

В.В. Гнатушенко, І.О. Гненний, І.М. Удовик, О.С. Шевцова

СЕГМЕНТАЦІЯ АЕРОКОСМІЧНИХ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

Анотація. В роботі запропоновано новий метод семантичної сегментації аерокосмічних зображень високого просторового розрізнення на основі згорткових нейронних мереж та генерації масок. Реалізовано мережу, що складається з окремих мереж одного класу. Семантичні ознаки використовуються для зменшення помилок семантичного рівня контуру. Оскільки незалежний тест виявив деякі недоліки в певних класах земного покриву, подальші дослідження будуть присвячено питанням створення каталогу навчальних вибірок для підтримки успішної диференціації об'єктів. Загалом оцінки точності демонструють сучасні характеристики розробленої моделі, а також ефективність цієї комбінації мережі та наборів даних для тестових регіонів. Результати показують, що запропонований алгоритм може ефективно покращити загальну точність семантичної сегментації зображень дистанційного зондування високого просторового розрізнення та скоротити загальний час навчання та час сегментації.

Ключові слова: сегментація, аерокосмічне зображення, машинне навчання, згорткова нейронна мережа.

Постановка проблеми. Зйомка Землі з космосу широко використовується в різних галузях господарства та сферах бізнесу, у тому числі для управління територіями, моніторингу надзвичайних ситуацій, для вивчення та екологічних досліджень нашої планети [1, 2]. За допомогою супутників вчені та дослідники можуть збирати та зберігати зображення високої просторової здатності, які потім можна обробляти, сегментувати та класифікувати. Сегментація зображення призначена для відділення когерентних областей друг від друга. Встановлення порогових значень використовується для вилучення переднього плану із фону на основі інтенсивності або кольору пікселів. Дослідники також використовують текстуру та додають інші фактори для кращого розділення в регіонах. Більш складна версія може варіюватися від заснованої на кольорі для використання більш складних доменів для сортування регіонів. Значного поширення отримали предметно-орієнтовані методи класифікації, засновані на

попередній сегментації зображень з наступним аналізом параметрів виділених зон. Однак їх застосування до реальних аерокосмічних даних утруднено, оскільки для природних ландшафтів зображення високої роздільної здатності зазвичай містять інформацію про текстуру і характеризуються плавними переходами. Це негативно позначається на якості сегментації: на межах сегментів, сформованих за таких умов, виникають зони невизначеності. Інша проблема об'єктно-орієнтованих методів полягає у складності застосування контекстного аналізу до результатів сегментації [3, 4].

Аналіз останніх досліджень. В останні роки великої популярності набула сегментація зображень за допомогою згорткових нейронних мереж (CNN) [3-6]. Суттєвою відмінністю цього підходу є анування віднесення кожного пікселя до того чи іншого класу об'єктів так, що процес навчання подібних мереж повністю контролюється. В основному, як впливає з назви, нейронні згорткові мережі використовують операцію згортки для вилучення ознак. Було досліджено багато нейронних мереж, побудованих на цьому принципі. Але основне використання згорткових нейронних мереж полягає в їхній цінності в комп'ютерному зорі. При великих розмірах зображень штучні нейронні мережі також повинні бути досить великими для розміщення даних, але зі згортковими нейронними мережами ця проблема зменшується в геометричній прогресії, оскільки з кожним шаром розмір даних стає все меншим і меншим. Ці моделі широко використовувалися в моделях виявлення та класифікації об'єктів, таких як сімейства VGG16 та ResNet [4-7]. Цей метод також є основою сегментації зображень. Першою віхою семантичної сегментації була повнозгорткова мережа (FCN), що здійснювала класифікацію пікселів за пікселями або маскою сегментації. З появою FCN для семантичної сегментації відкрилися нові можливості. Вилучені функції можна було не тільки виводити для класифікації зображень, але тепер їх можна також використовувати як вхідні дані для задач сегментації зображень. FCN класифікує зображення лише на рівні пікселів. На відміну від класичної CNN, що використовує пов'язаний рівень для отримання вектора ознак фіксованої довжини для класифікації після згорткового рівня, повнозгорткова мережа може приймати вхідні зображення будь-якого розміру і використовувати рівень деконволюції для останнього рівня згортки, і, нарешті, перейти на карту об'єктів з підвищеною дискретизацією.

