

Medical Flow Matching CT Translation

Крейнин Матвей

Московский физико-технический институт
Кафедра интеллектуальных систем

Научный руководитель: к.ф.-м.н. А.В. Грабовой
2025

Проблематика

- **Проблема:** классические методы диффузионных моделей требуют решения обратного стохастического процесса, что дорого и медленно.
- **Цель:** Заменить стохастический процесс на детерминированный поток (ODE) и ускорить генерацию с улучшением качества генерации, при возможности направлять процесс в две стороны.
- **Решение:** Обучить нейронную сеть аппроксимировать векторное поле между двумя распределениями π_0 и π_1 .

Flow Matching

Постановка задачи

Даны распределения π_0 (источник) и π_1 (цель) в \mathbb{R}^d . Найти векторное поле $v_\theta : \mathbb{R}^d \times [0, 1] \rightarrow \mathbb{R}^d$, такое что решение ОДУ

$$\frac{dX_t}{dt} = v_\theta(X_t, t), \quad X_0 \sim \pi_0,$$

переносит π_0 в π_1 при $t = 1$.

Путь

Линейная траектория

$$X_t = (1 - t)X_0 + tX_1, \quad (X_0, X_1) \sim \pi_0 \times \pi_1.$$

Тогда эталонное поле постоянно:

$$v^*(x, t) = X_1 - X_0 \quad (\text{не зависит от } t).$$

Flow Matching

Функция потерь

Обучаем v_θ так, чтобы он совпадал с v_{target} в среднем по промежуточным скоростям:

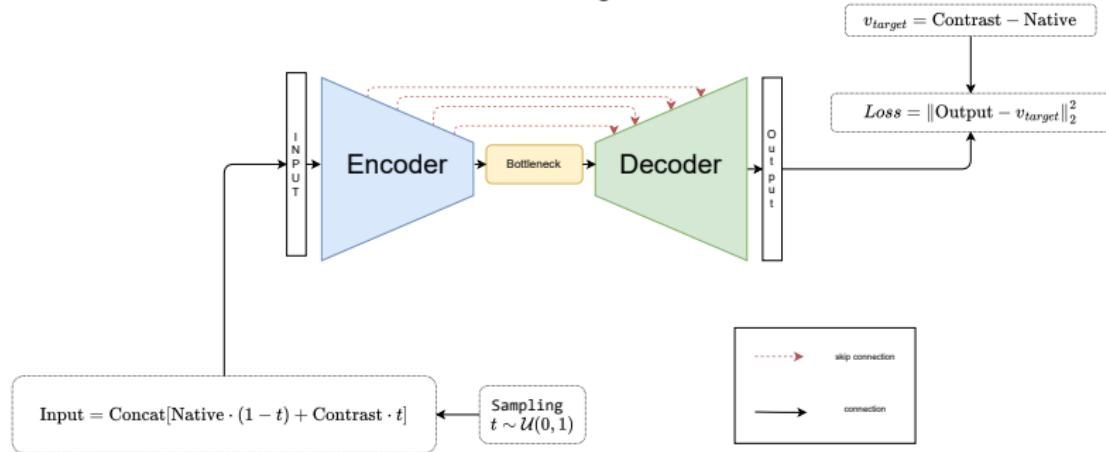
$$\min_{\theta} \mathbb{E}_{t \sim \mathcal{U}[0,1]} \mathbb{E}_{X_0, X_1 \sim \pi_0 \times \pi_1} [\|v_\theta(X_t, t) - v^*\|^2]$$

