

УДК 004.932

О.І. Земляков, студент гр. ПМ-31мп, PhD., ст.в. Момот А.С.
КПІ ім. Ігоря Сікорського

АНАЛІЗ ЕФЕКТИВНОСТІ АРХІТЕКТУР НЕЙРОННИХ МЕРЕЖ ДЛЯ АВТОМАТИЗОВАНОЇ СЕГМЕНТАЦІЇ КАРТОГРАФІЧНИХ ЗОБРАЖЕНЬ

Анотація. У матеріалах розглядається порівняльний аналіз сучасних архітектур нейронних мереж, таких як FCN, DeepLabv3 і U-Net, для задач сегментації картографічних зображень. Особлива увага приділяється перевагам і недолікам кожної архітектури, зокрема точності, обчислювальній складності та вимогам до обсягу даних для навчання. Аналіз показав, що U-Net є найбільш оптимальним вибором для обробки картографічних даних завдяки збалансованості між точністю та вимогами до ресурсів. У статті наголошується на важливості автоматизації сегментації для підвищення ефективності й об'єктивності роботи з великими обсягами даних.

Ключові слова: картографічні зображення, сегментація зображень, нейронні мережі, семантична сегментація.

ВСТУП

В епоху цифрового розвитку автоматизація процесів обробки та аналізу картографічних зображень стає все більш актуальною задачею. Сегментація картографічних зображень є ключовим етапом у процесі цифрової обробки карт, що дозволяє виділяти та класифікувати різні об'єкти місцевості: водойми, ліси, міську забудову, транспортну інфраструктуру. Автоматизація цього процесу не лише значно прискорює обробку великих масивів картографічних даних, але й підвищує точність та об'єктивність результатів.

Актуальність дослідження підкреслюється стрімким розвитком геоінформаційних систем та зростаючою потребою в оперативному оновленні картографічної інформації. Традиційні методи ручної сегментації є трудомісткими та часозатратними, що робить їх малоефективними для обробки сучасних обсягів картографічних даних. Впровадження автоматизованих методів сегментації стає необхідною умовою для забезпечення ефективного картографічного виробництва [1].

СУЧАСНІ МЕТОДИ СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ

Традиційні методи сегментації зображень включають порогову сегментацію, метод активних контурів та регіональну сегментацію. Порогова сегментація базується на розділенні пікселів зображення на групи на основі порівняння їх інтенсивності з одним або декількома пороговими значеннями. Метод активних контурів використовує деформовані криві під впливом внутрішніх та зовнішніх сил для виділення границь об'єктів. Регіональна сегментація базується на принципі виділення зв'язних областей з однорідними характеристиками Рис. 1.

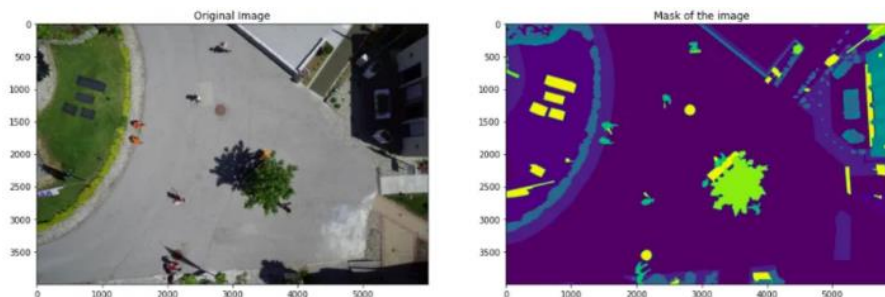


Рисунок 1. Приклад сегментованого картографічного зображення

Поява глибоких нейронних мереж значно розширила можливості сегментації зображень. Особливу роль відіграють згорткові нейронні мережі, які здатні автоматично вивчати ієрархічні представлення ознак з даних. Семантична сегментація з використанням глибоких нейронних мереж забезпечує піксельну класифікацію зображення з присвоєнням кожному пікселю певного семантичного класу [2].

ОГЛЯД АРХІТЕКТУР НЕЙРОННИХ МЕРЕЖ

FCN (Fully Convolutional Network) представляє собою архітектуру, що використовується переважно для семантичної сегментації. Особливістю FCN є використання лише локально пов'язаних рівнів, таких як згортка, об'єднання та підвищення дискретизації. Архітектура складається з енкодера та декодера, де енкодер відповідає за вилучення ознак, а декодер виконує відновлення просторової розмірності та формування кінцевої сегментаційної маски.

Важливою особливістю FCN є здатність обробляти зображення довільного розміру, оскільки мережа не містить повнозв'язних шарів. Це досягається завдяки використанню шарів деконволюції для відновлення просторової розмірності карт ознак.

DeepLabv3 використовує розширені (atrous) згортки для контролю рецептивного поля та просторової роздільної здатності карт ознак. Архітектура включає модуль Atrous Spatial Pyramid Pooling для ефективного вилучення багатомасштабних ознак. Це дозволяє мережі краще обробляти об'єкти різного масштабу на зображенні [3].

U-Net є однією з найефективніших архітектур для сегментації зображень, що складається з симетричного шляху стиснення (енкодер) та розширення (декодер). Енкодер включає послідовність згорткових блоків, де кожен блок містить дві згортки 3×3 з ReLU активацією та операцію максимальної підвибірки 2×2 для зменшення просторової розмірності. При цьому кількість каналів ознак подвоюється на кожному кроці стиснення, що дозволяє мережі вивчати все більш складні та абстрактні характеристики зображення. Декодер мережі виконує поступове відновлення просторової розмірності через послідовність шарів деконволюції. На кожному етапі розширення здійснюється конкатенація з відповідними картами ознак з шляху стиснення через skip-з'єднання, що дозволяє зберігати деталізовану інформацію з ранніх шарів. Фінальний шар використовує згортку 1×1 для формування вихідної сегментаційної маски. Така архітектура демонструє високу ефективність навіть при навчанні на обмежених наборах даних, що особливо важливо для специфічних задач картографічної сегментації.

