

Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold

Косса Николай Евгеньевич,
Студент группы БПМИ202

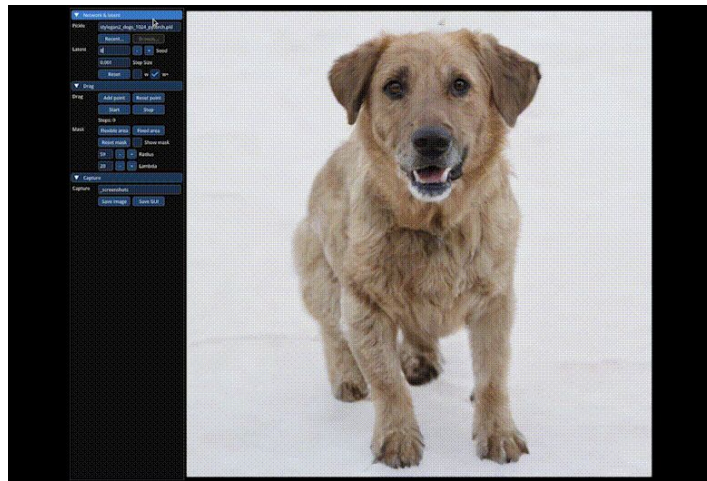
О чем эта статья?

- **DragGAN**
- **Существующие решения**
- **Подход**
- **Эксперименты**
- **Закличение**

DragGAN

и какие задачи он решает

DragGAN предоставляет способ управления GAN путем «перетаскивания» любых точек изображения для точного достижения целевого изображения в интерактивном режиме с пользователем.



генеративные модели для создания интерактивного контента

- Unconditional GANs
- Conditional GANs
- UserControllableLT
- 3D-aware GANs
- Diffusion Models

ОТСЛЕЖИВАНИЕ ТОЧЕК

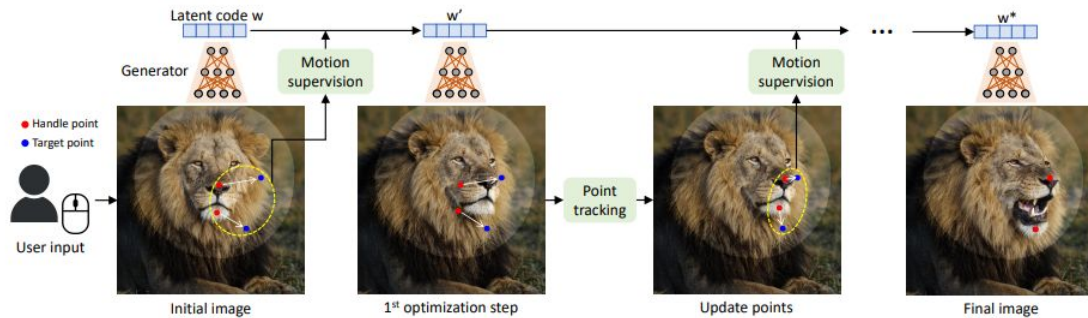
- Ручные критерии
- RAFT
- PIPs

ПОДХОД

- Интерактивная точечная манипуляция
- Motion Supervision
- Отслеживание точек

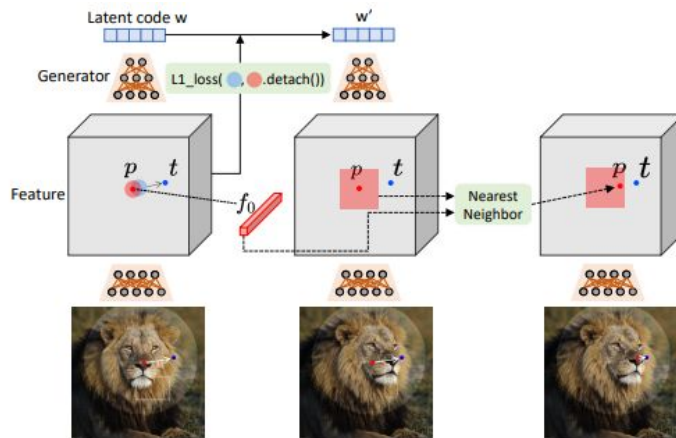
интерактивная точечная манипуляция

Цель состоит в том, чтобы переместить объект на изображении так, чтобы смысловые положения начальных точек достигали соответствующих целевых точек. Также разрешается дополнительно нарисовать маску, обозначающую, какая область изображения подвижна.