Генерация

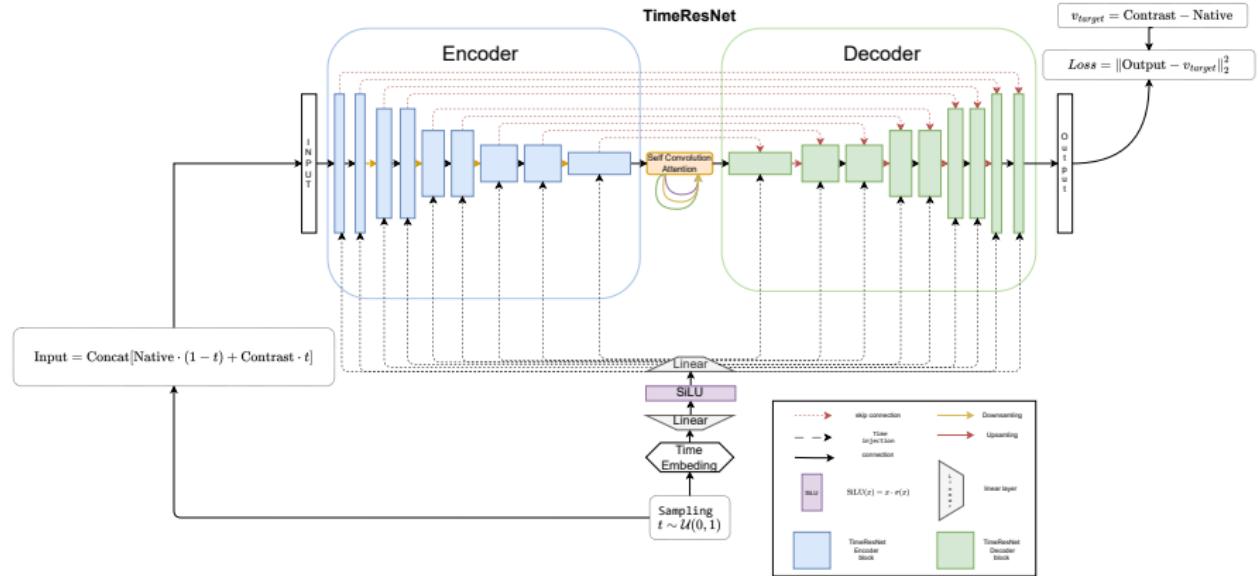
Объект из π_0 и численно интегрируем $\dot{X}_t = v_\theta(X_t, t)$ от $t = 0$ до $t = 1$.
Объект из π_1 и численно интегрируем $\dot{X}_t = v_\theta(X_t, t)$ от $t = 1$ до $t = 0$.

Схема метода

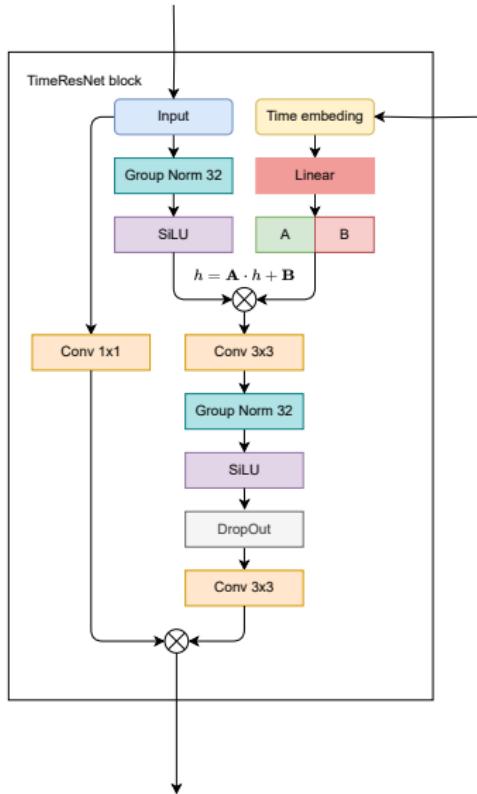
FlowMatching



TimeResNet



TimeResNetBlock



Постановка эксперимента

- **Задача:** аппроксимация векторного поля.
- **Выборка:** 120 исследований (52561 2д изображений), 80 (34992) - обучение, 20 (8197) - тест, 20 (9372) - отложенная.
- **Архитектуры:** TimeResNet, DiffusionNet, SwinUNETR, SegResNet.
- **Постановка эксперимента:**
 - ① Регистрируем изображения с контрастом на изображения без контраста.
 - ② Локализуем область исследования с помощью сегментации тела.
 - ③ Обрезаем картинку по (-1000, 1000) HU, переводим в (-1, 1).
 - ④ Сравниваем метрики и визуальные результаты с диффузией и регрессией.

