



И.В. Овчинников

РАЗРАБОТКА АЛГОРИТМОВ И ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ АВТОМАТИЧЕСКОГО ПОСТРОЕНИЯ ЦИФРОВЫХ АДРЕСНЫХ ПЛАНОВ НА ОСНОВЕ НЕПОЛНЫХ ПРОСТРАНСТВЕННЫХ ДАННЫХ

(Самарский университет)

Введение

Адресный план – специальный цифровой план территории. Адресный план содержит информацию двух видов: геометрия объектов (информация о пространственном расположении объектов) и их семантика (атрибутивное описание).

Ввиду быстрого развития городской инфраструктуры, картографические данные имеют свойство очень быстро устаревать. При этом картографирование является очень трудоемкой задачей, которая требует много времени при выполнении вручную. Поэтому актуальной является задача автоматического построения цифровых адресных планов на основе неполных пространственных данных.

Семантическая сегментация снимков полученных в результате дистанционного зондирования земли

Семантическая сегментация изображений заключается в выделении на изображении локальных областей (сегментов), соответствующих различным классам объектов [1]. С помощью фреймворка машинного обучения PyTorch [2] было обучено несколько моделей для семантической сегментации изображения архитектур U-Net [3] и DeepLabV3+ [4] с предобученными энкодерами (кодировщиками) ResNet [5] и EfficientNet [6]. Для обучения моделей использовался датасет с разметкой строений в городе Самара. Изображения были разделены на тренировочную (5000 шт.), валидационную (998 шт.) и тестовую (20 шт.) выборки. Пример изображения из датасета приведен на рисунке 1. Результаты обучения моделей нейронных сетей приведены в таблице 1.

Таблица 1 – Результаты обучения полносверточных нейронных сетей

| Модель | Предобученный энкодер (ImageNet) | Количество эпох обучения | На тренировочной выборке | |
|------------|----------------------------------|--------------------------|--------------------------|----------------|
| | | | Dice loss | Accuracy (IoU) |
| U-Net | – | 30 | 0,264 | 0,623 |
| | efficientnet-b3 | 30 | 0,013 | 0,975 |
| | resnet34 | 30 | 0,079 | 0,858 |
| DeepLabV3+ | resnet101 | 30 | 0,027 | 0,949 |

Из таблицы 1 видно, что наилучшую точность на тестовой выборке показала модель U-Net (efficientnet-b3). Поэтому данная модель используется для распознавания зданий на изображениях.



Контур на полученных бинарных изображениях с сегментированными зданиями аппроксимируются с помощью алгоритма Рамера-Дугласа-Пекера [7] и преобразуются в векторный слой. Данный векторный слой является одним из источников неполных пространственных данных.



Рис. 1. Результат применения нейросетевого алгоритма (а – исходное изображение, б – маска изображения, в – результат семантической сегментации)

Описание алгоритма построения цифровых адресных планов

На вход алгоритма построения объединенного адресного плана (ОАП) подаются данные открытых картографических источников, которые хранятся в структурированном виде в базе данных. В качестве одного из источников используются данные, полученные в результате семантической сегментации изображений. Алгоритм включает 3 этапа:

1. Построение попарных связей между источниками.

На данном этапе происходит автоматическое попарное связывание объектов между выбранными парами источников. Связь – это утверждение, что записи из двух таблиц описывают один и тот же объект реального мира (здание) с некоторой степенью достоверности.

Процедура установления связей требует сформулировать признаки и метрику близости объектов. Основные признаки, по которым устанавливаются пары между объектами – это **строковое описание адреса** (мера разности строк по дистанции Левенштейна [8]) и **геометрическая близость** (дистанция между объектами), определяется по формуле:

$$distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2},$$

где (x_1, y_1) , (x_2, y_2) , – координаты после преобразования в Web Mercator.

Общая дистанция находится путем суммирования полученных значений. Нормирование дистанции и подсчет веса осуществляется по следующей формуле:

$$weight = 1 - \frac{distance}{threshold'}$$

где $distance$ – полученная дистанция;

$threshold$ – порог допустимой дистанции для установления связи.

Весовые значения признаков и пороговые значения выбираются экспериментально, по некоторой случайной выборке данных.



2. Нахождения образов одного и того же здания в нескольких источниках по геометрии и семантике.

На данном этапе алгоритма происходит получение кластеров. Кластером будем называть компоненту, которая состоит из образов одного и того же здания из нескольких источников.