motion supervision

Ключевая идея заключается в том, что промежуточные факторы генератора очень различимы, так что для контроля движения достаточно простой функции потерь.



$$\mathcal{L} = \sum_{i=0}^n \sum_{\mathbf{q}_i \in \Omega_1(\mathbf{p}_i, r_1)} \|\mathbf{F}(\mathbf{q}_i) - \mathbf{F}(\mathbf{q}_i + \mathbf{d}_i)\|_1 + \lambda \|(\mathbf{F} - \mathbf{F}_0) \cdot (1 - \mathbf{M})\|_1,$$

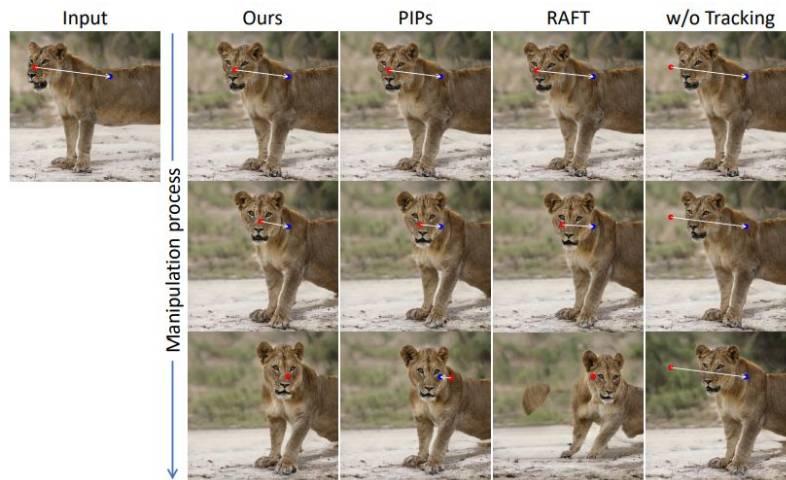
ОТСЛЕЖИВАНИЕ ТОЧЕК

Дискриминационные функции GAN фиксируют соответствие, и, таким образом, отслеживание может быть эффективно выполнено с помощью поиска ближайшего соседа в окрестности факторов.

качественная оценка



качественная оценка



КОЛИЧЕСТВЕННАЯ ОЦЕНКА

Манипуляция чертами лица

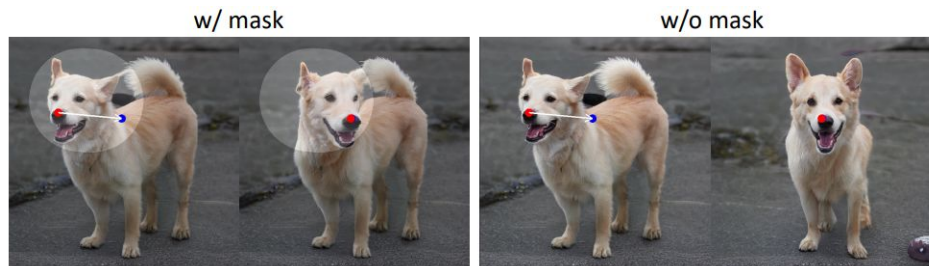
Method	1 point	5 points	68 points	FID	Time (s)
No edit	12.93	11.66	16.02	-	-
UserControllableLT	11.64	10.41	10.15	25.32	0.03
Ours w. RAFT tracking	13.43	13.59	15.92	51.37	15.4
Ours w. PIPs tracking	2.98	4.83	5.30	31.87	6.6
Ours	2.44	3.18	4.73	9.28	2.0

Реконструкция парного изображения

Dataset Metric	Lion		LSUN Cat		Dog		LSUN Car	
	MSE	LPIPS	MSE	LPIPS	MSE	LPIPS	MSE	LPIPS
UserControllableLT	1.82	1.14	1.25	0.87	1.23	0.92	1.98	0.85
Ours w. RAFT tracking	1.09	0.99	1.84	1.15	0.91	0.76	2.37	0.94
Ours w. PIPs tracking	0.80	0.82	1.11	0.85	0.78	0.63	1.81	0.79
Ours	0.66	0.72	1.04	0.82	0.48	0.44	1.67	0.74

результаты экспериментов

Эффекты маски



Out-of-distribution манипуляции



закключение

Авторы представляют DragGAN, интерактивный подход к интуитивному точечному редактированию изображений. Метод использует предобученный GAN для синтеза изображений, которые не только точно повторяют вводимые пользователем данные, но и сохраняют реалистичность.

