu}$, defined as:

$$\upsilon = \cos^2\theta = \left(\frac{\langle \boldsymbol{\mu}, \boldsymbol{\epsilon} \rangle}{\|\boldsymbol{\mu}\| \cdot \|\boldsymbol{\epsilon}\|}\right)^2. \tag{3}$$

We aim to show that $\upsilon$ follows a non-central Beta distribution and to derive an expression for its expectation. Without loss of generality, by leveraging the rotational invariance of Gaussian distributions, we can assume $\boldsymbol{\mu} = (\|\boldsymbol{\mu}\|, 0, \ldots, 0)^\top$. Under this coordinate system, we can write the random vector $\boldsymbol{\epsilon} = \boldsymbol{\mu} + \boldsymbol{z}$, where $\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$. Then:

$$\boldsymbol{\epsilon}_1 = \|\boldsymbol{\mu}\| + z_1, \quad \text{and} \quad \epsilon_i = z_i \text{ for } i \geq 2. \tag{4}$$

The inner product becomes:

$$\langle \boldsymbol{\mu}, \boldsymbol{\epsilon} \rangle = \|\boldsymbol{\mu}\|(\|\boldsymbol{\mu}\| + z_1), \tag{5}$$

and the squared norm:

$$\|\boldsymbol{\epsilon}\|^2 = (\|\boldsymbol{\mu}\| + z_1)^2 + \sum_{i=2}^{n} z_i^2. \tag{6}$$

Therefore:

$$\upsilon = \left(\frac{\|\boldsymbol{\mu}\|(\|\boldsymbol{\mu}\| + z_1)}{\|\boldsymbol{\mu}\| \cdot \sqrt{(\|\boldsymbol{\mu}\| + z_1)^2 + \sum_{i=2}^{n} z_i^2}}\right)^2 = \frac{(\|\boldsymbol{\mu}\| + z_1)^2}{(\|\boldsymbol{\mu}\| + z_1)^2 + \sum_{i=2}^{n} z_i^2}. \tag{7}$$

Define $X = (\|\boldsymbol{\mu}\| + z_1)^2$, and $Y = \sum_{i=2}^{n} z_i^2$. Then $X \sim \chi^2(1; \lambda)$, a non-central chi-squared distribution with 1 degree of freedom and non-centrality parameter $\lambda = \|\boldsymbol{\mu}\|^2$, and $Y \sim \chi^2(n-1)$, a central chi-squared distribution.

By a known result in the theory of ratio distributions, if $X \sim \chi^2(k_1; \lambda)$ and $Y \sim \chi^2(k_2)$ independently, then the ratio

$$\frac{X}{X + Y} \sim \text{Beta}\left(\frac{k_1}{2}, \frac{k_2}{2}; \lambda\right), \tag{8}$$

a non-central Beta distribution. Applying this to our case:

$$\upsilon = \frac{X}{X + Y} \sim \text{Beta}\left(\alpha = \frac{1}{2},\ \beta = \frac{n-1}{2};\ \lambda = \|\boldsymbol{\mu}\|^2\right). \tag{9}$$

The expectation of a non-central Beta random variable $\upsilon \sim \mathrm{Beta}(\alpha, \beta; \lambda)$ is given by the known analytical form involving a Gauss hypergeometric function (please refer to NIST Digital Library of Mathematical Functions ):

$$\mathbb{E}[\upsilon] = \frac{\alpha + \lambda}{\alpha + \beta + \lambda} \cdot {}_2F_1 \left(1,\ \alpha + 1;\ \alpha + \beta + 1;\ \frac{1}{1 + \lambda}\right) \cdot \frac{1}{\alpha + \beta}. \tag{10}$$

Substituting the parameters $\alpha = \frac{1}{2}$, $\beta = \frac{n-1}{2}$, and $\lambda = \|\boldsymbol{\mu}\|^2$, we obtain the expression claimed in the theorem.

$$\mathbb{E}[\cos^2 \theta] = \mathbb{E}[\upsilon] = \frac{\frac{1}{2} + \lambda}{\frac{n}{2} + \lambda} \cdot {}_2F_1 \left(1,\ \frac{3}{2};\ \frac{n+2}{2};\ \frac{1}{1 + \lambda}\right) \cdot \frac{1}{\frac{n}{2}}. \tag{11}$$