Метою даної роботи є розробка методів, алгоритмів та програмних засобів, що дозволить автоматизувати процес сегментації та подальшої

класифікації супутникових зображень високої просторової здатності та підвищити відповідну точність обробки.

Викладення основного матеріалу. На сьогоднішній день не існує чітко регламентованих правил реалізації структури нейронної мережі: кількість та організація шарів, кількість та розмір карт ознак, розмір матриць згортки, вибір алгоритму навчання. При розробці структури CNN слід розуміти, що невелика кількість параметрів мережі може знизити точність сегментації. Однак велика кількість параметрів збільшує обчислювальну складність мережі та не завжди забезпечує покращення класифікувальних здібностей мережі. Таким чином, розробка оптимальної структури нейронної мережі є емпіричним процесом, який включає велику кількість експериментальних досліджень.

В роботі запропоновано новий метод мультикласової сегментації аерокосмічних зображень високого просторового розрізнення на основі CNN та генерації масок та реалізовано мережу, що складається з окремих мереж одного класу (рис.1). Спочатку ми отримуємо маску сегментації окремих індивідуальних класів окремими мережами декодерів, потім об'єднуємо та здійснюємо постобробку для створення необхідної вихідної маски. Зауважимо, що точність згорткової мережі лише з одним класом набагато вища порівняно з метриками моделей, які сегментують усі класи разом. Ми спробували використати цю властивість при створенні моделі, що поєднує окремі сегментовані маски в єдине ціле. Таким чином, у нас є масив індивідуальних мереж згортання для окремих класів. Кожен клас підключається до вихідних даних об'єднаної моделі класів для отримання більш високих результатів. Вхідні дані поділяються на зображення розміром 512 x 512 пікселів. Потім ці вхідні зображення одночасно проходять через декілька окремих мереж та комбіновану мережу. Кожна з цих мереж виводить маску. Крім того, видаляються області, що не перекриваються. Потім видаляються всі несегментовані області. Нарешті, за постобробки отримуємо вихідну маску (рис.1).

Індивідуальні мережі являють собою набір U-Nets, навчених створювати маски одного класу, відповідно 1 для класу, що навчається, і 0 – в іншому випадку. Вхідними даними для мереж є те саме вхідне зображення, що і для всієї моделі. Їх вивід – це маска одного класу з одиниць та нулів. Кожна з мереж U-Net заснована на архітектурі MobileNetV2 [8, 9]. Особливістю MobileNetV2 є використання замість звичайних залишкових блоків зворотних залишкових блоків. Їхня відмінність полягає в тому, що обхід з'єднує блоки з пляшковими горлечками.

