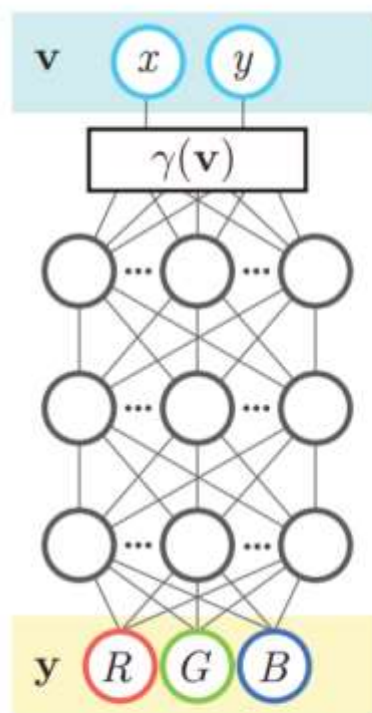


Использование преобразования Фурье для обучения высокочастотные функции в низкоразмерных областях

Охрименко Дмитрий, 172

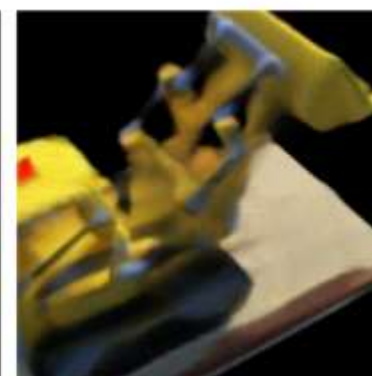
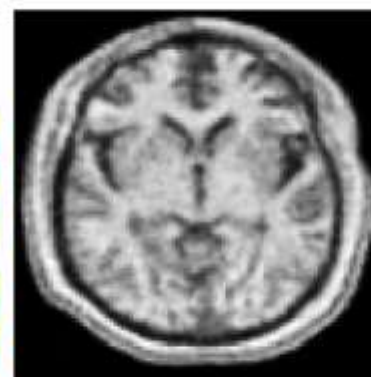
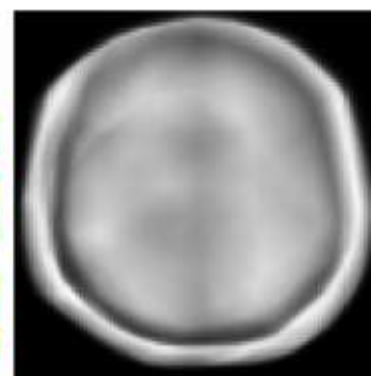
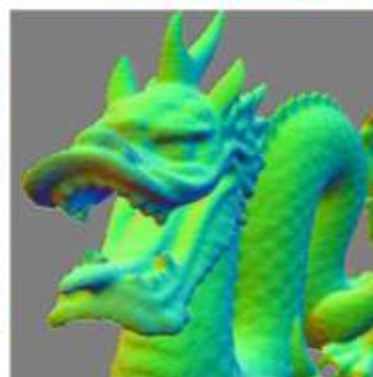
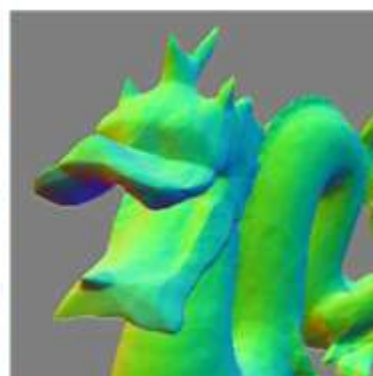
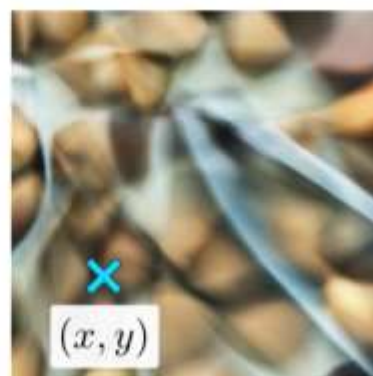