□

**Theorem 2.** *Under the setup described above, let $\{\upsilon_i\}_{i=1}^K$ be $K$ samples drawn from the non-central beta distribution Beta $(\alpha,\ \beta;\ \lambda)$. When $n$ is moderately large and $K$ is not too small, the typical maximum $\nu_{\mathcal{C}_n^K}$ is approximately the same:*

$$\nu_{\mathcal{C}_n^K} \approx \frac{2 \log K}{n} \tag{12}$$

*Proof.* Let $\boldsymbol{\mu} \in \mathbb{R}^n$ be a fixed mean vector, and let $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{I})$ be a sample from an isotropic Gaussian centered at $\boldsymbol{\mu}$. Denote the angle $\theta$ between $\boldsymbol{\epsilon}$ and $\boldsymbol{\mu}$, and define $\upsilon = \cos^2 \theta$. It is known that:

$$\upsilon = \left( \frac{\langle \boldsymbol{\epsilon}, \boldsymbol{\mu} \rangle}{\|\boldsymbol{\epsilon}\| \cdot \|\boldsymbol{\mu}\|} \right)^2. \tag{13}$$

The random variable $\upsilon \in [0, 1]$ describes the squared alignment between the Gaussian sample and the mean vector. Let $\upsilon_1, \ldots, \upsilon_K$ be $K$ independent samples of $\upsilon$, and let $\nu_{\mathcal{C}_n^K} := \max_{1 \le i \le K} \upsilon_i$ be their maximum. We aim to derive an approximation of $\nu_{\mathcal{C}_n^K}$ when $n$ is moderately large and $K$ is not too small.

**Step 1: Geometric Decomposition:** We begin by analyzing the structure of $\upsilon$. Write each sample $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{I})$ as:

$$\boldsymbol{\epsilon} = \boldsymbol{\mu} + \boldsymbol{z}, \quad \text{where } \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}). \tag{14}$$

Then the inner product is:

$$\langle \boldsymbol{\epsilon}, \boldsymbol{\mu} \rangle = \|\boldsymbol{\mu}\|^2 + \langle \boldsymbol{z}, \boldsymbol{\mu} \rangle, \tag{15}$$

and the norm is:

$$\|\boldsymbol{\epsilon}\|^2 = \|\boldsymbol{\mu}\|^2 + 2\langle \boldsymbol{z}, \boldsymbol{\mu} \rangle + \|\boldsymbol{z}\|^2. \tag{16}$$

Substituting into $\upsilon = \cos^2 \theta$, we have:

$$\upsilon = \left( \frac{\|\boldsymbol{\mu}\|^2 + \langle \boldsymbol{z}, \boldsymbol{\mu} \rangle}{\|\boldsymbol{\mu}\| \cdot \sqrt{\|\boldsymbol{\mu}\|^2 + 2\langle \boldsymbol{z}, \boldsymbol{\mu} \rangle + \|\boldsymbol{z}\|^2}} \right)^2. \tag{17}$$

Define $\lambda := \|\boldsymbol{\mu}\|^2$, and note:

- $\langle \boldsymbol{z}, \boldsymbol{\mu} \rangle \sim \mathcal{N}(0, \lambda)$, since $\boldsymbol{z} \sim \mathcal{N}(0, I)$ and $\boldsymbol{\mu}$ is fixed.
- $\|\boldsymbol{z}\|^2 \sim \chi^2(n)$, the chi-squared distribution with $n$ degrees of freedom.

This yields a complicated expression for $\upsilon$, but its key properties can be inferred through asymptotic concentration when $n$ is large.

**Step 2: Asymptotic Behavior in High Dimensions:** In high-dimensional spaces, random vectors drawn from $\mathcal{N}(\boldsymbol{\mu}, I)$ are almost orthogonal to any fixed vector not aligned with the coordinate axes. More precisely:

- The norm $\|\boldsymbol{z}\|^2 \sim \chi^2(n)$ concentrates around $n$ with high probability.
- The inner product $\langle \boldsymbol{z}, \boldsymbol{\mu} \rangle \sim \mathcal{N}(0, \lambda)$ has sub-Gaussian tails.
- Therefore, the numerator of $\upsilon$, which includes $\langle \boldsymbol{z}, \boldsymbol{\mu} \rangle$, varies on the order of $\sqrt{\lambda}$, while the denominator is dominated by $\sqrt{\lambda + n}$.

