

Position Informed Convolution for Multi-Agent Computer Vision

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 мульти-агентный метод для задач компьютерного зрения

Проблема:

При обычной свертке теряется информация о положении агентов 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

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

Учет позиции источника сообщения

Для обычной свертки результат не зависит от порядка присылаемых сигналов

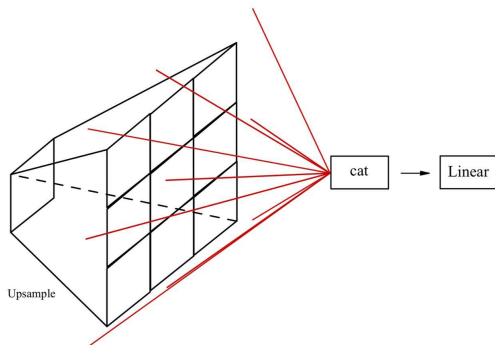
$$\text{Conv}_{\theta}(\mathbf{x}_1, \dots, \mathbf{x}_{|\varepsilon(i)|+1}) = \underbrace{\varphi_{\theta_1}(\mathbf{x}_1) + \dots + \varphi_{\theta_{|\varepsilon(i)|+1}}(\mathbf{x}_{|\varepsilon(i)|+1})}_{\text{Перестановочно}}$$

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

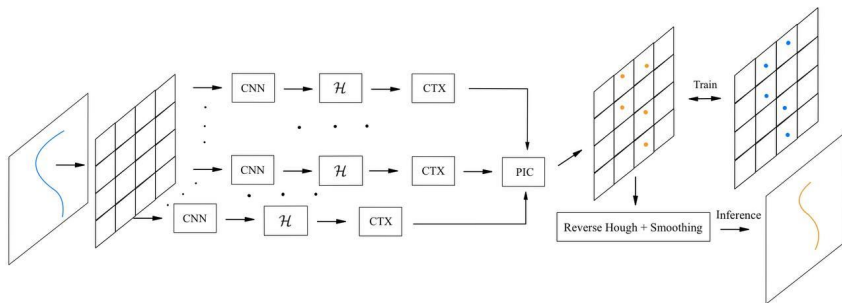


Схема работы алгоритма

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

Цели вычислительного эксперимента

Проверка работоспособности слоя

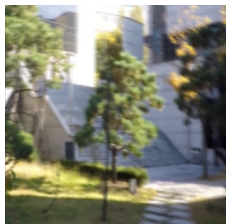
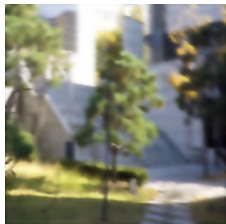
- ▶ Проверить предложенный слой на разных задачах
- ▶ Проверить предложенный алгоритм для детекции кривых

Изучение свойств слоя

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

Результаты экспериментов

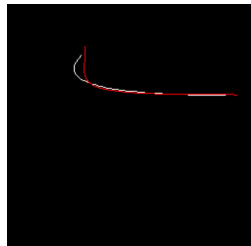
Проведены эксперименты на разных задачах



Устранение
размытия



Шумоподавление



Кривые

Достигнутые результаты:

- ▶ Разработан новый слой, учитывающий направление, по которому приходит информация
- ▶ Предложен метод детекции кривых, использующий мультиагентный подход
- ▶ Проведены эксперименты по изучению свойств слоя применительно к различным задачам

Следующие шаги:

- ▶ Продолжить изучения свойств слоя и найти подходящую задачу