No Fourier features

$$\gamma(\mathbf{v}) = \mathbf{v}$$

With Fourier features

$$\gamma(\mathbf{v}) = \text{FF}(\mathbf{v})$$



(a) Coordinate-based MLP

(b) Image regression
 $(x, y) \rightarrow \text{RGB}$

(c) 3D shape regression
 $(x, y, z) \rightarrow \text{occupancy}$

(d) MRI reconstruction
 $(x, y, z) \rightarrow \text{density}$

(e) Inverse rendering
 $(x, y, z) \rightarrow \text{RGB, density}$

Проблемы обычного MLP

- MLP испытывают трудности с изучением высокочастотных функций (спектральное смещение)
- Эвристическое синусоидальное отображение входных координат («позиционное кодирование») решает эту проблему
- Преобразование Фурье – его частный случай: $\phi(v) = [a_1 \cos(2\pi b_1^T v), a_1 \sin(2\pi b_1^T v), \dots, a_m \cos(2\pi b_m^T v), a_m \sin(2\pi b_m^T v)]^T$
- Оно преобразует NTK в стационарное (инвариантное к сдвигу) ядро и позволяет настраивать спектр NTK

Что даст использование преобразования Фурье

- Преобразование Фурье можно использовать для преодоления спектрального смещения основанных на координатах MLP в сторону низких частот, позволяя им изучать гораздо более высокие частоты (раздел 4).
- Случайное преобразование Фурье с правильно выбранным масштабом может значительно улучшить производительность основанных на координатах MLP для многих низкоразмерных задач в компьютерном зрении и графике

Бекграунд

- Ядровая регрессия: $\hat{f}(x) = \sum_{i=1}^n (K^{-1}y)_i k(x_i, x)$, $K_{ij} = k(x_i, x_j)$, k - симметричная положительно полуопределенная ядерная функция похожести двух векторов
- Когда ширина слоя стремится к бесконечности, а время обучения к 0, функция $f(x, \theta)$ сходится к решению ядерной регрессии с использованием $k_{NTK}(x_i, x_j) = E_{\theta \sim N} \left\langle \frac{\partial f(x_i; \theta)}{\partial \theta}, \frac{\partial f(x_j; \theta)}{\partial \theta} \right\rangle$.
- Спектральное смещение: $Q^T \left(\hat{y}_{train}^{(t)} - y \right) = -e^{-\eta \lambda t} Q^T y$.
Компоненты целевой функции, которые соответствуют собственным векторам ядра с большими собственными значениями, будут изучаться быстрее.

Почему рассматриваются задачи с низкой размерностью

- Хотелось, чтобы составной NTK был инвариантным относительно сдвига во входной области
- Хотелось контролировать пропускную способность NTK для повышения скорости обучения

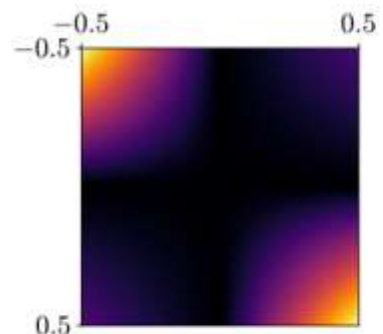
Преобразование Фурье

- Преобразование Фурье: $\phi(v) = [a_1 \cos(2\pi b_1^T v), a_1 \sin(2\pi b_1^T v), \dots, a_m \cos(2\pi b_m^T v), a_m \sin(2\pi b_m^T v)]^T$
- $\cos(\alpha - \beta) = \cos(\alpha) \cos(\beta) + \sin(\alpha) \sin(\beta)$, тогда ядровая функция:
- $k_\phi(v_1, v_2) = \phi(v_1)^T \phi(v_2) = \sum_{j=1}^m a_j^2 \cos(2\pi b_j^T (v_1 - v_2)) = h_\phi(v_1 - v_2)$, где $h_\phi(v\Delta) = \sum_{j=1}^m a_j^2 \cos(2\pi b_j^T v\Delta)$
- Тогда мэппинг – аппроксимация преобразования Фурье ядра, где b – базисные частоты Фурье, a – коэффициенты ряда Фурье

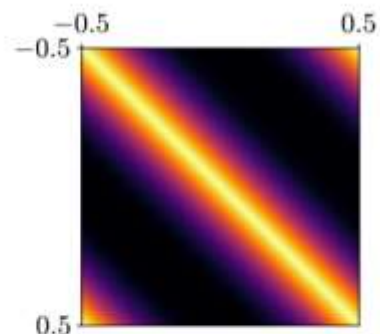
Преобразование Фурье

- После применения преобразования мы пропускаем данные через MLP. Результат может быть аппроксимирован ядровой регрессией с ядром
$$h_{NTK}(x_i^T x_j) = h_{NTK}(\phi(v_i)^T \phi(v_j)) = h_{NTK}(h_\phi(v_i - v_j))$$