Combining these facts, we estimate:

$$\upsilon \approx \left( \frac{\lambda + \mathcal{O}(\sqrt{\lambda})}{\sqrt{\lambda} \cdot \sqrt{\lambda + n + \mathcal{O}(\sqrt{n})}} \right)^2 \approx \mathcal{O}\left(\frac{1}{n}\right), \tag{18}$$

suggesting that a typical value of $\upsilon$ is small, on the order of $\frac{1}{n}$, as dimension grows.

**Step 3: Distribution of the Maximum:** Let $\nu_{\mathcal{C}_n^K} = \max_{i=1}^K \upsilon_i$. To estimate the typical value of this maximum, we use a probabilistic argument. Assume $\upsilon$ has a cumulative distribution function $F(t) := \mathbb{P}(\upsilon \le t)$. Then:

$$\mathbb{P}\left(\nu_{\mathcal{C}_n^K} \le t\right) = F(t)^K, \tag{19}$$

and thus:

$$\mathbb{P}\left(\nu_{\mathcal{C}_n^K} \ge t\right) = 1 - F(t)^K. \tag{20}$$

To estimate the typical maximum, we set this probability to $0.5$ and solve for $t$:

$$F(t)^K = 1/2 \quad \Rightarrow \quad \log F(t) = -\frac{\log 2}{K}. \tag{21}$$

Suppose $\upsilon \sim \text{Beta}(\frac{1}{2}, \frac{n-1}{2}; \lambda)$, then for small $t$, the distribution behaves approximately like a central Beta distribution. The density is approximately:

$$f(t) \propto t^{\alpha-1}(1-t)^{\beta-1} \approx t^{-1/2} \quad \text{for small } t, \tag{22}$$

and thus the CDF behaves like $F(t) \approx c \cdot \sqrt{t}$, for some constant $c > 0$ depending weakly on $\lambda$. Then:

$$F(t)^K \approx (c \cdot \sqrt{t})^K = c^K \cdot t^{K/2}. \tag{23}$$

Set this equal to $1/2$:

$$c^K \cdot t^{K/2} = \frac{1}{2} \quad \Rightarrow \quad t^{K/2} = \frac{1}{2c^K} \quad \Rightarrow \quad \log t = -\frac{2}{K}\log(2c^K) = -\frac{2\log K}{K} + \mathcal{O}\left(\frac{1}{K}\right), \tag{24}$$

which yields:

$$t \approx \exp\left(-\frac{2\log K}{K}\right) \approx 1 - \frac{2\log K}{K}, \quad \text{if } K \gg 1. \tag{25}$$

However, since we are dealing with small values of $\upsilon$, we take a simpler approach using exponential tail bounds. For $\upsilon \sim \text{Beta}(\frac{1}{2}, \frac{n-1}{2}; \lambda)$, it can be shown directly from the geometry of Gaussian vectors that:

$$\mathbb{P}(\upsilon > t) \approx \exp(-\tfrac{n}{2}t), \quad \text{for large } t. \tag{26}$$

Therefore, the probability that none of the $K$ samples exceeds $t$ is:

$$\mathbb{P}(\nu_{\mathcal{C}_n^K} \le t) \approx \left(1 - \exp(-\tfrac{n}{2}t)\right)^K. \tag{27}$$

Set this equal to 1/2:

$$\left(1 - \exp(-\tfrac{n}{2}t)\right)^K = \frac{1}{2} \quad \Rightarrow \quad \exp(-\tfrac{n}{2}t) = \frac{1}{K}, \quad \text{hence} \quad t = \frac{2\log K}{n}. \tag{28}$$

**Conclusion:** We have shown that the maximum squared cosine similarity among $K$ independent samples $\upsilon_i \sim \text{Beta}(\frac{1}{2}, \frac{n-1}{2}; \lambda)$ satisfies:

$$\nu_{\mathcal{C}_n^K} \approx \frac{2\log K}{n}, \tag{29}$$

as $n \to \infty$ and $K$ remains polynomial in $n$. This approximation arises from the geometry of high-dimensional Gaussians and a sharp estimate of the upper tail of the non-central beta distribution.

