

лектив кафедры интеллектуальных информационных технологий Белорусского государственного университета информатики и радиоэлектроники, в частности Голенкова В. В., Шункевича Д. В., Ковалева М. В.

Список литературы

1. Golovko, V. Neuro-Symbolic Artificial Intelligence: Application for Control the Quality of Product Labeling / V. Golovko, A. Kroshchanka, M. Kovalev, V. Taberko, and D. Ivaniuk // 2020 International Conference on Open Semantic Technologies for Intelligent Systems. – 2020. – P. 81–101.
2. Besold, T. Neural-symbolic learning and reasoning: A survey and interpretation / T. Besold, A. d'Avila Garcez, S. Bader, H. Bowman, P. Domingos, P. Hitzler, K. Kuehnberger, L. Lamb, D. Lowd, P. Lima, L. de Penning, G. Pinkas, H. Poon, G. Zaverucha // <https://arxiv.org/pdf/1711.03902.pdf> (accessed 2021, Sep).
3. Golovko, V. Integration of artificial neural networks and knowledge bases // V. Golovko, V. Golenkov, V. Ivashenko, V. Taberko, S. Ivaniuk, A. Kroshchanka, M. Kovalev // In: Open semantic technologies for designing intelligent systems (OSTIS2018): materials of the International Science and Technology Conf. Minsk, February 15-17, 2018). – Minsk: BSUIR, 2018. – P. 133–145.
4. Golenkov, V. Methods and tools for ensuring compatibility of computer systems / V. Golenkov, N. Guliakina, I. Davydenko, and A. Ereemeev // In Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]. – Minsk: BSUIR, 2019. – P. 25–52.

УДК 004.93'1

НЕЙРОННАЯ СЕТЬ U-NET ДЛЯ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ СЕЛЬСКОХОЗЯЙСТВЕННОЙ РАСТИТЕЛЬНОСТИ

А. А. Дудкин¹, С. В. Шелег²

¹ОИПИ НАН Беларуси, Минск, Беларусь, *e-mail*: ganchenko@lsi.bas-net.by,

²ООО «БСВТ – новые технологии», Минск, Беларусь,
e-mail: sergey.sheleg@hotmail.com

Семантическая сегментация является составной частью решения задач распознавания сельскохозяйственной растительности различного типа, для ее решения широкое применение находят искусственные нейронные сети. Выбор той или иной архитектуры сети зависит от характеристик исходных изображений (размерности и разрешающей способности, в первую очередь) и сложности задачи семантической сегментации (типа предиката, описывающего сегменты, и количество классов разбиений). В данной работе рассматривается задача сегментации изображения земной поверхности на два класса: «растительность» и «почва». По результатам анализа следует отметить перспективность применения для двухклассового распознавания сверточной нейронной сети (СНС) U-Net. Она имеет простую архитектуру и, как следствие, малое потребление ресурсов в работе и при обучении – даже при небольшой обучающей выборке достигаются приемлемые по качеству результаты.

Материалами для исследований явились фотографии экспериментального участка поля, засаженного картофелем, выполненные с высоты 5, 15, 50 и 100 метров (827 снимков разрешением 3474×2314 пикселей). Наблюдению подвергались 3 группы растений: зараженные инфекционной болезнью *alternaria* и

бактериальной болезнью *ergwinia*, а также здоровые растения. В результате развития указанных заболеваний происходит разрушение хлорофила в листьях картофеля, что приводит к изменению цвета растения. Анализ цветковых характеристик различных типов растительности на базе их гистограмм показывает заметное их различие для почвы и растительности (для всех цветковых каналов) независимо от того, здоровое и больное растение. Данная информация может быть учтена при использовании СНС U-Net для семантической сегментации.

Нейронная сеть, для которой разбатывается модификация [1], принимает на вход матрицу размером $256 \times 256 \times 3$, выходом является матрица $256 \times 256 \times 4$, где каждая из 4 подматриц размером 256×256 показывает степень принадлежности пикселя к одному из искомых классов. Сама сеть содержит 11 сверточных слоев помимо выходного, при этом ресемплирование осуществляется вплоть до разрешения 32×32 . Максимальное количество фильтров в сверточном слое – 256.

Обучающая выборка была получена путем «нарезки» имеющихся аэрофотоснимков с размеченными участками. При этом нарезались участки размером 32, 64 и 128 пикселей с перекрытием и добавлением поворотов на углы, кратные 90° . Для обучения формируются пары (растительность – почва), (здоровая растительность – больная растительность). Размер обучающей выборки: 110528 изображений. Размер валидационной выборки: 27632 изображений (20 % общей базы). Реализация на основе библиотеки Keras. Точность сегментации данной сети составляет 96.89 %.