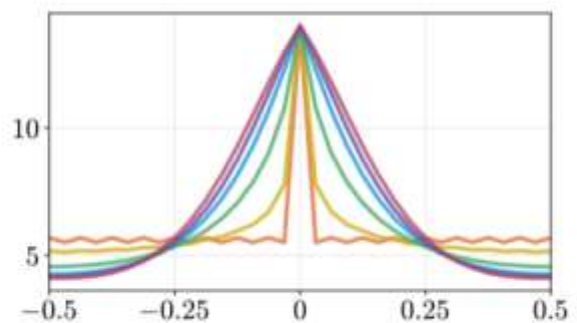
Подкрутка параметров



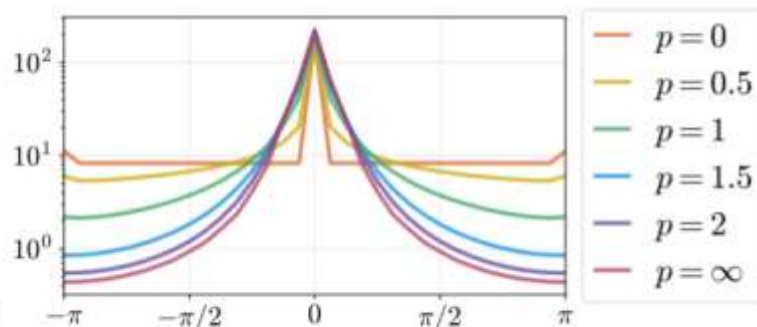
(a) No mapping NTK



(b) Basic mapping NTK



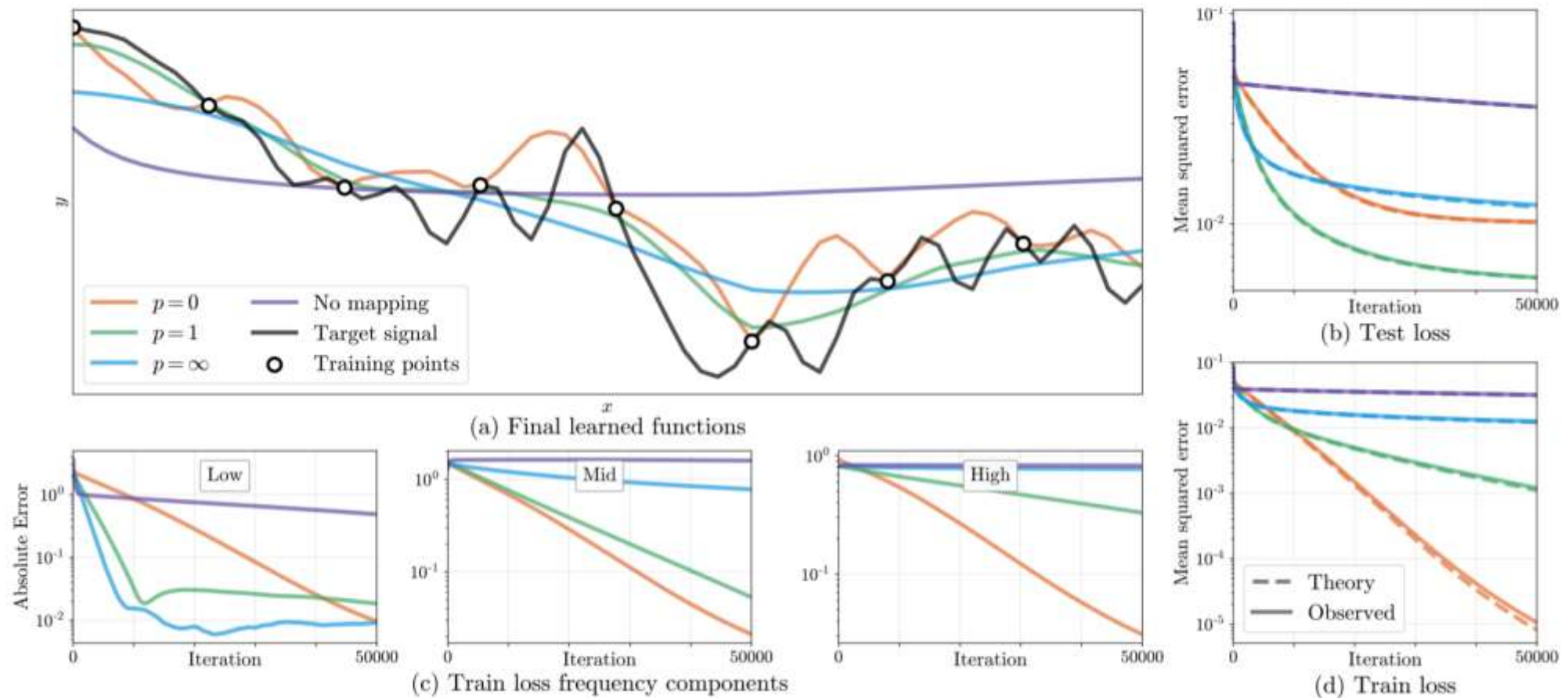
(c) NTK spatial



(d) NTK Fourier spectrum

$$a_j = \frac{1}{j^p}, b_j = j$$

Влияние преобразования Фурье на сходимость сети

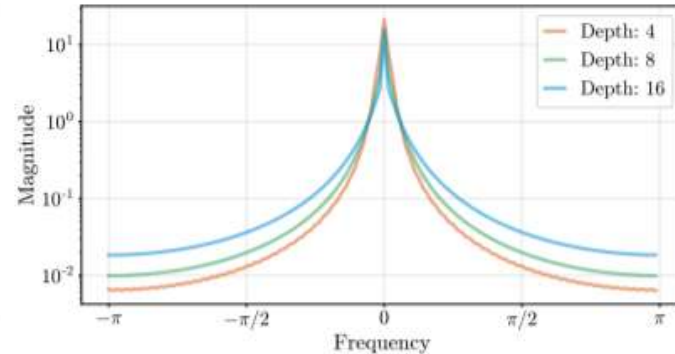
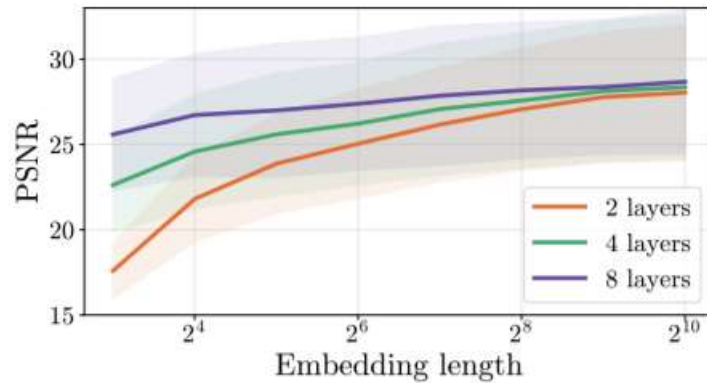


Различные виды отображений

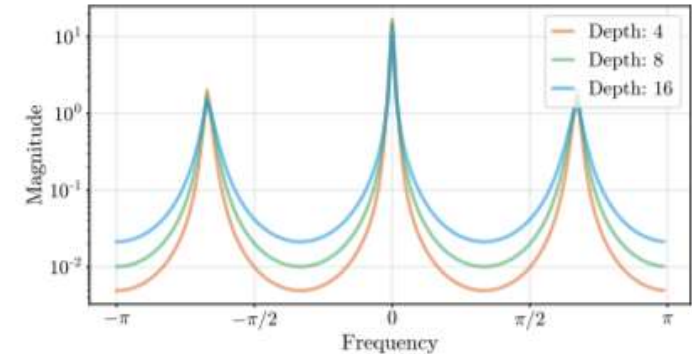
	Direct supervision			Indirect supervision			
	2D image		3D shape [24]	2D CT		3D MRI	3D NeRF [27]
	Natural	Text		Shepp	ATLAS	ATLAS	
No mapping	19.32	18.40	0.864	16.75	15.44	26.14	22.41
Basic	21.71	20.48	0.892	23.31	16.95	28.58	23.16
Positional enc.	24.95	27.57	0.960	26.89	19.55	32.23	25.28
Gaussian	25.57	30.47	0.973	28.33	19.88	34.51	25.48

- Basic: $\phi(v) = [\cos(2\pi v), \sin(2\pi v)]^T$
- Positional encoding: $\phi(v) = [\dots, \cos(2\pi\sigma^{j/m}v), \sin(2\pi\sigma^{j/m}v), \dots]^T$
- Gaussian: $\phi(v) = [\cos(2\pi Bv), \sin(2\pi Bv)]^T, B \sim N(0, \sigma^2)$

