

Підходи до вирішення задачі розпізнавання та класифікації об'єктів на аерофотознімках з використанням нейромереж

Вінницький національний технічний університет

Анотація

Робота присвячена проектуванню та розробці системи розпізнавання та класифікації об'єктів на аерофотознімках. Було проведено огляд предметної області та розглянуто методи виконання задачі. Для реалізації системи було обрано нейронну мережу архітектури SegNet що працює на базі HTTP-сервера, який надає інтерфейс відправки зображень на обробку та отримання результатів роботи.

Ключові слова: нейромережа, розпізнавання, класифікація, аерофотознімки, семантична сегментація, згорткові нейромережі, SegNet, HTTP, мікросервісна архітектура, горизонтальне масштабування.

Abstract

The work is devoted to the design and development of a system for recognizing and classifying objects in aerial photographs. A review of the subject area was conducted and methods of performing the task were considered. To implement the system, a neural network of SegNet architecture was selected, which works on the basis of HTTP-server, which provides an interface for sending images for processing and obtaining results

Keywords: neural network, recognition, classification, aerial photography, semantic segmentation, convolutional neural networks, SegNet, HTTP, microservice architecture, horizontal scaling.

З давніх часів для людської цивілізації існує необхідність орієнтування на місцевості, прокладення зручних і оптимальних маршрутів на ній, розподілу земельних ділянок та його документування. Ця необхідність виникла в різних галузях людської діяльності: військовій, торгівельній, сільськогосподарській. З часом ці потреби тільки зростали. Тому розвиток таких дисциплін як картографія та топографія є дуже важливою частиною розвитку людської цивілізації в цілому.

Автоматичне розпізнавання об'єктів на аерофотознімках місцевості дозволяє складати максимально точні карти місцевості в короткі терміни та проводити моніторинг різного роду людської діяльності: визначати стан посівів на полях, аналізувати густоту міської забудови, кількість зелених насаджень, створення детальних карт автомобільних шляхів.

Для того, щоб знаходити та визначати типи об'єктів зображених на аерофотознімках, потрібно вирішити задачу семантичної сегментації зображення.

Сегментація — це процес обробки зображення з ціллю розділення його на декілька сегментів. Сегментом є множина пікселів, яку також називаються об'єктом. Найчастіше метою сегментації є перетворення представлення зображення для того, щоб полегшити його аналіз, або спростити передачу каналами зв'язку.

Для вирішення задачі семантичної сегментації використовуються декілька архітектур нейромереж, таких як: SegNet, DeconvNet, FCN, U-Net. За результатами дослідження[1] де було проведено тестування семантичної сегментації на прикладі дорожньої сцени, мережа SegNet показує кращі результати на більшій частині класів розпізнавання та найбільше значення усередненої загальної точності з усіх представлених архітектур. Також наведено порівняння архітектур в задачі сегментації інтер'єрної сцени, в якому SegNet знову показує одні з найкращих результатів, лише трохи поступаючись архітектурі DeepLabv1. SegNet показує кращі результати для великих за розміром об'єктів, але дещо гірші для менших об'єктів. В задачі розпізнавання об'єктів на аерофотознімках це може означати гіршу якість розпізнавання для малих об'єктів, таких як малі будівлі, автомобілі, поодинокі невеликі дерева.

SegNet є менш швидкодіючою ніж більшість інших архітектур через наявність декодера, але швидкою за DeconvNet через відсутність повнозв'язних шарів. Важливою перевагою SegNet є набагато менше споживання оперативної пам'яті(у випадку використання відеоадаптера — менше

використання пам'яті відеоадаптера). SegNet може використовувати майже в два рази менше оперативної пам'яті чи пам'яті відеоадаптера.

Також можна зазначити що натренована модель SegNet займає менше місця на диску ніж найближчі аналоги. У дослідженні модель SegNet мала об'єм в 117МБ що трохи більше ніж 83 у DeepLab, але значно краще ніж 539 у FCN, чи 877 у DeconvNet. Ця особливість дозволить зменшити розмір інсталяції на диску.

Для реалізації системи було обрано нейронну мережу архітектури SegNet що працює на базі HTTP-сервера, який надає інтерфейс відправки зображень на обробку та отримання результатів роботи. Це дозволяє створити уніфікований інтерфейс користування функціоналом розпізнавання та зробити можливим горизонтальне масштабування системи шляхом додавання нових серверів для збільшення загальної швидкодії обробки даних.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Огляд архітектури SegNet [Електронний ресурс] — <https://towardsdatascience.com/review-segnet-semantic-segmentation-e66f2e30fb96>.

Ярмошук Дмитро Олександрович — студент групи ІКІ-20м, кафедри комп'ютерної інженерії ВНТУ, Вінницький національний технічний університет, м. Вінниця, e-mail: dima.yarmoshchuk@gmail.com

Науковий керівник: **Ткаченко Олександр Миколайович** — кандидат технічних наук, доцент кафедри обчислювальної техніки, Вінницький національний технічний університет, м. Вінниця, e-mail: alextk1960@gmail.com

Yarmoshchuk Dmytro Oleksandrovych — student of the Computer Techniques Chair, Vinnytsia National Technical University, Vinnytsia, e-mail: dima.yarmoshchuk@gmail.com

Research advisor: **Tkachenko Oleksandr Mykolaiovych** - Cand. Sc. (Eng), Associate Professor of Computing Engineering Departmen, Vinnytsia National Technical University, Vinnytsia, e-mail: alextk1960@gmail.com