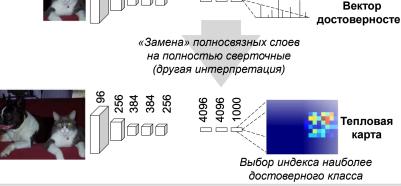
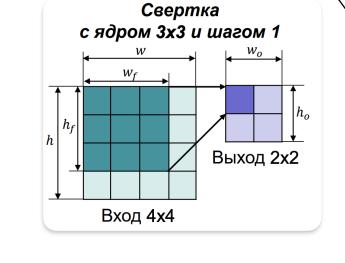
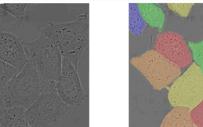


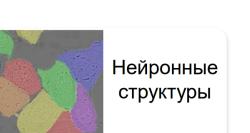
Сверточная нейронная сеть (FCN), SegNet, UNet, Enet , PSPnet, другие











Постановка задачи

Задача семантической сегментации (англ. semantic segmentation) — задача, в которой на вход модели подаётся изображение, а на выходе для каждого пикселя является метка принадлежности этого пикселя к определённой категории. Например, если в исходном изображении человек переходит дорогу, то для каждого пикселя необходимо вывести, является ли этот пиксель частью человеческого тела, профиля дороги, знака дорожного движения, неба, или какого-то другого типа. Существенный недостаток применения одной лишь семантической сегментации относительно задач, связанных с распознаванием объектов — маркировка пикселей по принадлежности только к типу объекта, что не создаёт различия

Метрики

между объектами как таковыми.

Попиксельная точность (pixel ассuracy) ,Средняя попиксельная точность по классам наблюдаемых (Jaccard index), Индекс Дайса (Dice

Возможные проблем и их решения

Проблема применения глубоких моделей для семантической сегментации При решении задачи семантической сегментации изображений на выходе модели должен быть трехмерный тензор с элементами, отвечающими достоверности принадлежности каждого пикселя к

Каким образом обеспечить на выходе тензор, у которого пространственные размерности совпадают с разрешением входного изображения?

Интерполяция – Построение архитектуры «кодировщик-декодировщик» (encoder-decoder architecture) – Применение графовых вероятностных методов, в частности, условных случайных полей (Conditional Random Fields, CRF)

Нейросетевые

подходы для

семантической

сегментации

изображений

объектов (mean pixel accuracy over classes), Метрика IoU (Intersection over Union) или индекс Жаккара

index) или F1-score

PA and mPA Попиксельная точность (pixel accuracy) определяется

следующим образом: $Accuracy = \frac{1}{TP + TN + FP + FN}$ где *TP* + *TN* - количество правильно проклассифицированных

пикселей (true positives + true negatives), TP + TN + FP + FN — общее количество пикселей

Попиксельная точность не является показательной в случае несбалансированности классов Поэтому вводится средняя попиксельная точность, вычисленная для каждого класса в отдельности и усредненная по количеству классов, - средняя попиксельная точность по классам (mean pixel accuracy over classes)

Метрика IoU (Intersection over Union) или индекс Жаккара (Jaccard index)

 $IoU = \frac{T}{TP + FP + FN}$ где *TP* – количество правильно проклассифицированных пикселей (true positives), FP – количество пикселей, которые метод проклассифицировал как принадлежащие классу, но они таковыми не являются (false positives), FN – количество пикселей, которые принадлежат классу, но метод проклассифицировал их как не принадлежащие классу (false negatives)

Индекс Дайса (Dice index) или F1-score определяется следующим образом:

Dice index

 $DICE = \frac{1}{2 \cdot TP + FP + FN}$ Индекс Дайса отличается от индекса Жаккара одним коэффициентом Указанные индексы связаны соотношениями:

оба показателя, достаточно вычислять какой-то один

 $IoU = \frac{DICE}{2 - IoU}, \qquad DICE = \frac{2 \cdot 100}{1 + IoU}$ Как следствие, не имеет смысла одновременно определять показатель

Методы описания содержания Разрешение семантики запроса Цвет, Текстура, Форма **Локальные фичи** описывают визуальные паттерны или **Global Features** структуры, идентифицируемые в небольших группах пикселей. Например, края, точки и различные фрагменты изображения. Color Moment Gabor Filter Fourier Descriptor Markov Random Field

Edge Histogram Descript



Автоматическое создание текстовых аннотаций для широкого спектра изображений невозможно.

Возможные проблем

и их решения

Добавление комментариев к изображениям вручную трудоемкая и дорогостоящая задача для больших баз данных изображений.

Ручные аннотации часто бывают субъективными, контекстно-зависимыми и неполными.

Google, Яндекс и другие используют текстовый поиск. Результаты не идеальны. Однако сейчас намного лучше, чем пару лет назад!

Проблема поиска на основе пиксельного содержания в большинстве ситуаций пока не имеет реализованного

Извлечение признаков низкого уровня — как представить изображение в компактном и описательном

- Как сравнивать признаки и, следовательно, изображения?
- Многомерное индексирование как индексировать огромные объемы многомерных данных?
- Визуальный интерфейс для просмотра изображений Как визуализировать результаты?



Постановка задачи

similarity metrics distance measures

Мера расстояния обычно количественно определяет несходство двух векторов признаков. Мы вычисляем его как расстояние между двумя векторами в некотором метрическом пространстве

Метрики

Manhattan distance, Mahalanobis distance, Histogram Intersection Distance (HID)

похожее на изображение запроса. Например, косинусное расстояние (cosine distance) измеряет угол между двумя векторами

признаков.

метрика подобия

количественно определяет

сходство между двумя

векторами признаков. Таким

образом, это работает

противоположно метрикам

расстояния: наиболее

значимое значение

показывает изображение,