$\square$

## 2 More Visualization Results

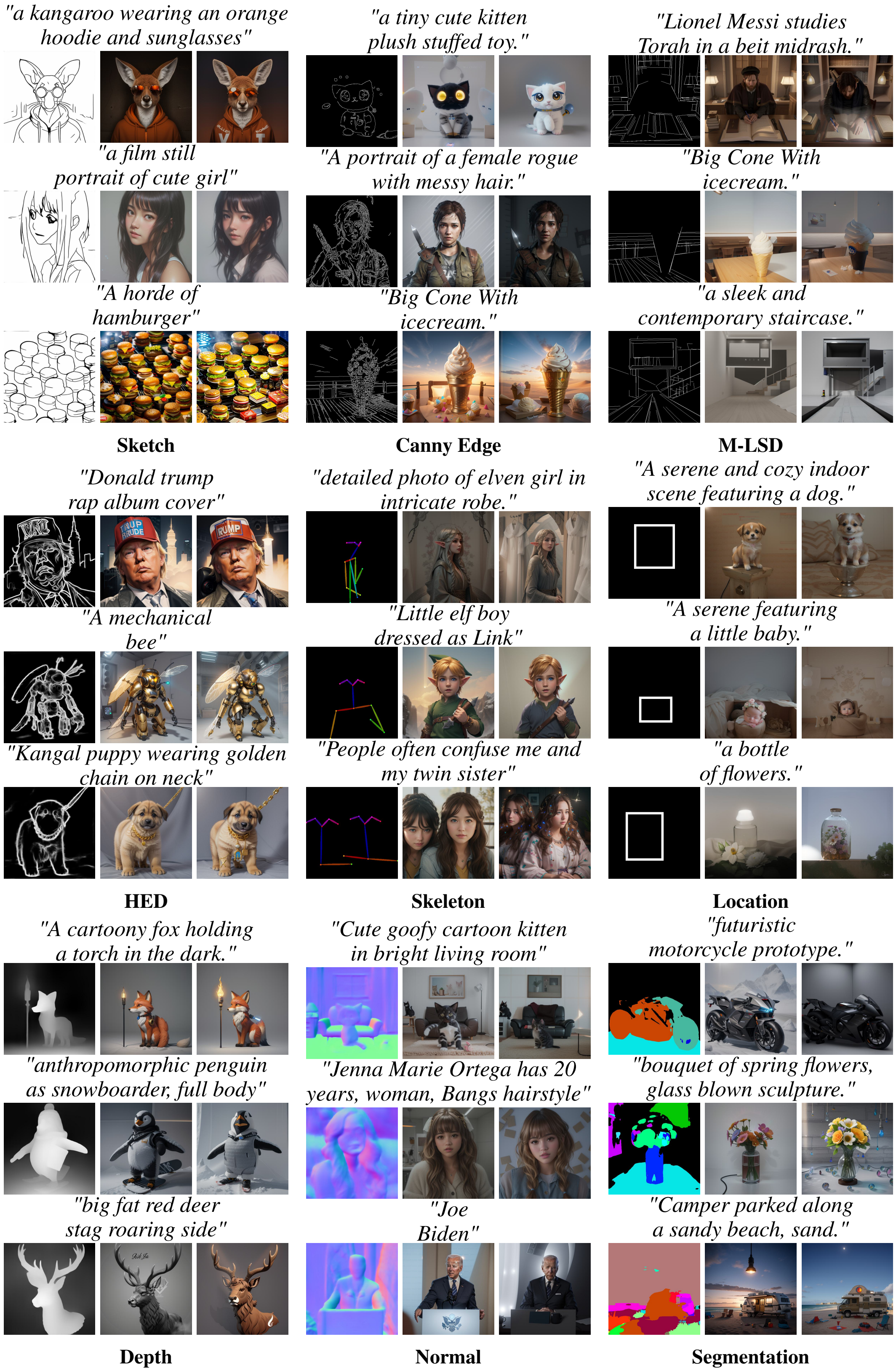


Figure 1: Conditional Generation with Diverse Conditions via Discrete Noise. Discrete noise enables seamless integration of various conditions into diffusion models. To demonstrate its versatility, we showcase ten diverse applications: Sketch, Canny Edge, MLSD, HED, Skeleton, Location, Depth, Normal, and Segmentation.

