

Улучшение качества ReID моделей с помощью дополнительной информации

Швейкин Денис Александрович

Выпускная квалификационная работа

Кафедра анализа данных, ФПМИ МФТИ

25.06.2024

- ❶ Введение. Задача Re-Id
- ❷ Анализ подходов к решению задачи Re-Id
- ❸ Система поиска пропавших домашних животных
- ❹ Исследование влияния дополнительной информации в Re-Id
- ❺ Заключение

Задача Person Re-Identification



Рис. 1: Пример изображений из набора PRID2011

Математическая постановка

- 1 Даны два множества - $Query$ и $Search$
- 2 Построить отображение $Query \rightarrow S_{|Search|}$, где S_n - группа перестановок чисел от 1 до n
- 3 То есть $\forall q_i \in Query : q_i \mapsto r_i \in S_{|Search|}$



Рис. 2: Пример ранжированного списка для одного запроса

Метрики качества

Cumulated Matching Characteristics (CMC):

$$CMC@k = \frac{1}{|Query|} \sum_{i=1}^{|Query|} \mathbb{I}[r_i[:k] \cap T_i \neq \emptyset]$$

k - пороговая длина списка

$r_i[:k]$ - первые k элементов списка

T_i - множество истинных объектов

Mean Average Precision (mAP):

$$AP_i = \frac{1}{|T_i|} \sum_{k=1}^{|Search|} P_i(k)$$

$P_i(k)$ - доля истинных объектов
среди первых k элементов списка

Общий пайплайн решения задачи

- 1 Обучение представлений признаков
- 2 Метрическое обучение
- 3 Оптимизация ранжирования

Обучение представлений признаков

- 1 Глобальное представление
- 2 Локальное представление
- 3 Использование дополнительной информации
- 4 Обработка видео как последовательности

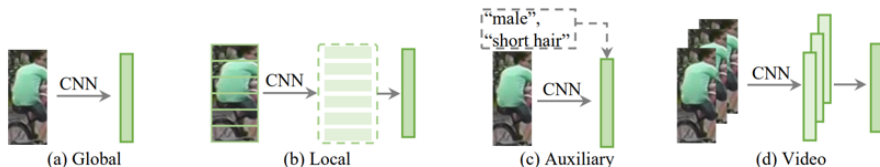


Рис. 3: Варианты работы системы на этапе обучения представлению признаков

Метрическое обучение

Примеры функций потерь метрического обучения:

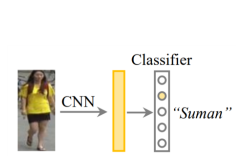


Рис. 4: Identity-loss

$$\mathcal{L}_{id} = -\frac{1}{n} \sum_{i=1}^n \log(p(y_i | x_i))$$

x_i - эмбединг
 y_i - идентификатор
 $p(y_i | x_i)$ - предсказанная
 вероятность

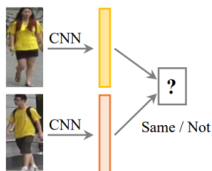


Рис. 5: Verification-loss

$$\mathcal{L}_{very}(i, j) = -\delta_{ij} \log(p(\delta_{ij} | x_{ij})) - \\ -(1 - \delta_{ij}) \log(1 - p(\delta_{ij} | x_{ij}))$$

x_{ij} - общий эмбединг
 δ_{ij} - индикатор совпадения id
 $p(\delta_{ij} | x_{ij})$ - предсказанная
 вероятность

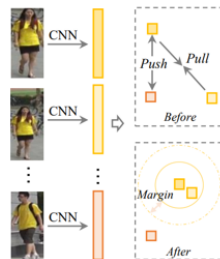


Рис. 6: Triplet-loss

$$\mathcal{L}_{tri}(a, p, n) = \max(p + d_{ap} - d_{an}, 0)$$

d_{ap} - расстояние от якоря до
 положительного объекта
 d_{an} - расстояние от якоря до
 отрицательного объекта
 ρ - порог

Оптимизация ранжирования

Ре-ранкинг:

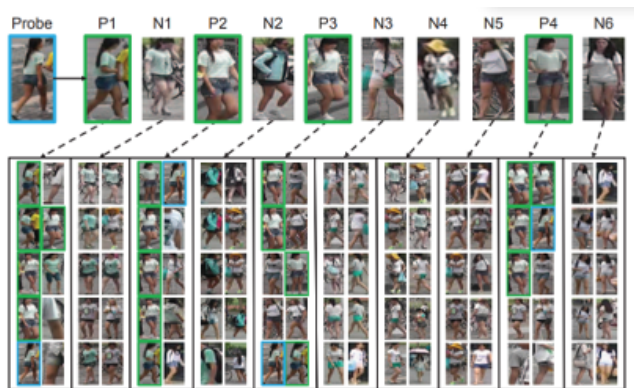


Рис. 7: k-reciprocal re-ranking

Система поиска пропавших домашних животных

- 1 Поиск по базе на основе визуального соответствия
- 2 Уточнение с помощью дополнительной информации - текстового описания, классификации

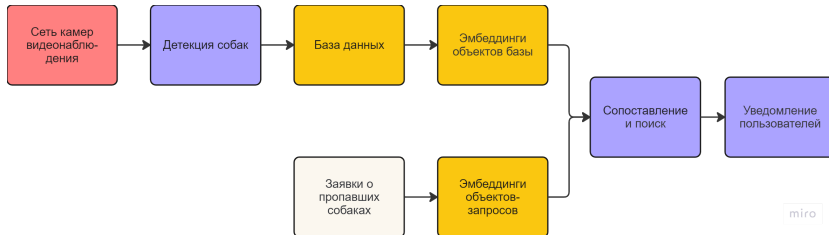


Рис. 8: Схема работы системы поиска

Датасеты

- 1 Был составлен собственный датасет из 50 000 кадров
- 2 Имеются категории классификации: цвет, размер, длина лап, форма морды, тип ушей, тип шерсти



Рис. 9: Пример изображений из датасета

Решение с помощью классификации

С помощью модели Multi-Task классификации получены результаты:

Таблица 1: Качество классификации собак по 6 признакам

Признак	Balanced Accuracy, %	ROC AUC, %
Цвет	72,2	97,2
Размер	73,7	91,2
Длина лап	86,0	94,4
Тип ушей	69,4	92,2
Тип морды	82,1	91,3
Тип шерсти	74,1	95,0

Решение с помощью метрического обучения

В режиме Re-Id позволяет работать метрическая модель Unicom:

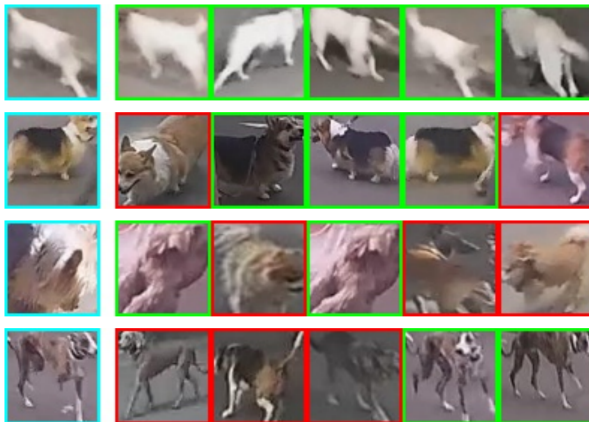


Рис. 10: Ранжированные списки, даваемые моделью метрического обучения

Анализ результатов

- 1 Метрическое обучение показывает визуально хорошие результаты
- 2 В некоторых случаях возникает необходимость дополнительной информации



Рис. 11: Пример сложного случая. Слева - самоед, справа - чихуахуа

Метод CLIP-ReID

1 Первая стадия обучения:

$$\mathcal{L}_{\text{t2i}}(y_i) = \frac{-1}{|P(y_i)|} \sum_{p \in P(y_i)} \log \frac{\exp(\text{sim}(V_p, T_{y_i}))}{\sum_{j=1}^B \exp(\text{sim}(V_j, T_{y_i}))}$$

$$\mathcal{L}_{\text{i2t}}(i) = -\log \frac{\exp(\text{sim}(V_i, T_i))}{\sum_{j=1}^B \exp(\text{sim}(V_i, T_j))}$$

