Your Diffusion Model is Secretly a Noise Classifier and Benefits from Contrastive Training

Presenter: Yunshu Wu

Yunshu Wu¹, Yingtao Luo², Xianghao Kong¹, Evangelos E. Papalexakis¹, Greg Ver Steeg¹

¹ University of California Riverside

² Carnegie Mellon University

Probability = (Optimal) Denoising $\sum_{i=1}^{n}$ 2 Diffusion Model Background: Optimal Denoisers are Density Estimators 2 Diffusion Model Background: Optimal Denoisers are Density Estimators $S_h = \frac{1}{2} \sum_{i=1}^{n} \frac{1}{i} \sum_{i=1}^{n} \frac{1}{i!}$ we see that log-likelihood can be written *exactly* in terms of an expression that depends only on the \mathcal{M} solution to the Gaussian denomination of problem, i.e., α among the scaling conventions, please conventions, please conventions, please conventions, please check Apple

Guo, Shamai, Verdú, (IEEE 2005) – Information-MMSE relations Kong, Brekelmans, Ver Steeg (ICLR 2023) "Information-theoretic Diffusion" T defining feature of diffusion models is a sequence of distribution models is a sequence of distributions that α log *p*(*x*) = *c* + ¹*/*² E ² Guo, Shamai, Verdú, (IEEE 2005) – Information-MMSE relations

Gaussian noise channel

\n
$$
x_{\alpha} \equiv \sqrt{\sigma(\alpha)}x + \sqrt{\sigma(-\alpha)}\epsilon
$$
\nlog-SNR

\nN(0,1)

\n
$$
x
$$
\n
$$
x_{\alpha}
$$

 N egative Log-likelihood (NIL). we packed to *p* intermode (*i*xtly. Negative Log-likelihood (NLL): unknown data distribution for *^x* ² ^R*^d*, and (*·*) is the sigmoid function. We define the sequence of we see that log-likelihood can be written *exactly* in the seed of an expression that depends on the dependence on the dependence on M_{M} $\frac{1}{2}$ lihood (NIT):

• NLL for data distribution:
$$
-\log p(x) = c + \frac{1}{2} \int_{-\infty}^{\infty} \mathbb{E}_{p(\epsilon)} [\|\epsilon - \hat{\epsilon}(x_{\alpha}, \alpha)\|_2^2] d\alpha.
$$

• **NLL for noisy data distribution:**
$$
-\log p_{\zeta}(\boldsymbol{x}) = c + \frac{1}{2}.
$$

• NLL for noisy data distribution:
$$
-\log p_{\zeta}(\boldsymbol{x}) = c + \frac{1}{2} \int_{-\infty}^{\infty} d\alpha \mathbb{E}_{p(\boldsymbol{\epsilon})}[\|\boldsymbol{\epsilon} - \hat{\boldsymbol{\epsilon}}_{\zeta}(\boldsymbol{x}_{\alpha}, \alpha)\|_2^2]
$$

 \sim

$$
-\log p_{\zeta}(\boldsymbol{x}) = c + 1/2 \int_{-\infty}^{\infty} d\alpha \; \mathbb{E}_{p(\boldsymbol{\epsilon})}[\|\boldsymbol{\epsilon} - \boldsymbol{b} \cdot \hat{\boldsymbol{\epsilon}}(\boldsymbol{x}_{\alpha}, \beta] \|_2^2]
$$

Log likelihood ratios and classification

If we sample from either distribution with probability $P(C)=1/2$, then, Bayes rule relates the log likelihood ratio to the optimal classifier:

$$
LLR(x) \equiv \log \frac{p_1(x)}{p_0(x)} = \log \frac{p(C = 1 \mid x)}{p(C = 0 \mid x)}
$$

Idea: train classifiers to learn densities

Probability space

Gutmann & Hyvarinen (2012) "Noise contrastive estimation…"

Denoising = Classifying
$$
\text{source} \setminus \text{other} \setminus \text{other} \setminus \text{other} \setminus \text{There}
$$