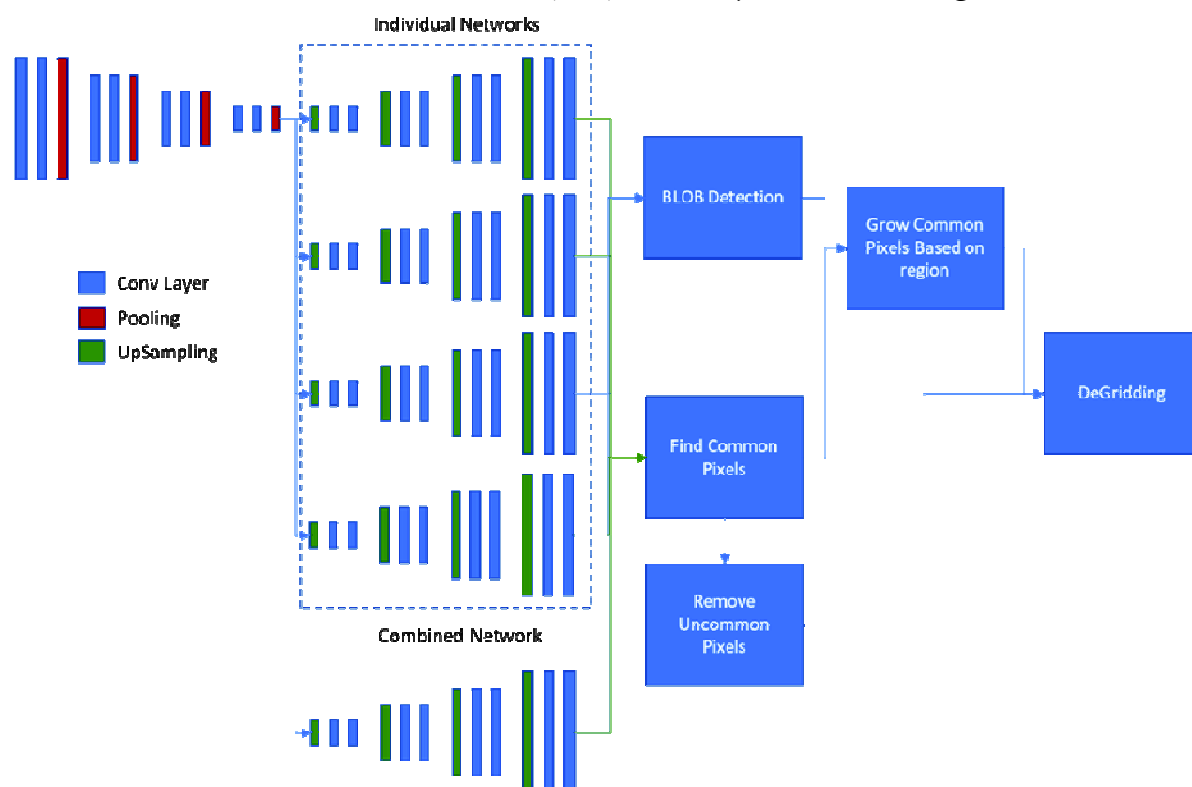


Рисунок 1 - Модель обробки

Комбінована мережа є основою нашої моделі. Подібно до індивідуальних мереж, комбінована мережа є мережею U-Net з MobileNetV2 в якості магістралі. Але замість навчання на двійковій масці мережа навчається на наземних даних і таким чином видає повну маску прогнозування. Комбінована маска створює зображення 512 x 512 пікселів. При обробці зображення Blob-об'єкти визначаються як область з'єднаних пікселів з однаковим кольором або інтенсивністю. Якщо регіон не пов'язаний з іншим жодним зі своїх пікселів ні безпосередньо, ні по діагоналі, то він вважається окремим Blob-об'єктом. Коли позначений кожен Blob-об'єкт, його мітки передаються наступному блоку. У разі комбінованої мережі готується окремий набір зображень у градаціях сірого для кожного класу, щоб позначити відповідні класи. Потім цей набір зображень передається до наступного блоку. Далі ми видаляємо з базового зображення пікселі, що не перекриваються з масками зображень класів. Спочатку створюється маска відтінків сірого для кожного класу в базовому зображенні і послідовно порівнюється з масками зображень класів, створених з індивідуальних мереж. Якщо піксель «невстановлений» на 1 в обох зображеннях у градаціях сірого, піксель встановлюється на 0. При цьому ізолюються спільні пікселі від зображень. Але це також дає багато пікселів з 0 у базовій

масці. Ця проблема вирішується в наступному блоці, в якому спочатку визначається порядок зростання пікселів класу, а потім клас, що має найбільшу кількість пікселів у зображенні. Таким чином, упорядкування класів за цим критерієм і наступний процес виконується для кожного з класів на основі одного й того ж порядку. Для кожного пікселя в базовому зображенні обирається сегмент, якому він відповідає як в індивідуальній масці, так і в комбінованій масці. Потім визначаються пікселі, що мають значення 0 в базовій масці та спочатку не були класифіковані як клас 0. Порожнє місце заповнюється значенням класу, якщо піксель знаходиться в одному з можливих місць зростання. Цей процес повторюється для кожного з різних класів, щоб створити остаточну розширену маску. Врешті решт, щоб позбутися артефактів, застосовується середнє значення сегментованих пікселів у блоці 8×8 пікселів. Відповідний процес створює вихідне зображення 512×512 , яке ми об'єднуємо з іншими обрізаними частинами зображення, щоб отримати остаточне зображення, що корелює з вхідним зображенням. Фрагмент результатів роботи запропонованого підходу подано на рис.2.