Численные результаты

Model	MAE \downarrow	SSIM \uparrow	PSNR \uparrow	Time, s \downarrow	Params, M \downarrow
TimeResNet (ours)	5.436	0.996	<u>39.776</u>	0.209	124.7
SwinUNETR	6.229	<u>0.992</u>	38.933	<u>0.087</u>	<u>120.1</u>
SegResNet	<u>6.146</u>	0.983	40.108	0.056	214.1
DiffusionNet	18.203	0.934	30.051	44.200	108.4

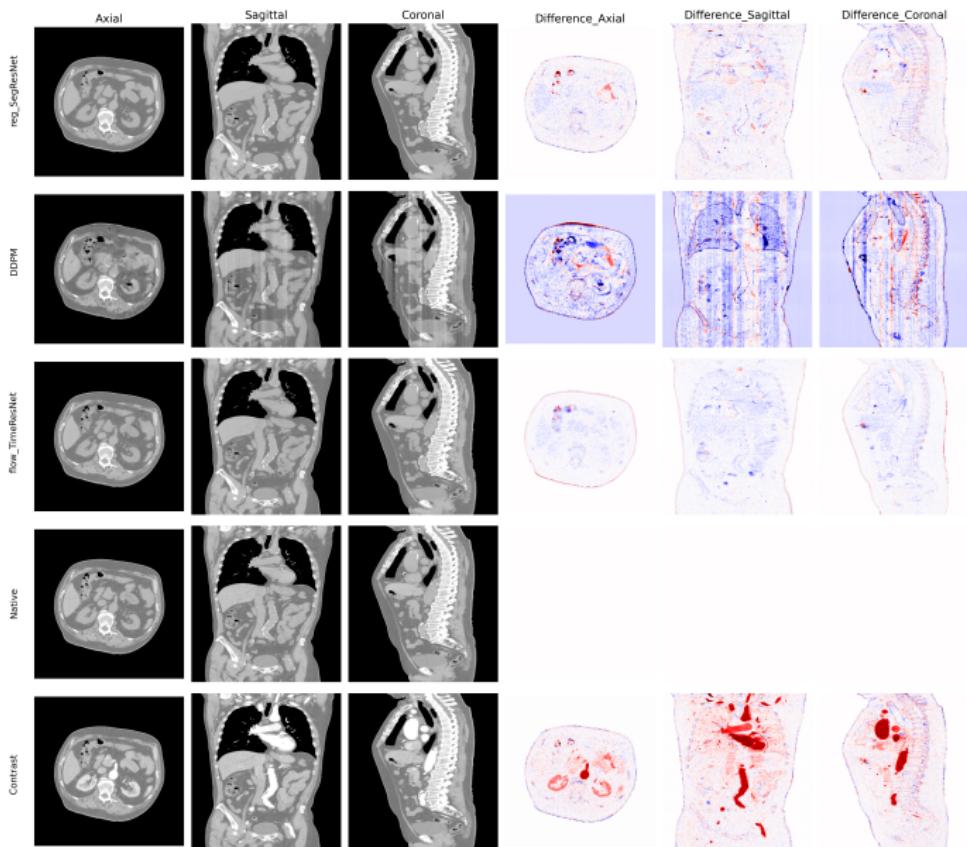
Таблица: Сравнение метрик моделей на hold-out выборке. TimeResNet и SwinUNETR обучены на задаче Flow Matching, SegResNet обучен на задачу регрессии, DiffusionNet обучен на задачу диффузии.

Выбор солвера

Solver (steps)	MAE↓	SSIM↑	PSNR↑	Time, s↓	Complexity
Euler (1)	6.362	0.991	38.182	0.105	1
Euler (2)	5.525	0.995	39.017	<u>0.210</u>	2
Euler (3)	<u>5.472</u>	0.995	38.994	0.313	3
Euler (5)	5.526	0.995	38.951	0.521	5
RK2 (1)	5.436	0.996	39.776	0.209	2
RK3 (1)	5.553	0.995	39.122	0.313	3
RK4 (1)	5.716	0.995	<u>39.257</u>	0.419	4
Mid Point (1)	5.611	0.995	39.047	<u>0.210</u>	2
Mid Point (3)	5.759	0.995	38.923	0.628	6

Таблица: Сравнение ODE-солверов для TimeResNe на задаче (Contrast → Native) на *hold-out*-выборке.

Визуальные результаты



Выносится на защиту

- Придуманы и имплементированы TimeResNet и TimeResNet block.
- Проведены значимые эксперименты, которые показывают что TimeResNet даёт преимущество по сравнению с нейронными сетями из фреймворка MONAI.
- Получена скорость генерации **400×** быстрее диффузионных моделей, качество значительно превышает регрессию.
- На задаче сегментации, nnUNetV2 (фреймворк automl) на синтетических данных показывает Dice score = 0.926 против = 0.966 на реальных данных.
- **Перспективы:** условная генерация образований в органах.