Эксперимент с глубиной сети



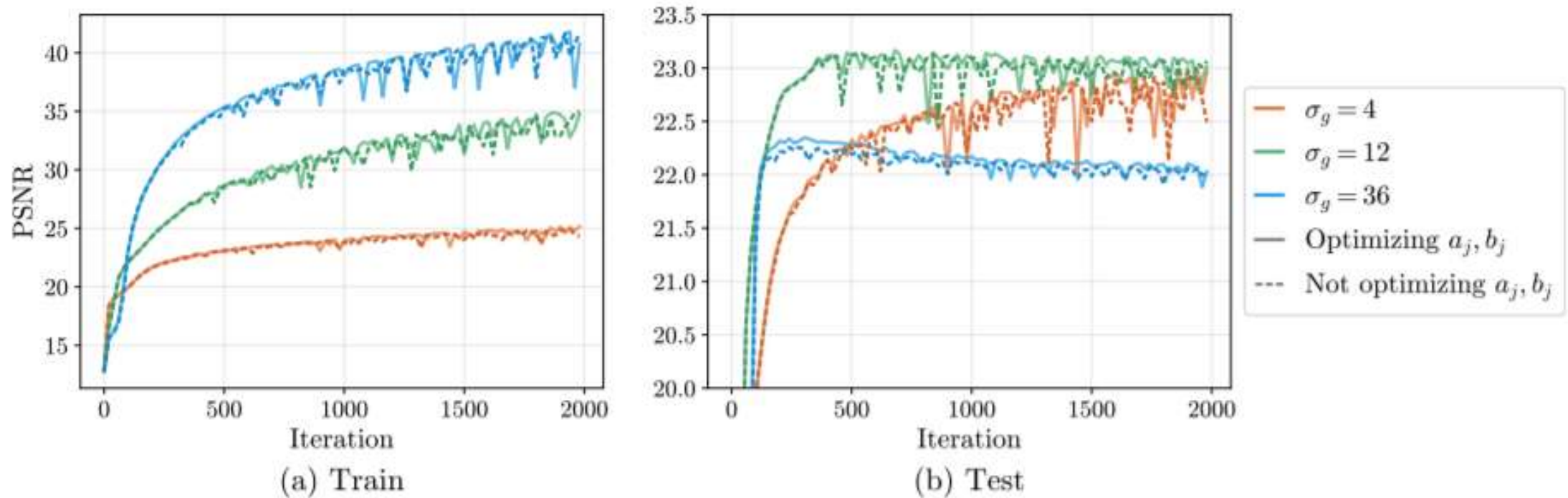
(a) NTK Fourier spectrum with basic mapping



(b) NTK Fourier spectrum with basic mapping and an additional frequency

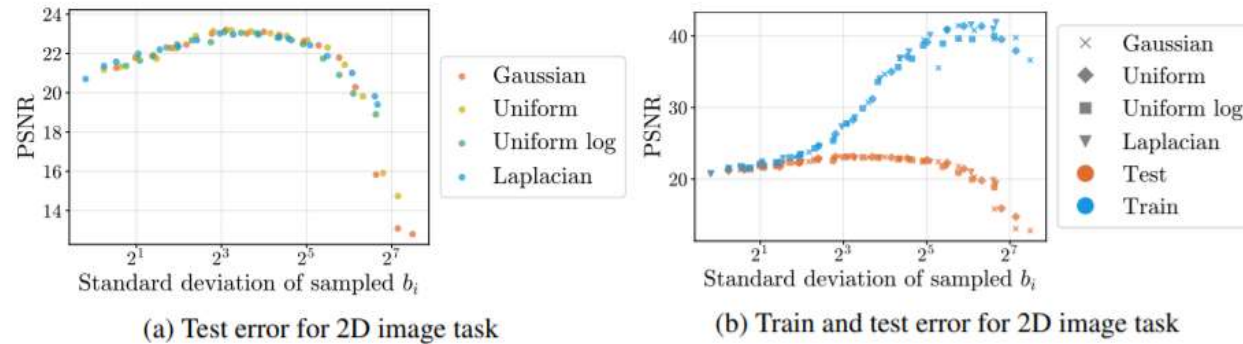
Более глубокие сети нуждаются в меньшем количестве функций Фурье, чем мелкие сети
Это объясняется эффектом расширения частот

Градиентный спуск не оптимизирует преобразование Фурье



Совместная оптимизация параметров a и b не улучшает производительность

Недо- и переобучение



Низкие значения σ не подходят, потому что приводят к чрезмерной сглаженной интерполяции, и большие значения σ переобучают, что приводит к зашумленной интерполяции.