Рисунок 2 - Фрагмент результатів роботи нейронної мережі

Для оцінки якості сегментації застосовано стандартні метрики, зокрема IoU (intersection over union), що оцінює близькість передбачуваної маски до оригінальної [6]. Для кожного класу об'єктів за I позначається кількість пікселів, де клас передбаченої маски збігається з оригінальним, а за U кількість пікселів, де хоча б одна з масок вказала на об'єкт даного класу. Знаходимо відношення I/U. Якість сегментації зображення – середнє значення даної мет-

рики за всіма класами. Оціночні індекси, отримані в результаті сегментації, наведено в таблиці 1.

Таблиця 1

Оцінка		Ліс	Луга	Тіні	Ґрунт	Будівлі	Дороги
Precision		89	82	86	84	83	88
Recall		86	87	88	89	91	90
IoT		82	74	73	75	74	78

Висновки. В роботі запропоновано новий метод мультикласової сегментації аерокосмічних зображень високого просторового розрізнення на основі CNN та генерації масок та реалізовано мережу, що складається з окремих мереж одного класу. Семантичні ознаки використовуються для зменшення помилок семантичного рівня контуру. Оскільки незалежний тест виявив деякі недоліки в певних класах земного покриття, подальші дослідження будуть присвячено питанням створення каталогу навчальних вибірок для підтримки успішної диференціації об'єктів. Загалом оцінки точності демонструють сучасні характеристики розробленої моделі, а також ефективність цієї комбінації мережі та наборів даних для тестових регіонів. Результати показують, що запропонований алгоритм може ефективно покращити загальну точність семантичної сегментації зображень дистанційного зондування високої роздільної здатності та скоротити загальний час навчання та час сегментації.

ЛІТЕРАТУРА / LITERATURE

1. Hordiiuk, D. M., Hnatushenko, V. V. (2017). Neural network and local laplace filter methods applied to very high resolution remote sensing imagery in urban damage detection. 2017 IEEE International Young Scientists Forum on Applied Physics and Engineering (YSF).doi:10.1109/ysf.2017.8126648.
2. Mozgovoy Dmitriy, Hnatushenko Volodymyr, Vasyliiev Volodymyr. Accuracy evaluation of automated object recognition using multispectral aerial images and neural network. Proc. SPIE 10806, Tenth International Conference on Digital Image Processing (ICDIP 2018), 108060H (9 August 2018); <https://doi.org/10.1117/12.2502905>.
3. Brand A.K., Manandhar, A.: SEMANTIC SEGMENTATION OF BURNED AREAS IN SATELLITE IMAGES USING A U-NET-BASED CONVOLUTIONAL NEURAL NETWORK, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B3-2021, 47–53, <https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-47-2021>, 2021.

4. Hnatushenko V., Zhernovyi V. (2020) Method of Improving Instance Segmentation for Very High Resolution Remote Sensing Imagery Using Deep Learning. In: Babichev S., Peleshko D., Vynokurova O. (eds) Data Stream Mining & Processing. DSMP 2020. Communications in Computer and Information Science, vol 1158. Springer, Cham. https://doi.org/10.1007/978-3-030-61656-4_21
5. Hnatushenko Volodymyr, Zhernovyi Vadym, Udovik Iryna, Shevtsova Olga. Intelligent System for Building Separation on a Semantically Segmented Map. 2 International Workshop on Intelligent Information Technologies & Systems of Information Security (IntelITSIS-2021) Khmelnytskyi, Ukraine. <http://ceur-ws.org/Vol-2853/keynote1.pdf>
6. Marushko E.E, Doudkin A.A, Zheng X. Identification of Earth's surface objects using ensembles of convolutional neural networks. Journal of the Belarusian State University. Mathematics and Informatics. 2021;2:114–123.
7. Simonyan, K., Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR, abs/1409.1556.
8. Sandler M., Howard A., Zhu M., Zhmoginov A. and Chen L. MobileNetV2: Inverted Residuals and Linear Bottlenecks. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510-4520, doi: 10.1109/CVPR.2018.00474.
9. Knapheide J., Stabernack B., Kuhnke M. A High Throughput MobileNetV2 FPGA Implementation Based on a Flexible Architecture for Depthwise Separable Convolution. 30th International Conference on Field-Programmable Logic and Applications (FPL), 2020, pp. 277-283, doi: 10.1109/FPL50879.2020.00053.