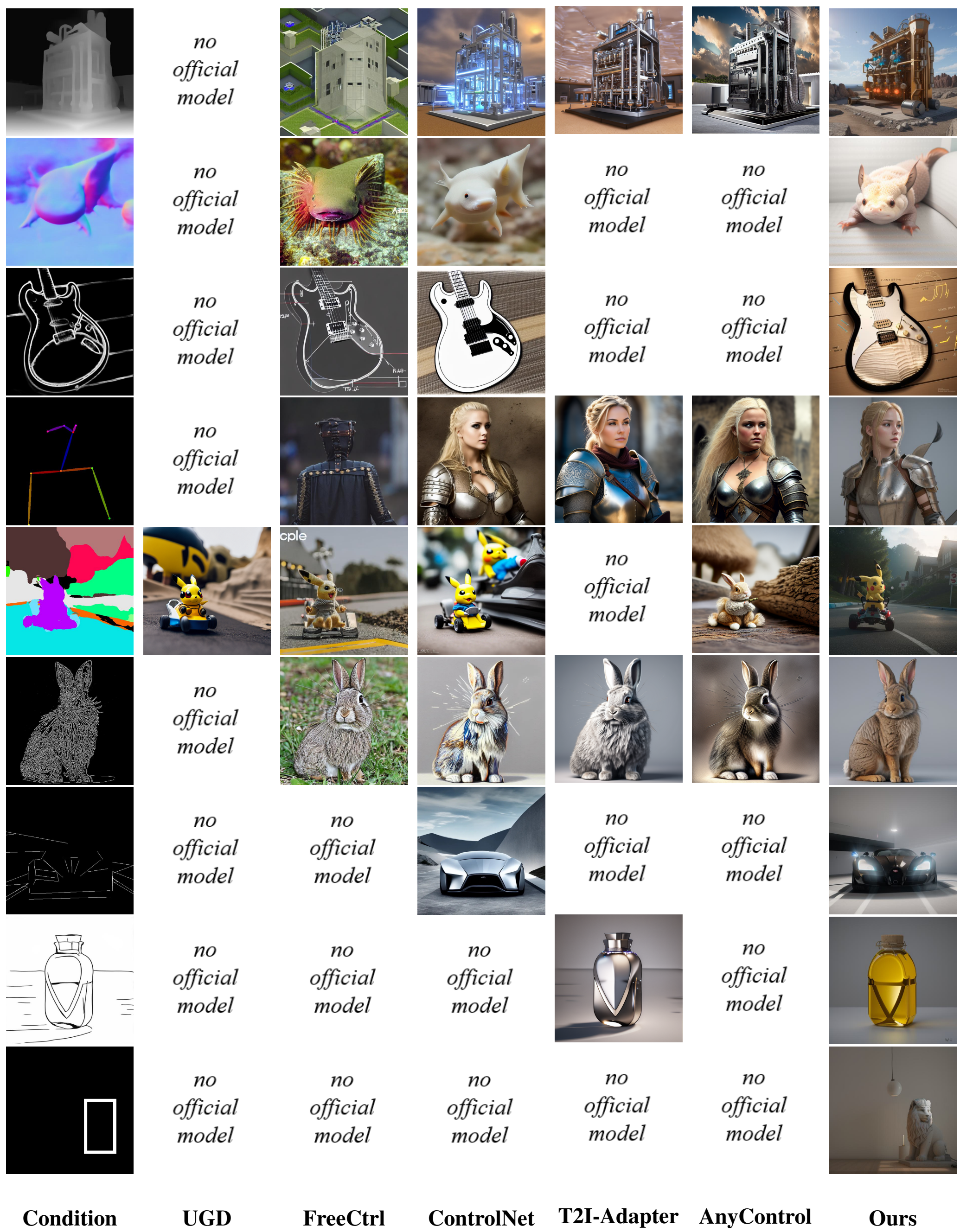


Figure 2: Qualitative Comparison for Controllable Generation under Different Conditions: We compare our method with five other approaches. Among them, UGD and FreeCtrl belong to the classifier-based guidance category, while ControlNet, T2I-Adapter, and AnyControl are part of the classifier-free guidance category. The results demonstrate that our method is competitive with state-of-the-art techniques.