Sample from either distribution with probability $q(y= \pm 1) =$, - , then, the optimal binary Bayes classifier is related to the log-likelihood ratio: Consider from other distribution, with probability $g(y-1, 1) = \frac{1}{2}$ then the optime of binary label in the transition of the same distribution with probability $q(y - \pm 1) - \frac{1}{2}$, then, the optimal $q(y - \pm 1)$. α oither dictribution with probability $g(y-1, 1) = \frac{1}{y}$ the *x*/(2 *x*/(1 *x*/(2 *x*/($q(x) = 1$ $y = 1$ 'aɒIIIty q(y= ± 1*) —*
'-likalihood ratio' *^q*(*x|y*))=1*/*(1 + exp(*y*(log *^q*(*x|^y* ⁼ 1) log *^q*(*x|^y* = 1)))) an with probability $a(y - 1) - \frac{1}{x}$ then the ontimal $\frac{1}{2}$ and when probability $4(y - 1 / 1) - 2'$ then, the optimal signation the log-likelihood ratio: Increase of diffusion model to the log-likelihood ratio:

$$
q(y|\boldsymbol{x}) = \frac{q(\boldsymbol{x}|y)q(y)}{q(\boldsymbol{x})} \qquad \log q(y|\boldsymbol{x}) = -\operatorname{softplus}(y \frac{q(\boldsymbol{x}|y = -1)}{q(\boldsymbol{x}|y = 1)})
$$

 $= p(x)$ be data dis data distribution, $q(x|y=-1) = p_{\zeta}(x)$ be noisy data distr Let $a(x|y=1) = p(x)$ be data distribution. $a(x|y=-1) = p_z(x)$ be noisy data distri Let $q(x|y-1) - p(x)$ be data distribution, $q(x|y-1)$ In the second secton $\mathbf{a}(\mathbf{x}|\mathbf{v}=-1)=p_z(x)$ be noisv data distribution Let q(x|y=1) = $p(x)$ be data distribution, q(x|y= - 1) = $p_{\zeta}(x)$ be noisy data distribution

LLR → Optimal classifier: $LLR = \log p_{\zeta}(x) - \log p(x)$ diffusion denoisers and noise classifiers to define a new training objective. We set the distributions diffusion denoisers and noise classifiers to define a new training objective. We set the distributions

IN the second line, because 8*y*, quality can constant can be constant out. The constants can just expanding cancel out. The cancel out of the cancel out. The cancel out of the constant of the constant of the cancel out. T *We have our Contrastive Diffusion Loss:* $\frac{1}{2}$ and we can write in terms of the write *d***(***x*) = 1) and *q*(*x*) = 1) to be two distributions at different noise levels that we can write in terms of α and *q*(*x|y* = 1) ⌘ *p*⇣ (*x*), for some noise level, ⇣. Then given a sample (*x, y*) ⇠ *q*(*x, y*) the We have our Contrastive Diffusion Loss:
We have our Contrastive Diffusion Loss:

$$
\mathcal{L}_{CDL} = \mathbb{E}_{q(\bm{x},y)}\left[\text{softplus}(y(\log p_{\zeta}(\bm{x}) - \log p(\bm{x})))\right]
$$

CDL – more general transport schemes

Profit : CDL training on OOD regions among these scaling conventions, please check App. B.3. log *p*(*x*) = *c* + ¹*/*² \bigcap ^E*p*(✏)[k✏ ✏ˆ(*x*↵*,* ↵)k²

3 What Your Diffusion Model is Hiding: Noise Classifiers

Profit : CDL training on OOD regions

Better distribution learning with hard constraints

Contrastive training improves "FID"

Table 2: Evaluating FID score (lower is better) of parallel DDPM sampler on real-world datasets using 5,0000 samples. "NA" stands for "Not Applicable". For reported FID scores, we run three sets of random seeds and reported the average with uncertainty.

Thank you!