Сеть была настроена на входной тензор размерностью $512 \times 512 \times 3$. Выходом такой сети является тензор $512 \times 512 \times 1$. Архитектура U-Net представляет собой энкодер-декодер. В основе как энкодера, так и декодера лежит следующий алгоритм: последовательное применение свёртки, нормализации и функции активации. Данные операции проводятся с целью изменения разрешения карты признаков. Так, каждый уровень энкодера понижает ширину и высоту карты (с помощью слоя MaxPool2d), но увеличивает глубину. В свою очередь, декодер увеличивает ширину и высоту карты, но уменьшает глубину (за счёт предварительного применением слоя ConvTranspose2d и объединения карты с соответствующим уровнем энкодера). Описание параметров слоев сети приведено в таблице 1. Через BaseBlock обозначены последовательные применения свёртки, нормализации и функции активации. Общее число обучаемых параметров: 7763041. Общее расчётное потребление памяти: 1648,61 МБ.

В качестве loss-функции используется DiceLoss:

$$\text{Dice}(X, Y) = \frac{2|X \cap Y| + \text{smooth}}{|X| + |Y| + \text{smooth}}$$

где X является предсказанием,

Y – правильно размеченной маской на текущем объекте,

smooth – плавный коэффициент.

В качестве метода оптимизации используется Адам [2] со следующими параметрами:

$\text{lr} = 0,0001$ – коэффициент скорости обучения;

$\beta_1 = 0,9$, $\beta_2 = 0,999$ – коэффициенты, используемые для расчёта скользящих средних градиентов и их квадратов;

$\epsilon = 1e-8$ – слагаемое, добавляемое к знаменателю для улучшения числовой стабильности;

$\text{weight_decay} = 0$ – L2-регуляризация.

При обучении датасет был разбит на 2 множества: для обучения (90 % от общего числа изображений) и валидации (10 %). Также в процессе обучения на каждой итерации использовались следующие преобразования изображений с целью аугментации датасета: случайное изменение масштаба от 90 % до 110 %; случайный поворот изображения от -75 до $+75$ градусов; случайное отражение изображения по горизонтали.

Нейронная сеть была реализована на языке Python с использованием фреймворка PyTorch.

Предложенная в настоящей работе нейронная сеть способна обрабатывать 10 изображений в секунду на ноутбуке с графическим ускорителем NVIDIA GeForce GTX 1060 with Max-Q Design, центральным процессором Intel(R) Core(TM) i7-7700HQ CPU @ 2.80GHz и 16GB оперативной памяти под управлением ОС Ubuntu 16.04. Общее потребление памяти составило не более 2 ГБ.

Проведенные исследования показывают, что совместное использование сети из работы [1] и предложенной дает прирост точности предложенной модификации 1,4% (98,3% против 96,9%) несмотря на то, что вторая реализация обучалась на меньшем общем количестве изображений.

Ключевыми отличиями, повлиявшие на результат, являются: размер входного слоя (512×512 в предлагаемой сети против 256×256 у сети [1]) и глубина нейронной сети (4 пары «энкодер-декодер» против 3 пар у [1]). Они прямо влияют на количество обучаемых параметров и, как следствие, качество сегментации. Из негативных последствий таких изменений можно отметить большее потребление ресурсов.

Список литературы

1. Ganchenko, V. Image Semantic Segmentation Based on Convolutional Neural Networks for Monitoring Agricultural Vegetation / V. Ganchenko, A. Doudkin // Communications in Computer and Information Science, Springer, 2019. – Ch. 5. – V. 1055. – 2019. – P. 52-63.
2. Adam: A Method for Stochastic Optimization [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1412.6980> – Дата доступа : 01.05.2006.

УДК 004.415.25

БЕСПИЛОТНЫЙ ТРАНСПОРТ В АВТОМАТИЗАЦИИ ПРОИЗВОДСТВ

Я. О. Заречный, Е. В. Василюк, А. О. Заречный

Брестский государственный технический университет, Брест

Научный руководитель: В.Н. Шуть, доцент, кандидат технических наук

Роботы – это физические агенты, которые выполняют поставленные перед ними задачи, проводя манипуляции в физическом мире. На данный момент все большую перспективность приобретает мобильная робототехника – область ро-