ПОРІВНЯЛЬНИЙ АНАЛІЗ ЕФЕКТИВНОСТІ РОЗГЛЯНУТИХ АРХІТЕКТУР

В даній статті [4], розглянуто нейронну мережу створену на базі архітектури FCN (Fully Convolutional Network). Провівши аналіз можна сказати, що FCN досить проста в реалізації та швидко навчається, проте має суттєві обмеження. Через послідовне зменшення розмірності в згорткових шарах втрачається багато детальної інформації, що призводить до розмитих границь об'єктів на

результуючій масці сегментації. Крім того, архітектура погано справляється з об'єктами різного масштабу.

Проаналізувавши статтю [5] в якій розкриваються особливості архітектури DeepLabv3 робимо висновок, що основною її особливістю є використання атрусних (розширених) згорток, які дозволяють збільшити рецептивне поле без втрати роздільної здатності. Модуль Atrous Spatial Pyramid Pooling (ASPP) забезпечує ефективну обробку контексту на різних масштабах. Це дозволяє досягти високої точності сегментації, особливо на складних сценах з об'єктами різного розміру. Однак така потужність має свою ціну - архітектура потребує значних обчислювальних ресурсів та пам'яті, довше навчається та складніша в налаштуванні.

В статті [6] розглядається нейронна мережа з архітектурою U-Net. U-Net пропонує елегантне рішення у вигляді симетричної архітектури енкодер-декодер зі skip-з'єднаннями між відповідними рівнями. Це дозволяє зберегти детальну просторову інформацію при відновленні роздільної здатності в декодері. Архітектура показує відмінні результати навіть при невеликих наборах даних для навчання, що робить її особливо привабливою для застосувань в сфері обробки картографічних зображень. U-Net можна розглядати як золоту середину між простотою FCN та потужністю DeepLabv3.

Проаналізувавши дані архітектури і нейронні мережі побудовані на них, які були представленні у статтях згаданих вище, була побудована табл. 1.

Таблиця 1. Порівняння ефективності архітектур нейронних мереж

<i>Метрика</i>	<i>FCN</i>	<i>DeepLabv3</i>	<i>U-Net</i>
IoU	62.2%	82.47%	92%
Overall Accuracy	81.4%	92.13%	95.47%
Кількість параметрів	134M	22.96M	31M
GPU пам'ять	Висока	Низька	Середня
Розмір навчального набору	Великий (>1000)	Середній	Малий (~30)
Потреба в аугментації	Низька	Середня	Висока

Розглянуті архітектури демонструють різний рівень адаптивності до задач картографічної сегментації залежно від складності сцени та доступних ресурсів. FCN, незважаючи на простоту та високу швидкість обробки, є менш ефективною для задач, де важлива деталізація та точне відтворення границь об'єктів. DeepLabv3, навпаки, забезпечує відмінні результати для зображень зі складною структурою, однак її обчислювальна складність може стати критичним фактором для реального впровадження, особливо в умовах обмежених апаратних можливостей.

З іншого боку, U-Net завдяки використанню skip-з'єднань та симетричної

архітектури енкодера-декодера забезпечує баланс між точністю та вимогами до обчислювальних ресурсів. Вона також демонструє високу ефективність при невеликих обсягах навчальних даних, що є важливим для спеціалізованих картографічних задач. Це робить U-Net універсальним інструментом для широкого спектра застосувань, починаючи від стандартної картографії до аналізу даних у геоінформаційних системах.

ВИСНОВКИ

Проведений аналіз різних архітектур нейронних мереж для сегментації картографічних зображень показує, що базова архітектура U-Net є найбільш доцільною для початкового етапу досліджень. Це обумовлено її здатністю ефективно обробляти просторові дані та зберігати деталі зображення при відносно низькій обчислювальній складності. U-Net найкраще підходить для застосувань з обмеженими ресурсами або даними для навчання. Однак для задач, що вимагають вищої точності на складних сценах або різномірних об'єктах, доцільно розглядати DeepLabv3 за умови наявності достатніх обчислювальних потужностей. У подальших дослідженнях рекомендується розробка гібридних моделей, які могли б поєднати простоту U-Net із потужністю та гнучкістю DeepLabv3.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Momot A. Automation of ultrasound breast cancer images classification using deep neural networks / A. Momot, R. Galagan, M. Zabolueva // Sciences of Europe. – 2022. – №96. – pp. 38–41. DOI
- [2] Momot A. Deep learning automated data analysis of security infrared cameras / A. Momot, I. Skladchikov // Slovak international scientific journal. – 2021. – №52. – pp. 13–16. 81.
- [3] Sarah Guido, Andreas Müller. Introduction to Machine Learning with Python. A guide for Data Scientists. — O’Rilly Media, 2016. — 285p.
- [4] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431-3440.
- [5] Y. Wang, L. Yang, X. Liu, and P. Yan, "An improved semantic segmentation algorithm for high-resolution remote sensing images based on DeepLabv3," Scientific Reports, vol. 14, no. 9716, 2024.
- [6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention (MICCAI), LNCS, vol. 9351, 2015, pp. 234-241.
- [7] Momot A. Automated segmentation of ultrasound medical images using the Attention U-Net model / A. Momot, M. Zabolueva, R. Galagan // Norwegian Journal of development of the International Science. – 2024. – №128. – pp. 56–60.

Наук. керівник – PhD., ст.в. Momot A.C.