Для этого строится взвешенный ориентированный граф G . Вся совокупность образов зданий во всех источниках, для которых на предыдущем этапе попарных связей была установлена хотя бы одна связь, составит множество вершин графа. Все связи между двумя образами составляют множество ребер графа, вес связи – это вес ребра. Граф по структуре получается достаточно разреженным, количество инцидентных ребер не превышает количество обрабатываемых источников, при этом общее число вершин на несколько порядков больше. Поэтому в реализации для хранения графа используется структура данных на основе списков смежности. Таким образом, необходимо найти в графе все связные компоненты, удовлетворяющие ряду условий [9].

Компонента связности - часть графа (подграф), являющаяся связной. Нахождение кластера в компоненте связности начинается со стартовой вершины, которая выбирается произвольно из числа оставшихся в графе, с учетом приоритета источников.

К построенной части кластера (первоначально – только одной вершине) итеративно добавляются наиболее связанные с ней вершины. Рассматриваются вершины, имеющие двустороннюю связь со всей построенной частью кластера, и не принадлежащие источникам, уже задействованным в кластере. Выбирается вершина, суммарный вес связи на которую наибольший, она добавляется в кластер. Процесс продолжается, пока в компоненте не будет по одному экземпляру из каждого источника, которые есть в компоненте связности, либо пока все достижимые от кластера вершины окажутся непригодными для добавления по названному выше условиям.

3. Формирование обобщенного образа здания по нескольким неточным образам.

На предыдущем этапе были построены кластеры, которые содержат образы здания из разных источников, участвующих в связывании. Теперь требуется построить по ним обобщенный образ, наиболее полный и не противоречивый, чем входящие в его состав отдельные.

Параметрами алгоритма является набор характеристик (семантических полей), которые требуется получить в обобщенном кластере. Каждое поле имеет свои характеристики: тип (строковый, числовой или геометрическая форма), способ агрегации (медианное среднее или сумма). В зависимости от характеристики и ее значения происходит получение результирующего значения поля нового образа. Пример формулы для нахождения медианного среднего для геометрии объектов:

$$\text{dist} = \frac{\sum_{i=0}^n \sqrt{(x - x_i)^2 + (y - y_i)^2}}{sW_{in}},$$



где $dist$ – дистанция для каждого объекта;

x, y – координаты поля, для которого считается дистанция;

x_i, y_i – координаты поля, с которым считается дистанция;

n – количество значений для сравнения;

sW_{in} – вес источника.

Результат применения алгоритма связывания

Для оценки качественных и количественных показателей работы программного обеспечения построим объединенный адресный план (ОАП) для территории «Многоэтажная застройка». Это территория внутри города Самары, которая имеет застройку многоэтажными зданиями.

Для построения ОАП использовались 8 источников данных. Схема связывания источников приведена на рисунке 2. Цвет стрелок на схеме показывает вид алгоритма установления попарных связей (зеленый – по иерархическому представлению адреса, синий – по геометрической составляющей).

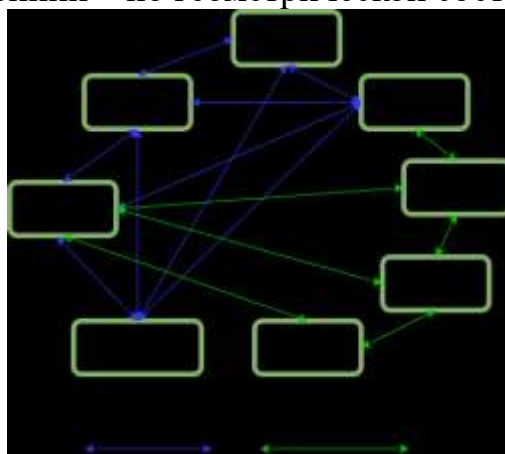


Рис. 2. Схема связывания источников

В таблице 2 приведено сравнение эталонного адресного плана на рассматриваемой территории с ОАП и адресными планами из OpenStreetMap. Метрика сравнения используется та же, что и в построении попарных связей (дистанция Левенштейна и дистанция между геометрией объектов).

Таблица 2 – Результаты построения ОАП и сравнение с эталонным адресным планом

| Адресный план | Количество объектов | Количество объектов в эталонном адресном плане | Количество симметричных связей с эталоном (в обе стороны) | Средний вес симметричных связей |
|---------------|---------------------|--|---|---------------------------------|
| ОАП | 235 | 192 | 185 | 0,999 |
| OpenStreetMap | 178 | | 121 | 0,975 |



Из таблицы 2 можно сделать вывод, что построенный адресный план для рассматриваемой территории наиболее полный и точный относительно адресного плана открытого источника данных OpenStreetMap.

