Übungsblatt 9: Grundlagen Generativer Modelle
BA-INF 153: Einführung in Deep Learning für Visual Computing

Deadline:	**Theoretische Aufgaben**	10.07.2024 - 14:00 via eCampus
Tutoren:	Alina Pollehn	s6aapoll@uni-bonn.de
	Johannes van de Locht	s6jovand@uni-bonn.de
Übungsgruppenleitung:	Jan Müller	muellerj@cs.uni-bonn.de

1 Theoretische Übungsaufgaben

Permutationen mit 1×1 **Konvolutionen** ***(5 Punkte)*** Die invertierbare Netzwerkarchitektur für flow-based Modelle verwendet Permutationen um sicherzustellen, dass beide Partitionen der Eingabe auf eine Stichprobe der gelernten Verteilung abgebildet werden. Eine einfache Möglichkeit dies zu implementieren wäre es nach jedem flow Layer eine zufällige Permutation der Partitionen durchzuführen. Es hat sich aber gezeigt, dass das Lernen von Permutationen zu besseren Ergebnissen führt. Kingma und Prafulla, 2018 schlagen vor, statt einer zufälligen Permutation zwischen den Affine Coupling Layer, die Permutation mit Hilfe von 1×1 Konvolutionen zu lernen. Die Permutation erfolgt entlang der Featuredimension des Ouputs eines Affine Coupling Layer für jeden Pixel individuell.

Aufgabe: Erläutern Sie wie sich eine solche Permutation mit Hilfe einer 1×1 Konvolution darstellen lässt. Berücksichtigen Sie dabei auch die Anforderungen, die ein normalizing flow erfüllen soll.

Details im Diffusion-Prozess ***(10 Punkte)***

Im Diffusionsprozess bildet die Vorwärtsdiffusion eine Zufallvariable x_0, die entsprechend der uns unbekannten "echten" Verteilung $p(x)$ verteilt ist. Durch einen **Markov-Prozess / Markov-Kette** $q(x_t \mid x_{t-1})$ wird diese auf eine Zufallvariable x_T abgebildet, so dass diese (im Limit $T \to \infty$) der Verteilung unseres Priors folgt. Bei der Rückwärtsdiffusion wird dieser Prozess umgekehrt. Stichproben x_T der Priorverteilung sollen mit Hilfe der Posteriori-Verteilung $q(x_{t-1} \mid x_t)$ iterativ so umgewandelt werden, dass diese echten Verteilung von $p(x)$ entsprechen.

Da die Posteriori-Verteilung unbekannt ist, wird diese durch ein Neuronales Netzwerk $p_\theta(x_{t-1} \mid x_t)$ modelliert. Das Netzwerk wird trainiert die Vorwärtsdiffusion auf Beispielen des Datensatz umzukehren.

- (4 Punkte) Zeigen Sie, dass die Übergangswahrscheinlichkeiten der Vorwärtsdiffusion auch durch die Posteriori-Verteilung ausgedrückt werden können d.h.

$$q(x_t \mid x_{t-1}) = \frac{q(x_{t-1} \mid x_t, x_0) q(x_t \mid x_0)}{q(x_{t-1} \mid x_0)}.$$

  Hinweis: Verwenden Sie die Markov-Eigenschaft des Diffusionsprozess.

- (6 Punkte) Zeigen Sie, die eine Minimierung der Zielfunktion des Diffusions Modells tatsächlich die KLD zwischen $p_\theta(x_{t-1} \mid x_t)$ und $q(x_{t-1} \mid x_t)$ minimiert indem Sie zeigen, dass die folgende Gleichheit gilt:

$$\begin{aligned}
&\mathbb{E}_{q(x_{1:T}|x_0)}\left[-\log\left(p_\theta(x_T)\right) - \sum_{t=2}^{T} \log\left(\frac{p_\theta(x_{t-1} \mid x_t) q(x_{t-1} \mid x_0)}{q(x_{t-1} \mid x_t, x_0) q(x_t \mid x_0)}\right) - \log\left(\frac{p_\theta(x_0 \mid x_1)}{q(x_1 \mid x_0)}\right)\right] \\
=&\mathbb{E}_{q(x_{1:T}|x_0)}\left[-\log\left(\frac{p_\theta(x_T)}{q(x_T \mid x_0)}\right) - \sum_{t=2}^{T} \log\left(\frac{p_\theta(x_{t-1} \mid x_t)}{q(x_{t-1} \mid x_t, x_0)}\right) - \log\left(p_\theta(x_0 \mid x_1)\right)\right].
\end{aligned}$$