V_i, T_i - визуальный и текстовый эмбединги

$P(y_i)$ - множество кадров с тем же id

2 Вторая стадия обучения:

$$\mathcal{L}_{\text{tri}} = \max(d_p - d_n + \rho, 0)$$

$$\mathcal{L}_{\text{id}} = \sum_{k=1}^N q_k \log(p_k)$$

$$\mathcal{L}_{\text{i2tce}}(i) = \sum_{k=1}^N -q_k \log \frac{\exp(\text{sim}(V_i, T_{y_k}))}{\sum_{j_k=1}^N \exp(\text{sim}(V_i, T_{y_j}))}$$

d_p - расстояние от якоря до позитивного примера

d_n - расстояние от якоря до негативного примера

q_k - сглаженный индикатор совпадения id

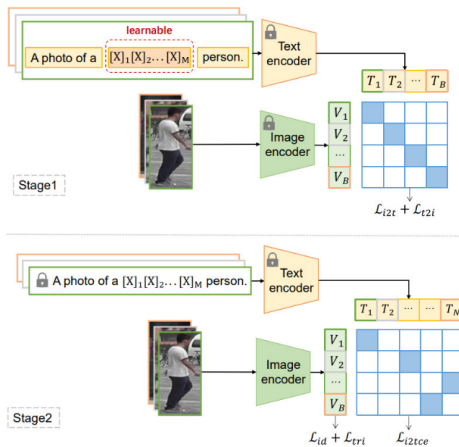


Рис. 12: Схема работы метода CLIP-ReID

Предложенные модификации

Использование информации об атрибутах:

- 1 Составление промптов на основе атрибутов
- 2 Label smoothing на основе атрибутов на первой стадии
- 3 Label smoothing на основе атрибутов на обеих стадиях

Использование информации о ключевых точках:

- 1 Регуляризация обучаемым энкодером ключевых точек
- 2 Регуляризация эмбедингом модели pose estimation

Сравнение качества методов

Таблица 2: Сравнение результатов экспериментов

Тип метода	Эксперимент	#эпох	mAP	CMC@1
CLIP-ReID	Настройка визуальной модели CLIP без первой стадии	120	88,1	94,7
	CLIP-ReID с обеими стадиями	120	89,8	95,7
	Повторение процесса обучения CLIP-ReID 2 раза	120	89,7	95,4
Модификации	Промпты по атрибутам вместо первой стадии	120	85,5	94,2
	Label smoothing по атрибутам на первой стадии	120	88,9	95,6
	Label smoothing по атрибутам на обеих стадиях	120	16,7	29,8
	Регуляризация эмбедингом ключевых точек, без нормализации	120	87,5	94,3
	Регуляризация эмбедингом ключевых точек, нормировка	360	88,1	94,9
	Регуляризация эмбедингом ключевых точек, нормализация средней нормой вектора	480	88,9	95,3
	Регуляризация эмбедингом HRNet, нормализация средней нормой вектора	480	88,1	95,1

Анализ обученных моделей

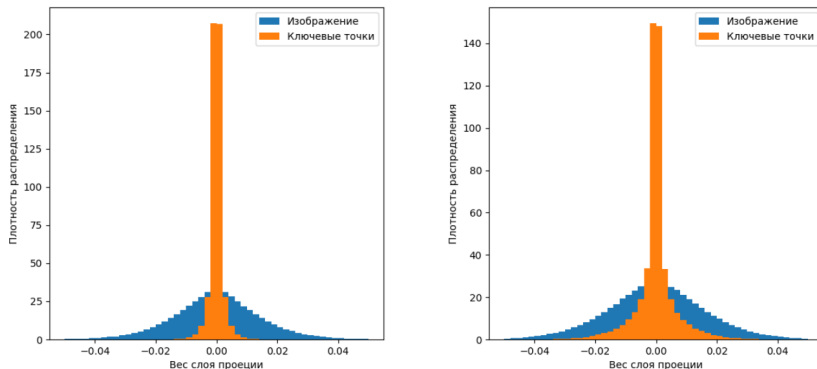


Рис. 13: Распределение весов проектора в общее пространство. Слева - случай обучаемого эмбединга, справа - случай эмбединга HRNet

Заклучение

Таким образом, в данной работе рассмотрена задача Person Re-Identification, ее научная теория, практические аспекты и приложения:

- ➊ Приведена математическая постановка задачи
- ➋ Приведен обширный анализ методов ее решения и областей их применимости
- ➌ Реализована система поиска пропавших домашних животных - практическое приложение технологий Re-Id
- ➍ Выявлены пути развития применяемых методов и мотивация использования этих путей
- ➎ Приведен анализ базового решения адресованной проблемы, предложены его модификации
- ➏ Проведены эксперименты по внедрению данных модификаций и анализ их результатов

Перспективы

Пути развития:

- 1 Собрать Re-Id-датасет для собак
- 2 Применить подход с ключевыми точками для собак - их размер и форма тела более разнообразны
- 3 Воспроизвести процесс обучения CLIP с дополнительной ветвью

Спасибо за внимание!