Состязательные Мосты Шрёдингера на задаче перевода изображения в изображение

Григорий Сергеевич Ксенофонтов
Научный руководитель: к.ф.-м.н. Р. В. Исаченко

Кафедра интеллектуальных систем ФПМИ МФТИ
Специализация: Интеллектуальный анализ данных
Направление: 03.04.01 Прикладные математика и физика

2024

Цель исследования

Задача

Найти прямое и обратное отображения между двумя наборами изображений с использованием мостов Шрёдингера.

Проблема

Современные подходы для отображения моделируют стохастические процессы, что вычислительно трудно.

Решение

Решить задачу мостов Шрёдингера с помощью состязательного обучения.

Постановка задачи

Пусть дано:

1. $\{x_i\}_{i=0}^N = X, \{y_j\}_{j=0}^M = Y$ – два непарных набора данных, где $x_0 = x \sim \pi_0(x)$ и $x_T = y \sim \pi_T(y)$,
2. $p^{\mathbb{W}^\gamma}(x_0, \dots, x_T)$ – совместное распределение априорного Винеровского процесса $\mathbb{W}^\gamma$.

Динамическая постановка мостов Шредингера:

$$\min_{q \in \mathcal{D}(\pi_0, \pi_T)} D_{KL}(q(x_0, \dots, x_T) || p^{\mathbb{W}^\gamma}(x_0, \dots, x_T)),$$

где $\mathcal{D}(\pi_0, \pi_T)$ – множество всех совместных распределений с маргиналами $\pi_0(x_0), \pi_T(x_T)$

Проблема

Моделирование процесса с помощью марковской цепи – вычислительно трудная задача.

$$q(x_0, \dots, x_T) = \pi_1(x_T) \prod_{t=0}^{T-1} q_{t|t+1}(x_t | x_{t+1})$$

Статическая постановка задачи мостов Шрёдингера

Избавится от моделирования марковской цепи позволяет статическая постановка задачи мостов Шрёдингера:

$$q^*(x, y) = \arg \min_{q \in \mathcal{D}(\pi_0, \pi_T)} D_{KL}(q(x, y) || p^{\mathbb{W}^\gamma}(x, y)).$$

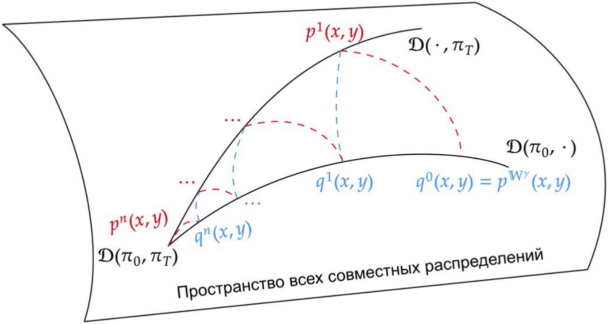
Таким образом, отображение происходит за один шаг:

$$q(x, y) = \pi_1(x_T) q(y|x)$$

Iterational Proportional Fitting (IPF)

Задача мостов Шрёдингера решается с помощью итеративного алгоритма Iterational Proportional Fitting (IPF):

$$\underbrace{\min_{p \in \mathcal{D}(\cdot, \pi_T)} D_{KL}(p(x,y)||q^*(x,y))}_{\text{обратный шаг}}; \underbrace{\min_{q \in \mathcal{D}(\pi_0, \cdot)} D_{KL}(q(x,y)||p^*(x,y))}_{\text{прямой шаг}}$$



Предложенный метод: вариационное представление

Обратный шаг

$$\min_{p(x|y)} \max_{D} \mathbb{E}_{y\sim\pi_T}\mathbb{E}_{x\sim p(x|y)}\left[D(x,y)\right] - \mathbb{E}_{x\sim\pi_0}\mathbb{E}_{y\sim q^*(y|x)}\left[e^{D(x,y)-1}\right]$$

Прямой шаг

$$\min_{q(y|x)} \max_{D} \mathbb{E}_{x\sim\pi_0}\mathbb{E}_{y\sim q(y|x)}\left[D(x,y)\right] - \mathbb{E}_{y\sim\pi_T}\mathbb{E}_{x\sim p^*(x|y)}\left[e^{D(x,y)-1}\right]$$

