

СЕГМЕНТАЦИИ АВТОМОБИЛЬНЫХ НОМЕРНЫХ ЗНАКОВ НА ИЗОБРАЖЕНИЯХ

В.С. Адаричев, Ю.А. Иванова
Томский политехнический университет
E-mail: shoutsoul97@mail.ru, jbolotova@tpu.ru

Аннотация

В работе представляется применение архитектуры U-Net для сегментации автомобильных номерных знаков на изображениях. Для оценки качества сегментации применяются различные метрики, такие как: ошибка по модулю, квадратичная ошибка, процентная точность, визуальная оценка.

Введение

U-Net — это сверточная нейронная сеть (СНС), которая была создана в 2015 году для сегментации биомедицинских изображений в отделе Computer Science Фрайбургского университета [1, 2].

U-Net считается одной из стандартных архитектур для задач сегментации изображений, когда нужно не только определить класс изображения целиком, но и сегментировать его области по классу, т. е. создать маску, которая будет разделять изображение на несколько классов. Архитектура состоит из сжимающего пути для захвата контекста и симметричного расширяющегося пути, который позволяет осуществить точную локализацию.

В данной работе исследуется качество сегментации изображений нейронной сетью U-Net.

Описание выборки данных

В данной работе эксперименты проводились на данных, сформированных вручную из изображений реальных сцен, на которых находятся автомобильные номерные знаки.

Выборка содержит 48 изображений размером 540x480 пикселей. Выборка была разделена на 2 части. Первая часть содержит 35 изображений, что составляет 73% от всей выборки, вторая содержит 13 изображений, что составляет 27% от всей выборки. Первая часть используется для обучения нейронной сети, вторая для тестирования.

На данных изображениях необходимо сегментировать номерные знаки автомобилей.

Пример входного изображения можно увидеть в левой части изображения ниже (рис 1), образец наложения сегментированного номерного знака на исходное изображение в правой части изображения (рис 1).



Рис. 1. Пример входного изображения(слева), наложение сегментации на исходное изображение (справа)

Для каждого изображения создана маска, покрывающая весь номерной знак, которая использовалась при обучении и тестировании (рис 2).

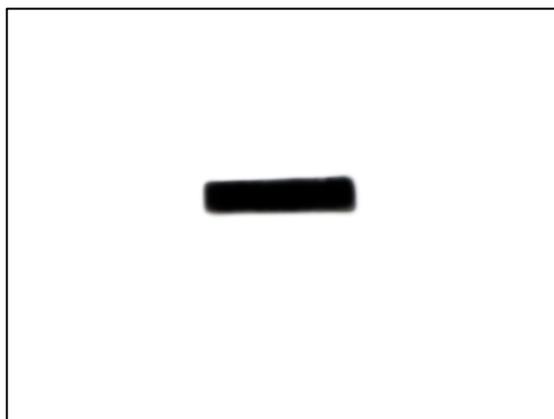


Рис. 2. Пример маски, используемой при обучении и тестировании.

Топология исследуемой нейронной сети

Архитектура сети представляет собой полносвязную сверточную сеть [3], модифицированную так, чтобы она могла работать с меньшим количеством примеров (обучающих образов) и более точно сегментировала изображение.

Сеть содержит сжимающий и расширяющий путь. На каждом шаге мы удваиваем количество каналов признаков.

Сжимающий путь похож на типичную сверточную сеть, он содержит 2 подряд сверточных слоя 3x3, после которых идет слой ReLU и пулинг с функцией максимума 2x2 с шагом 2.

Каждый шаг расширяющего пути содержит слой, обратный пулингу, который расширяет карту признаков, после которого следует свертка 2x2, уменьшающая количество каналов признаков. После идет конкатенация с соответствующим образом обрезанной картой признаков из сжимающего пути и две свертки 3x3, после каждой из которой идет ReLU. Обрезка нужна из-за того, что мы теряем пограничные пиксели в каждой свертке.

На последнем слое свертка 1x1 используется для приведения каждого 64-компонентного вектора признаков до требуемого количества классов. Всего сеть имеет 23 сверточных слоя.

Результаты тестирования

Программа реализована на языке Python, с использованием библиотеки машинного обучения Keras.

Качество сегментации номерного знака сравнивается по следующим метрикам качества: ошибка по модулю (module), квадратичная ошибка (squaring), процентная точность (percent), Intersection over union (IoU) [4]. Где module и squaring – это средняя ошибка сегментации пикселя по модулю и возведенная в квадрат соответственно, percent – это количество верно сегментированных пикселей (в процентах), а IoU показывающий меру площади правильно отмеченных сегментов (отношение площади пересечения к площади объединения).

Проведено обучение и тестирование нейронной сети. Обучение производилось на 20 эпохах, на каждую эпоху подавалось 105 изображений.

Результаты тестирования нейронной сети U-Net представлены в таблице ниже (таблица 1):

Таблица 2. Результаты тестирования U-Net

Выборка	module	squaring	percent	IoU
Обучающая	0.256	0.163	84.3%	0.984
Тестовая	0.122	0.105	79.2%	0.917

Результаты работы представлены на изображениях ниже (рисунки 3,4):



Рис. 3. Результат наложения сегментации на исходное изображение.



Рис. 4. Результат наложения сегментации на исходное изображение.

Заключение

По полученным результатам можно сделать вывод, что нейронная сеть U-Net обучилась сегментировать номерные знаки автомобилей с некоторыми погрешностями.

Устранить недостатки можно путем расширения обучающей выборки более разнообразными изображениями и увеличить время обучения нейронной сети.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-08-00977 А и в рамках Программы повышения конкурентоспособности ТПУ.

Список использованных источников:

6. Olaf Ronneberger, Philipp Fischer, Thomas Brox U-Net: Convolutional Networks for Biomedical Image Segmentation (2015) / arXiv:1505.04597v1 [cs.CV] 2015
7. U-Net. [Электронный ресурс] / Статьи и информация о Deep Learning — URL: <http://deeplearning.net/tutorial/unet.html> (дата обращения 09.09.2019)
8. J. Long, E. Shelhamer, T. Darrell: Fully convolutional networks for semantic segmentation (2014), arXiv:1411.4038v2 [cs.CV] 2015
9. Evaluating image segmentation models [Электронный ресурс]. / Блог Jeremy Jordan — URL: <https://www.jeremyjordan.me/evaluating-image-segmentation-models/> (дата обращения 10.09.2019)