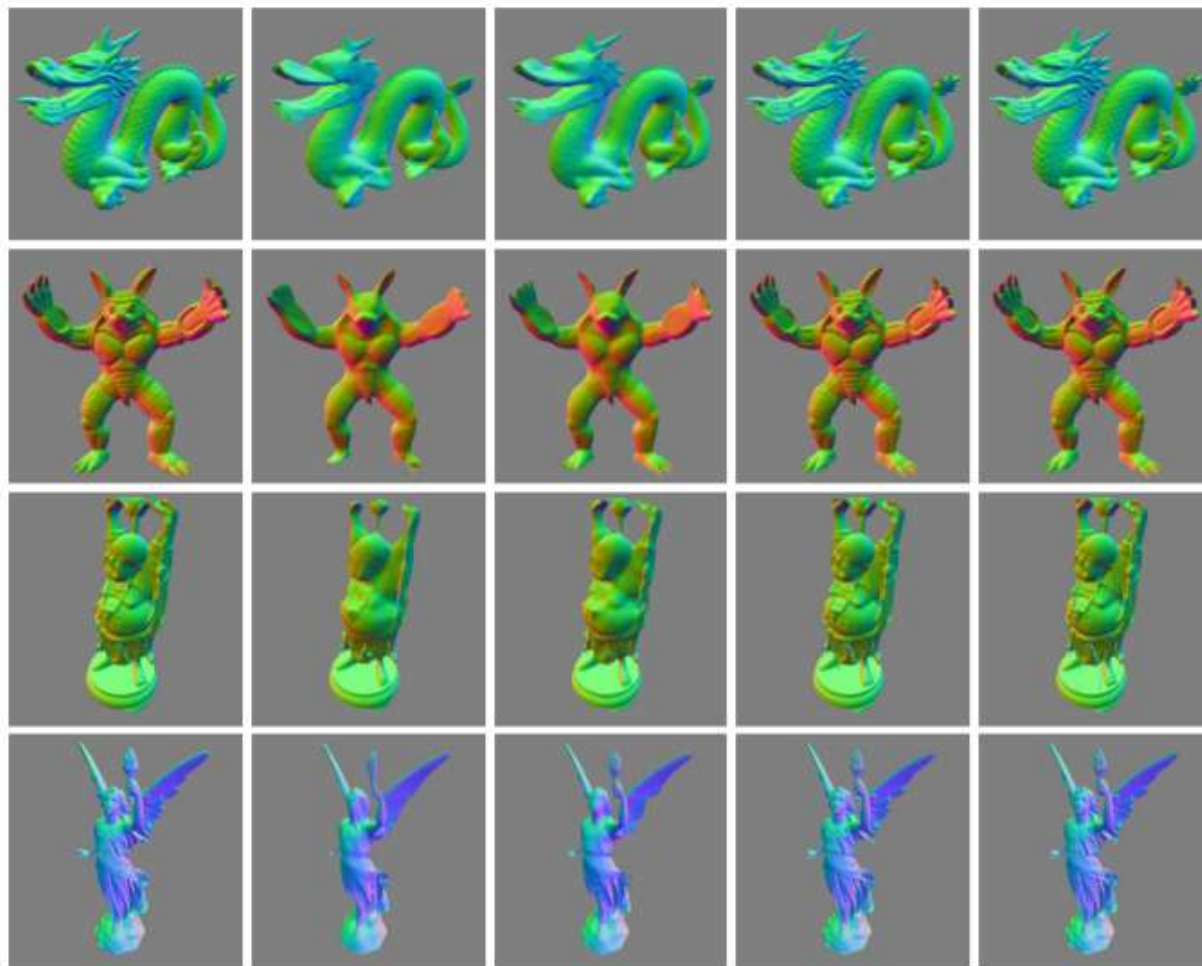
2D image regression



	Natural	Text
No mapping	19.32 ± 2.48	18.40 ± 2.23
Basic	21.71 ± 2.71	20.48 ± 1.96
Positional enc.	24.95 ± 3.72	27.57 ± 3.07
Gaussian	25.57 ± 4.19	30.47 ± 2.11



