**Introduction**

Goal: Learn diffusion models in the presence of conditioning noise

Approach: Estimate conditioning alignment with a coherence score

Prior work: Filtering the dataset into a high coherence subset [Rombach 2022]

Our method: Keep every sample but condition the model by the coherence score

---

**Coherence-aware diffusion (CAD)**

- Condition a diffusion model by the coherence score of the datapoint.

- \( L_{\text{model}} = -\mathbb{E}_{X \sim p(X)}\left[\log q(y | x, c)ight] \)

- \( (X, y) \) can be noisy, but \( X \) contains interesting information of the data manifold.

- By discarding \( X \), it would lose the information contained in \( X \).

- Providing the coherence score to the model instead, it can learn to choose to use the conditioning or not.

---

**Architecture: TextRIN**

- Deep diffusion models for text-to-image

- Text-to-image

- Class-conditioned:
  - Resample dataset given a label
  - Use Entropy as coherence score
  - Coherence-aware training benefits Conditional diffusion

---

**Coherence-aware free guidance**

- We adapt classifier free guidance in the following way:

- \( q(x, y) = q(x, y, 0) + \alpha q(x, y, 1) - q(x, 0, 0) \)

- - Varying the coherence at test time

- - Interpolating over the coherence

---

**Semantic conditioning**

- Class-conditional:
  - Captions: A bear emojis with shiny bear thrower and happy eyes.

- User study

- Qualitative results

- Quantitative results

---

**Text-to-image**

Captions: A cute little matte low poly isometric cherry blossom forest island waterfalls lighting soft volumetric lighting spectacular ambient lights bulu rendered in cosmic beach whirlpool engine bula rendered in cosmic beach whirlpool engine underwater catedral Pirate ship trapped in a cosmic maelstrom neon pollution cinematic atmosphere art nouveau piranha make up blue on red side profile looking serious eyes 50mm portrait photography portrait photo of a asia old warrior chief tribal painting of a zebra and snowman cow and horse dancing in a dusting of powdered sugar snow adorned with vibrant candy canes and shimmering gumdrops A cozy gingerbread house nestled in a dusting of powdered sugar snow adorned with vibrant candy canes and shimmering gumdrops A japanese samurai background with sakura tree a blue jay stops on the top of a helmet of Underwater catedral A old-world galleon navigating through a dodge of pirates.

---

**COCO-10K Metrics**

- **ADE20k**
  - Baseline: 150M TextRIN samples
  - Filtered: 150M TextRIN samples
  - CLIPScore: 150M TextRIN samples

- **Baseline CAD**
  - Text-to-image

- **CAD**
  - Text-to-image

---

**Class-conditional**

- **CLIPScore**
  - Text-to-image

---

**Coherence**

- **Confidence**
  - Varying the coherence at test time

- **Different levels of noise**
  - Interpolating over the coherence