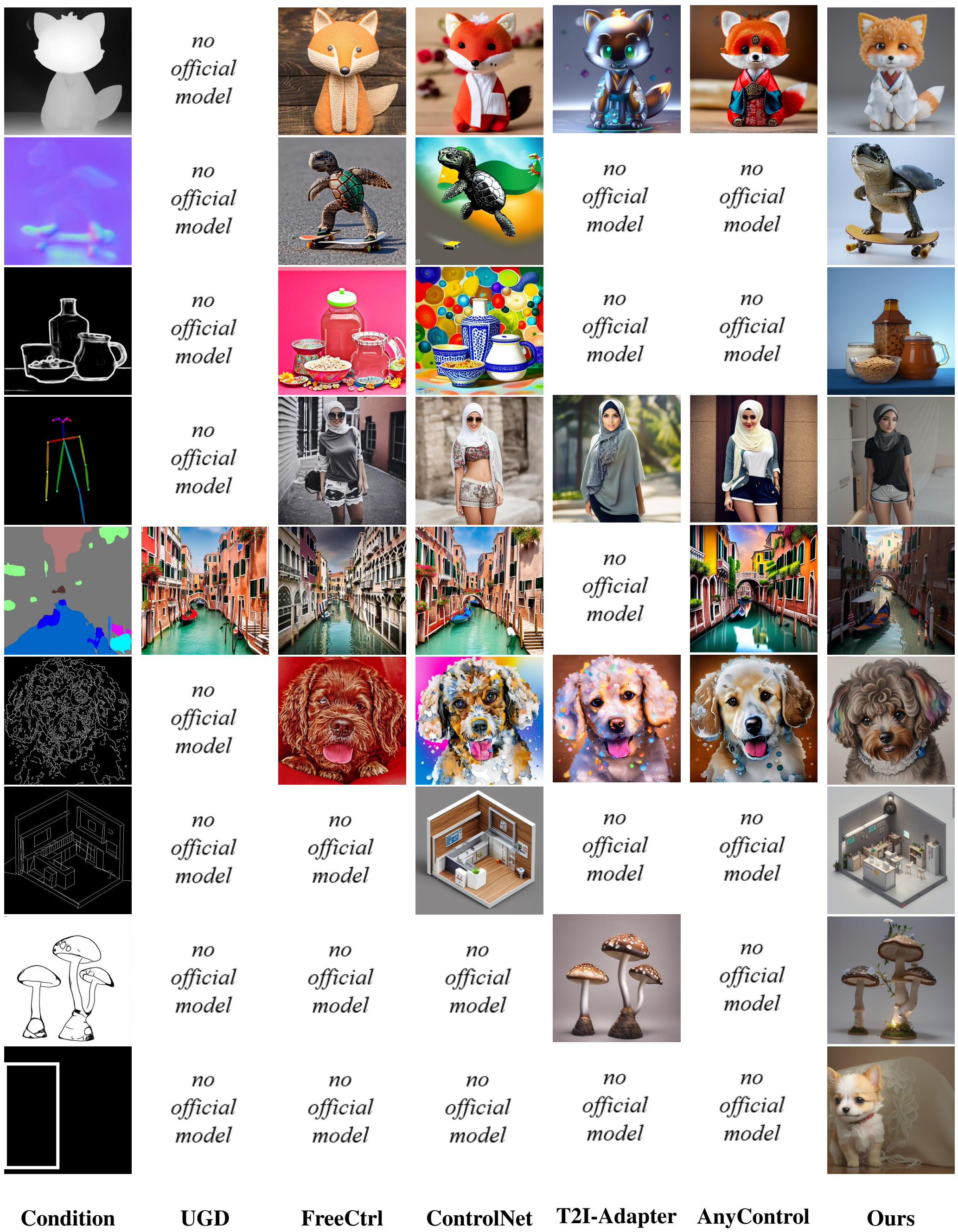


Figure 3: Qualitative Comparison for Controllable Generation under Different Conditions: We compare our method with five other approaches. Among them, UGD and FreeCtrl belong to the classifier-based guidance category, while ControlNet, T2I-Adapter, and AnyControl are part of the classifier-free guidance category. The results demonstrate that our method is competitive with state-of-the-art techniques.