3D shape regression



(a) Ground Truth

(b) No mapping

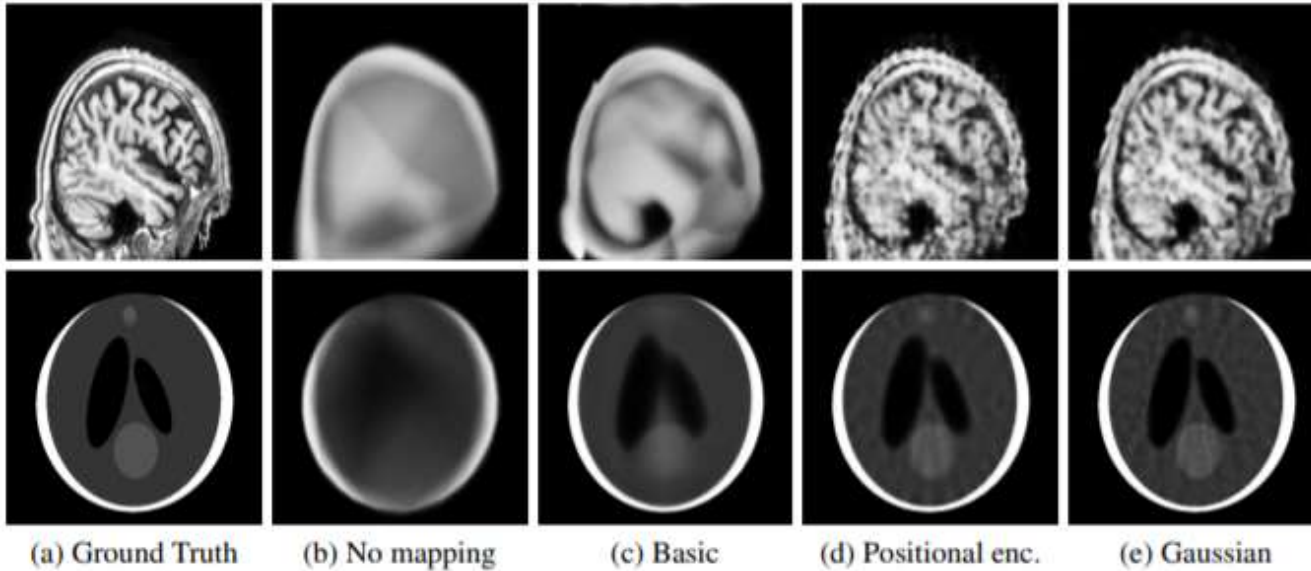
(c) Basic

(d) Positional enc.

(e) Gaussian

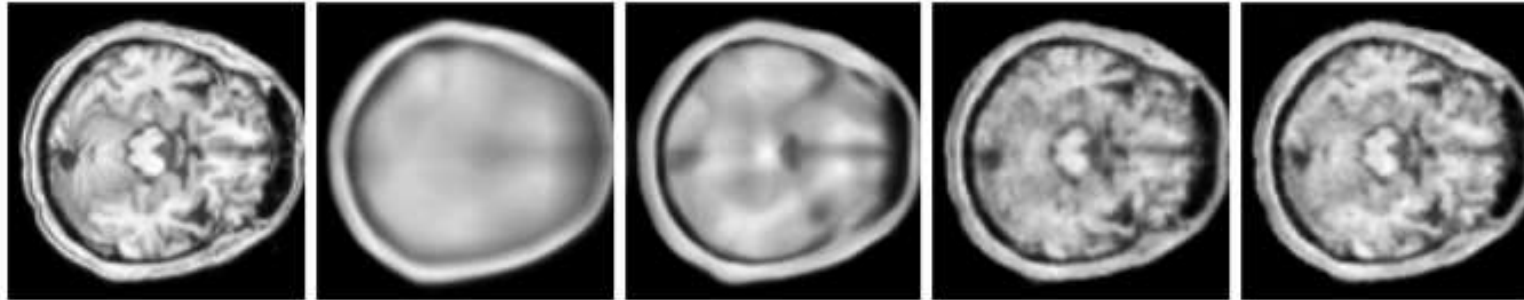
	Uniform points	Boundary points
No mapping	0.959 ± 0.006	0.864 ± 0.014
Basic	0.966 ± 0.007	0.892 ± 0.017
Positional enc.	0.987 ± 0.005	0.960 ± 0.011
Gaussian	0.988 ± 0.007	0.973 ± 0.010

2D computed tomography



	Shepp	ATLAS
No mapping	16.75 ± 3.64	15.44 ± 1.28
Basic	23.31 ± 4.66	16.95 ± 0.72
Positional enc.	26.89 ± 1.46	19.55 ± 1.09
Gaussian	28.33 ± 1.15	19.88 ± 1.23

