

МЕТОДИ СЕМАНТИЧНОЇ СЕГМЕНТАЦІЇ ЗОБРАЖЕННЯ ТА ЇХ ПОРІВНЯННЯ

Ткаченко Н.О.

Науковий керівник – к.т.н., доц. Кобилін О.А.

Харківський національний університет радіоелектроніки, каф. ІНФ,
м. Харків, Україна,

тел.: (057) 702-14-19, e-mail: nazar.tkachenko@nure.ua

This work is devoted to research, comparison and implementation of semantic image segmentation methods with the aim of developing an effective algorithm for automatic selection and classification of objects in images. The purpose of the study is to determine the optimal method or combination of methods that provide the highest segmentation accuracy and speed of image processing. Convolutional Neural Networks (CNNs) were chosen as the main method, noting their high accuracy, generalizability and efficiency in different settings. The implementation of the selected method is performed using a framework for deep learning. The research results can be useful for various applications in modern image processing systems and intelligent systems.

Семантична сегментація зображень – це процес розділення зображення на різні частини або області з метою призначення семантичного значення кожній з цих областей. Іншими словами, це завдання класифікації кожного пікселя на зображенні, присвоюючи йому конкретну мітку, що відповідає класу або категорії.

Семантична сегментація має велике значення в обробці зображень через свою здатність відокремлювати та ідентифікувати об'єкти та їхні частини на зображенні [1]. Вона допомагає в багатьох важливих застосуваннях, таких як розпізнавання об'єктів на зображеннях, автоматичне вирізання та виділення областей інтересу, аналіз медичних зображень, віртуальна реальність, автономні автомобілі, моніторинг навколишнього середовища тощо. Отже, семантична сегментація відіграє важливу роль у багатьох сучасних системах обробки зображень та розвитку інтелектуальних систем.

Мета дослідження полягає у вивченні, порівнянні та реалізації різних методів семантичної сегментації зображень з метою розробки ефективного алгоритму для автоматичного виділення та класифікації об'єктів на зображеннях. Це дослідження спрямоване на визначення оптимального методу або комбінації методів, які забезпечують найвищу точність сегментації та швидкодію обробки зображень. Додатковою метою є реалізація обраного методу з метою його практичного використання у різноманітних застосуваннях, включаючи комп'ютерний зір, медичні дослідження, автономні системи та інші області.

Конволюційні нейронні мережі є одними з найбільш популярних та ефективних методів семантичної сегментації. Вони використовуються для

автоматичного визначення класу кожного пікселя на зображенні шляхом застосування фільтрів та згорток для виявлення ознак на різних рівнях абстракції.

Рекурентні нейронні мережі, хоч і менш поширені в сфері семантичної сегментації порівняно з CNN, все ж використовуються для врахування контексту та послідовності в зображеннях. Вони можуть бути використані для моделювання просторових та часових залежностей між пікселями у зображенні.

Класичні методи сегментації включають в себе підходи, що базуються на характеристиках, текстурних ознаках та алгоритмах кластеризації. Ці методи, хоч і менш потужні порівняно з методами глибокого навчання, все ще знаходять своє застосування у випадках, де обмежені обчислювальні ресурси або недостаток великої кількості даних для навчання нейронних мереж [2].

Конволюційні нейронні мережі (CNN) загалом вважаються найефективнішими для багатьох завдань семантичної сегментації через їхню здатність автоматично вивчати корисні ознаки з даних та адаптуватися до нових умов [4]. Рекурентні нейронні мережі (RNN) можуть бути корисними для завдань, де важливий контекст або послідовність, таких як обробка послідовних зображень або відео [5, 6]. Хоча класичні методи можуть бути менш ефективними порівняно з глибоким навчанням у деяких випадках, вони можуть залишатися конкурентоспроможними в умовах, де обмежені обчислювальні ресурси або кількість доступних даних для навчання.

Для виконання поставленої задачі було обрано CNN. CNN показали вражаючі результати в багатьох сферах обробки зображень, включаючи розпізнавання об'єктів, класифікацію зображень та семантичну сегментацію. Їхній успіх у вирішенні складних завдань свідчить про їхню ефективність та потужність. CNN можуть бути легко масштабовані для різних завдань та різних розмірів зображень. Вони можуть бути навчені на великих обсягах даних та використовуватися для обробки зображень різної роздільної здатності без великих змін у структурі моделі. За останні кілька років було розроблено безліч відкритих бібліотек та фреймворків для роботи з CNN, таких як TensorFlow, PyTorch, Keras тощо. Це робить їх легкодоступними та легкими у використанні для дослідників та розробників.

Отже, обираючи CNN для семантичної сегментації зображень, ми можемо впевнено розраховувати на їхню здатність до вивчення складних ознак зображень та ефективне рішення задач сегментації зображень.

Для реалізації методу семантичної сегментації зображень за допомогою конволюційних нейронних мереж (CNN) було обрано наступні засоби:

– навчальний набір даних, який складається з зображень та відповідних міток сегментації. Структура CNN, яка включає в себе шари згортки, пулінгу, повнозв'язаних шарів тощо. Ця мережа буде використовуватися для вивчення ознак та виконання сегментації зображень;

– використання фреймворку для глибокого навчання, такого як TensorFlow, PyTorch, Keras тощо, для реалізації та навчання CNN. Ці фреймворки надають зручні інтерфейси та оптимізовані інструменти для роботи з нейронними мережами;

– для оцінки ефективності та точності моделі потрібні метрики, такі як індекс схожості Якобі, долю правильних класифікацій (IoU), точність, відгук тощо.

CNN демонструють вражаючу точність у вирішенні задач семантичної сегментації зображень. Їхні здатності до вивчення складних корисних ознак зображень дозволяють отримувати високоякісні результати, які зазвичай перевершують традиційні методи. CNN проявляють високу здатність до узагальнення на нові дані, що робить їх ефективними для різних завдань семантичної сегментації зображень. Після навчання на великому наборі даних вони можуть успішно застосовуватися до нових зображень з різних джерел та умов. Використання CNN для семантичної сегментації зображень дозволяє автоматизувати процес виділення областей інтересу на зображеннях, що робить його більш ефективним та швидким у порівнянні з ручними або напівавтоматичними методами.

Список використаних джерел:

1. Mechea D. What is Panoptic Segmentation and why you should care. Medium. 2019. URL: <https://medium.com/@danielmechea/what-is-panoptic-segmentation-and-why-you-should-care-7f6c953d2a6a>

2. The Berkeley Segmentation Dataset and Benchmark: веб-сайт. URL: <https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>

3. Kaur D., Kaur Y. Various Image Segmentation Techniques: A Review. International Journal of Computer Science and Mobile Computing. 2014. Vol. 5, №. 1. P. 809–814.

4. Gorokhovatskyi, V., Tvoroshenko, I., Kobylin, O., Vlasenko, N. Search for Visual Objects by Request in the Form of a Cluster Representation for the Structural Image Description, Advances in Electrical and Electronic Engineering, 2023, 21 (1), pp. 19–27. DOI: 10.15598/aeee.v21i1.4661.