[1] f-GAN: Training Generative Neural Samplers using Variational Divergence Minimization

[2] Conditional Generative Adversarial Nets

Постановка экспериментов

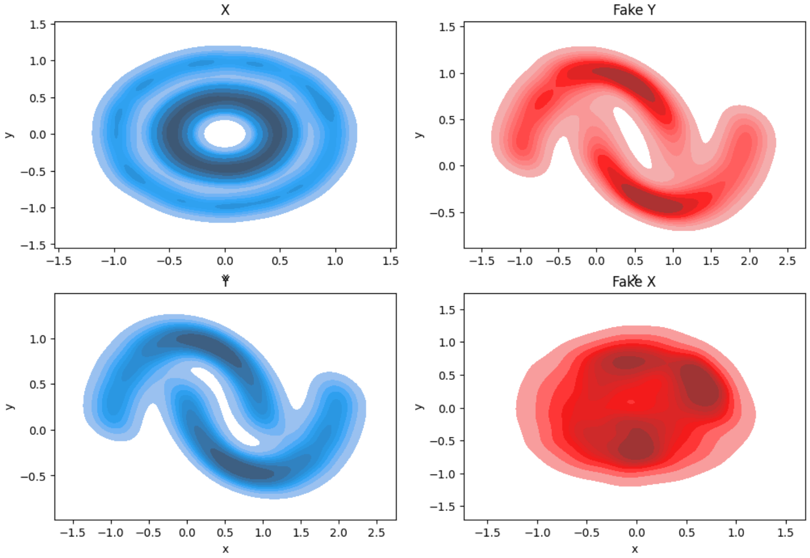
Эксперимент на 2D данных

- **Цель:** санитарное тестирование предложенного метода;
- **Данные:** X - точки из распределения колец, Y - точки из распределения лун.

Эксперимент на EMNIST

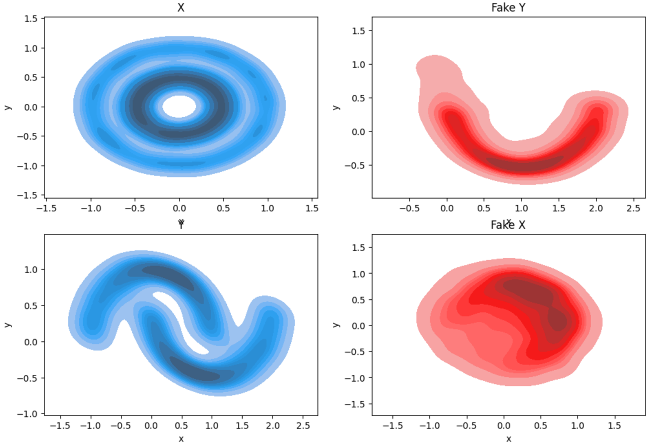
- **Цель:** проверить работоспособность метода на данных большей размерности;
- **Данные:** X - буквы из EMNIST, Y - цифры из EMNIST.

2D данне

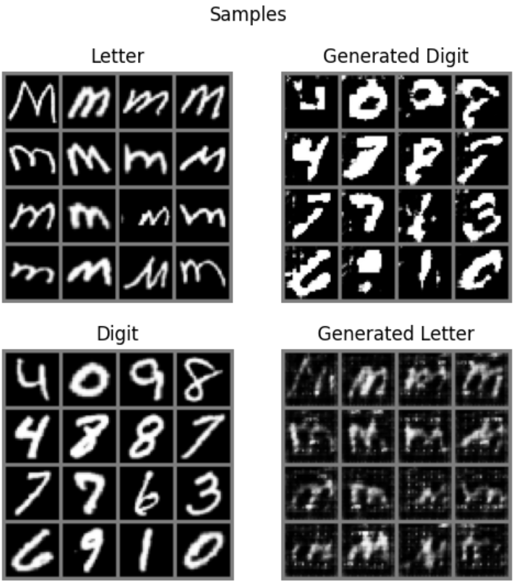


Коллапс моды

Предложенный подход наследует недостатки GAN.
Неаккуратный выбор параметров может привести к коллапсу моды:



EMNIST



Выносится на защиту

1. Предложен новый метод нахождения мостов Шрёдингера, который не строит стохастического процесса между распределениями.
2. Проведены эксперименты и выявлены недостатки подхода.