

# Position Informed Convolution for Multi-Agent Curve Detection

Бабкин Пётр

Научный руководитель: к.т.н. Никита Шубин

Кафедра интеллектуальных систем ФПМИ МФТИ

Специализация: Интеллектуальный анализ данных

Направление: 03.04.01 Прикладные математика и физика

2024

# План работы

## Цель исследования:

Предложить метод детектирования кривых на плоскости с помощью мультиагентного подхода

## Проблема:

Агрегация данных от разных агентов должна проводиться с учетом их положения друг относительно друга

## Предлагаемый метод:

Использовать слой, в котором будут учитываться индекс, от которого приходит сообщение

## Постановка задачи

Требуется максимизировать правдоподобие с учетом результатов работы агентов

$$\Psi(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n \psi_i(\mathbf{y}_i|\mathbf{x}_i, \mathbf{x}_j, j \in \varepsilon(i)) \rightarrow \max_{\mathbf{y}},$$

где  $\mathbf{x}_i$  – результат работы  $i$ -ого агента, а  $\mathbf{x}$  – их конкатенация,  $\varepsilon(i)$  – соседи  $i$ -й вершины

Зададим параметрическую функцию, получающую наиболее правдоподобный ответ и получим задачу оптимизации

$$\mathbb{E}_{\mathbf{x}, \mathbf{y} \in \mathcal{D}}[-\log \Psi(\mathbf{y}|\mathbf{x})] = \mathbb{E}_{\mathbf{x}, \mathbf{y} \in \mathcal{D}} \left[ \sum_i \mathcal{L}(f_{\theta}(\mathbf{x}_i, \mathbf{x}_j, j \in \varepsilon(i)), \mathbf{y}_i) \right] \rightarrow \min_{\theta},$$

# Глубокое преобразование Хафа

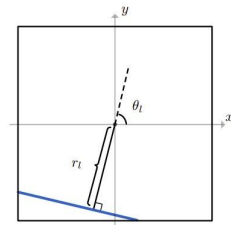
Проводит преобразование изображения в пространство параметров прямой

$$\mathcal{H} : \mathbb{R}^{C \times H \times W} \rightarrow \mathbb{R}^{C \times \Theta \times R},$$

где  $C$  – число каналов,  $H, W$  – размер изображения,  $\Theta, R$  – величина квантизации

Если  $\mathcal{H}(X) = Y$ , то

$$Y(\theta_l, r_l) = \sum_{i \in I} X(i)$$



Параметризация  
прямой

## Учет позиции источника сообщения

Для обычной свертки результат не зависит от порядка присылаемых сигналов

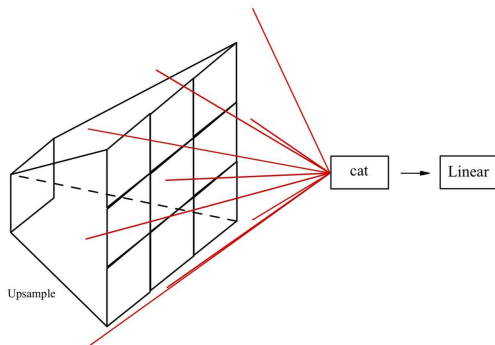
$$\text{Conv}_{\theta}(\mathbf{x}_1, \dots, \mathbf{x}_{|\varepsilon(i)|+1}) = \underbrace{\varphi_{\theta_1}(\mathbf{x}_1) + \dots + \varphi_{\theta_{|\varepsilon(i)|+1}}(\mathbf{x}_{|\varepsilon(i)|+1})}_{\text{Перестановочно}}$$

Требуется функция со следующим свойством

$$\begin{aligned} f_{\theta}(\mathbf{x}_1, \dots, \mathbf{x}_{|\varepsilon(i)|+1}) &= g_{\theta}(\varphi_{\theta_1}(\mathbf{x}_1), \dots, \varphi_{\theta_{|\varepsilon(i)|+1}}(\mathbf{x}_{|\varepsilon(i)|+1})) \neq \\ &\neq g_{\theta}(\varphi_{\theta_{\pi(1)}}(\mathbf{x}_{\pi(1)}), \dots, \varphi_{\theta_{\pi(|\varepsilon(i)|+1)}}(\mathbf{x}_{\pi(|\varepsilon(i)|+1)})) \end{aligned}$$

## Свертка с учетом позиции

Последовательное применение ConvTranspose, конкатенации и линейного слоя позволяет получить выполнение условия с предыдущего слайда



Positional Informed Convolution

# Разработанный алгоритм

Разработан алгоритм, позволяющий проводить end-to-end обучение

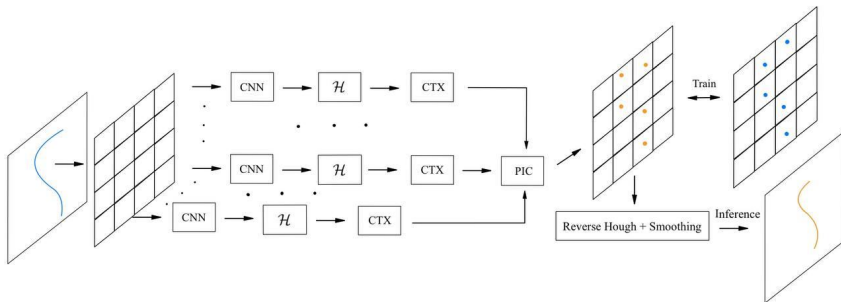


Схема работы алгоритма

CNN – сверточная сеть,  $\mathcal{H}$  – глубокое преобразование Хафа, CTX – детектор линий, PIC – Свертка с учетом позиции

# Цели вычислительного эксперимента

## Генерация данных

- ▶ Сгенерировать датасет кривых с помощью сглаживания сплайнами

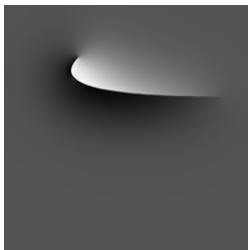
## Базовый эксперимент проверки работы PIC

- ▶ Проверить эффективность разработанного слоя в базовой задаче на датасете MNIST, сравнившись с обычной сверткой

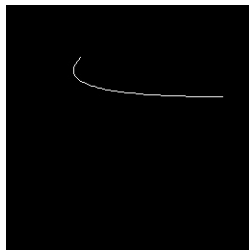


# Результаты эксперимента

Сгенерирован датасет кривых с помощью сглаживания сплайнами



Пример



Разметка

Для свертки с учетом позиции показано, что она работает не хуже обычной свертки на MNIST

## Достигнутые результаты:

- ▶ Разработан новый слой, учитывающий направление, по которому приходит информация
- ▶ Сгенерирован датасет кривых на плоскости
- ▶ Предложен метод детекции кривых, использующий мультиагентный подход

## Следующие шаги:

- ▶ Реализовать пайплайн целиком
- ▶ Провести набор вычислительных экспериментов