На рисунке 3 приведен итоговый векторный слой построенного ОАП на подложке Google Maps.



Рис. 3. Итоговый векторный слой ОАП

Заключение

В результате выполнения работы были обучены модели полносверточных нейронных сетей на датасете с разметкой строений в городе Самара. Лучший результат (IoU score) на тестовой выборке показала архитектура U-Net с предобученным кодировщиком efficientnet-b3 (0,975). Был построен ОАП для территории «Многоэтажная застройка» города Самары.

Литература

1. Друки, А.А. Семантическая сегментация данных дистанционного зондирования Земли при помощи нейросетевых алгоритмов [Текст] / В.Г. Спицын, Ю.А. Болотова, А.А. Башлыков // Известия Томского политехнического университета. Инжиниринг георесурсов. – 2018. – Т. 329, №1. – С. 59-68.
2. PyTorch documentation [Электронный ресурс] // PyTorch – официальный сайт. – Электрон. дан. – [Б. м.], 2019. – URL: <https://pytorch.org/docs/stable/index.html> (дата обращения: 05.04.2022).
3. Ronneberger, O. U-net: Convolutional networks for biomedical image segmentation [Текст] / O. Ronneberger, P. Fischer, T. Brox // Medical Image Computing and Computer-Assisted. – 2015. – Vol. 3(1). – P. 234-241.
4. Chen, L.-C. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation [Текст] / L.-C Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam // Computer Vision – ECCV. – 2018. – Vol. 7. – P. 805-822.
5. He, K. Deep residual learning for image recognition [Текст] / К. He, X. Zhang, S. Ren, J. Sun // 2016 IEEE Conference on Computer Vision and Pattern Recognition. – 2016. – Vol. 5(2). – P. 235-247.
6. Tan, M. EfficientNet: Rethinking model scaling for convolutional neural networks [Текст] / М. Tan, Q. V Le // In Proceedings of International Conference on Machine Learning (ICML). – 2019. – P.235-247.



7. Алгоритм Рамера — Дугласа — Пекера [Электронный ресурс] // Википедия: свободная энцикл. — Электрон. дан. — [Б. м.], 2019. — URL: https://ru.wikipedia.org/wiki/Алгоритм_Рамера_—_Дугласа_—_Пекера (дата обращения: 05.04.2022).

8. Расстояние Левенштейна [Электронный ресурс] // Википедия: свободная энцикл. — Электрон. дан. — [Б. м.], 2022. — URL: https://ru.wikipedia.org/wiki/Расстояние_Левенштейна (дата обращения: 04.05.2022).

9. Лафоре, Р. Структуры данных и алгоритмы в Java [Текст] / Р. Лафоре. — СПб.: Питер, 2011. — 704 с.

А.В. Озеров, А.С. Маршова

ТЕХНИЧЕСКОЕ ЗРЕНИЕ В СОСТАВЕ СИСТЕМ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ ДВИЖЕНИЕМ ПОЕЗДОВ

(АО «НИИАС»)

Современное развитие техники создает предпосылки для масштабного внедрения систем автоведения поездов (так называемые уровни автоматизации GoA 3/4, согласно МЭК 62267:2009, предполагающие автоматическое движение поезда без участия машиниста). В первую очередь, это связано с появлением специализированных датчиков машинного (технического) зрения, мощных вычислительных средств, обладающих высокой производительностью, и новых алгоритмов обработки данных.

Бортовой комплекс технического зрения, устанавливаемый в голове поезда, состоит из набора датчиков различного назначения и быстродействующих вычислительных средств с большим объемом памяти, обрабатывающих сигналы этих датчиков в режиме реального времени с использованием алгоритмов машинного обучения [1].

Основное назначение бортовой системы технического зрения — своевременное обнаружение и классификация посторонних объектов и принятие решения о соответствующей реакции бортового устройства безопасности (подача звукового сигнала, снижение скорости, экстренное торможение и т.д.) [2].

Такие системы в последние годы активно тестируются на различных полигонах в России и в мире. При этом существует ряд открытых вопросов, касающихся выбора и тестирования средств технического зрения, подготовки и верификации набора данных (датасетов), сертификации автоматической системы управления и обеспечения движения поезда, в состав которой входит техническое зрение, и т.д.

Как правило, бортовые датчики технического зрения включают видеокамеры различного типа, ультразвуковые датчики, радары, лидары. Датчики функционируют в различных диапазонах электромагнитного спектра (рис. 1) и имеют свои преимущества и недостатки, которые могут проявляться при за-