3D magnetic resonance imaging



(a) Ground Truth

(b) No mapping

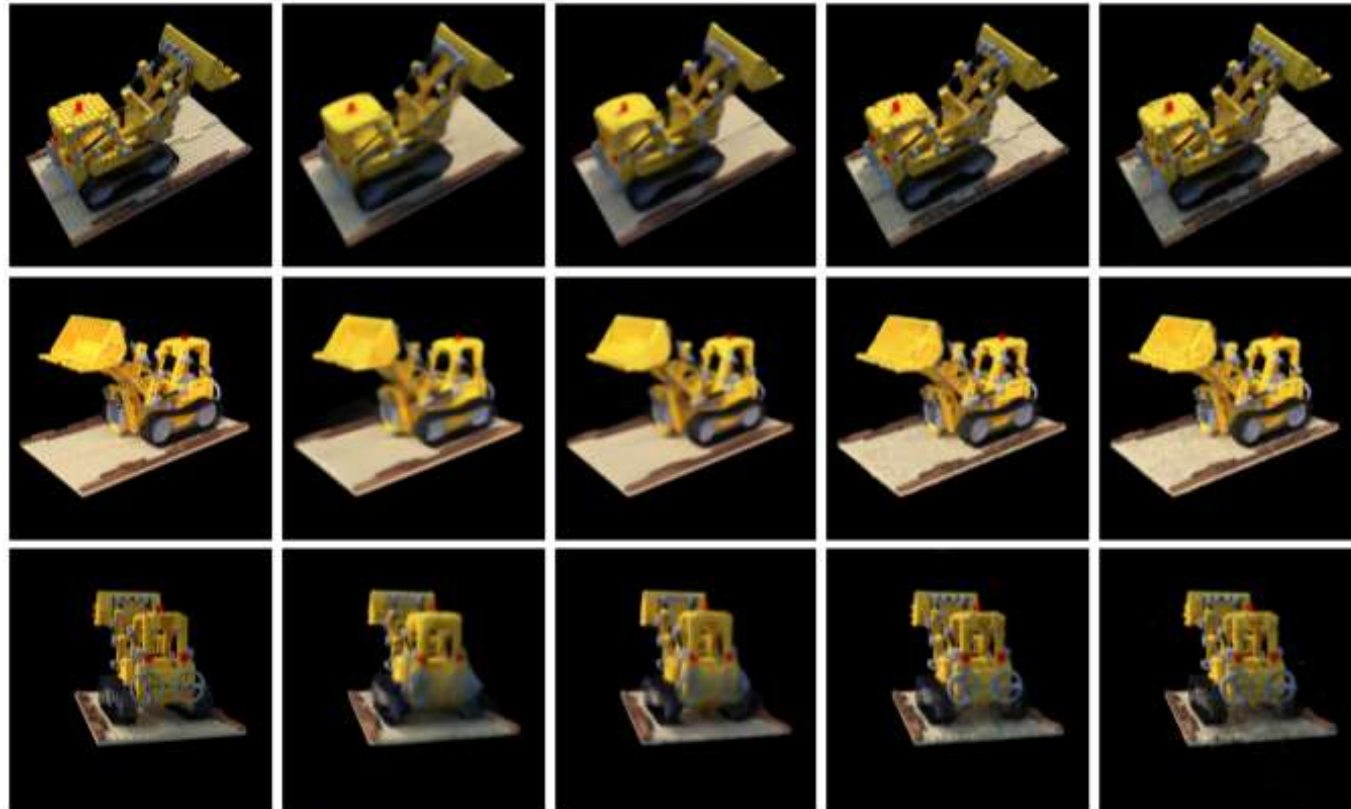
(c) Basic

(d) Positional enc.

(e) Gaussian

	ATLAS
No mapping	26.14 ± 1.45
Basic	28.58 ± 2.45
Positional enc.	32.23 ± 3.08
Gaussian	34.51 ± 2.72

3D inverse rendering for view synthesis



(a) Ground Truth

(b) No mapping

(c) Basic

(d) Positional enc.

(e) Gaussian

	3D NeRF
No mapping	22.41 ± 0.92
Basic	23.16 ± 0.90
Original pos. enc.	24.81 ± 0.88
Positional enc.	25.28 ± 0.83
Gaussian	25.48 ± 0.89
Original pos. enc. (axis-aligned)	25.60 ± 0.76
Positional enc. (axis-aligned)	26.27 ± 0.91

Итого

Преобразование признаков Фурье может сделать основанные на координатах MLP более подходящими для моделирования функций в малых измерениях, тем самым преодолевая спектральное смещение, присущее основанным на координатах MLP

Вопросы

- 1. Какие варианты отображений были рассмотрены и какой из них показал лучшие результаты?
- 2. В чем особенность спектрального смещения при обучении нейронной сети?
- 3. Как происходит тюнинг модели на практике?