REFERENCES

1. Hordiiuk, D. M., Hnatushenko, V. V. (2017). Neural network and local laplace filter methods applied to very high resolution remote sensing imagery in urban damage detection. 2017 IEEE International Young Scientists Forum on Applied Physics and Engineering (YSF).doi:10.1109/ysf.2017.8126648.
2. Mozgovoy Dmitriy, Hnatushenko Volodymyr, Vasyliiev Volodymyr. Accuracy evaluation of automated object recognition using multispectral aerial images and neural network. Proc. SPIE 10806, Tenth International Conference on Digital Image Processing (ICDIP 2018), 108060H (9 August 2018); <https://doi.org/10.1117/12.2502905>.
3. Brand A.K., Manandhar, A.: SEMANTIC SEGMENTATION OF BURNED AREAS IN SATELLITE IMAGES USING A U-NET-BASED CONVOLUTIONAL NEURAL NETWORK, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B3-2021, 47–53, <https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-47-2021>, 2021.

4. Hnatushenko V., Zhernovyi V. (2020) Method of Improving Instance Segmentation for Very High Resolution Remote Sensing Imagery Using Deep Learning. In: Babichev S., Peleshko D., Vynokurova O. (eds) Data Stream Mining & Processing. DSMP 2020. Communications in Computer and Information Science, vol 1158. Springer, Cham. https://doi.org/10.1007/978-3-030-61656-4_21
5. Hnatushenko Volodymyr, Zhernovyi Vadym, Udovik Iryna, Shevtsova Olga. Intelligent System for Building Separation on a Semantically Segmented Map. 2 International Workshop on Intelligent Information Technologies & Systems of Information Security (IntelITSIS-2021) Khmelnytskyi, Ukraine. <http://ceur-ws.org/Vol-2853/keynote1.pdf>
6. Marushko E.E, Doudkin A.A, Zheng X. Identification of Earth's surface objects using ensembles of convolutional neural networks. Journal of the Belarusian State University. Mathematics and Informatics. 2021;2:114–123.
7. Simonyan, K., Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR, abs/1409.1556.
8. Sandler M., Howard A., Zhu M., Zhmoginov A. and Chen L. MobileNetV2: Inverted Residuals and Linear Bottlenecks. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510-4520, doi: 10.1109/CVPR.2018.00474.
9. Knapheide J., Stabernack B., Kuhnke M. A High Throughput MobileNetV2 FPGA Implementation Based on a Flexible Architecture for Depthwise Separable Convolution. 30th International Conference on Field-Programmable Logic and Applications (FPL), 2020, pp. 277-283, doi: 10.1109/FPL50879.2020.00053.

Received 07.09.2021.
Accepted 10.09.2021.

Сегментация аэрокосмических изображений с использованием свёрточных нейронных сетей

В работе предложен новый метод сегментации аэрокосмических изображений высокого пространственного разрешения на основе свёрточных нейронных сетей и генерации масок. Реализована сеть, состоящая из отдельных сетей одного класса, работающих над расширением сегментации. Семантические признаки используются для уменьшения ошибок семантического уровня контура. Поскольку независимый тест выявил некоторые недостатки в определенных классах земной поверхности, дальнейшие исследования будут посвящены вопросам создания каталога обучающих выборок для поддержки успешной дифференциации объектов. В целом, полученные оценки точности демонстрируют современные характеристики разработанной модели, а также эффективность этой комбинации сети и наборов данных для тестовых регионов. Результаты показывают, что предложенный алгоритм может эффективно улучшить общую точность се-

мантической сегментации изображений дистанционного зондирования высокого пространственного разрешения и сократить время обучения сегментации.

Segmentation of aerospace images using convolutional neural networks

To solve practical problems, including for the management of territories, monitoring of emergencies, for the development of the ecology of our planet, for the study and ecological research of our planet, effective algorithms for the segmentation of multispectral images are needed. In recent years, image segmentation using convolutional neural networks has become very popular. A significant difference of this approach is the annotation of the assignment of each pixel to a particular class of objects, so that the learning process of such networks is completely controlled. *The paper proposes a new method for segmentation of aerospace images of high spatial resolution based on convolutional neural networks and mask generation. Our model is based on a combined U-Net network with MobileNetV2 as the backbone. It trains on ground data and provides a full prediction mask. A network has been implemented, consisting of separate networks of the same class, working on the expansion of segmentation. Semantic features are used to reduce errors at the semantic level of the outline. Since an independent test revealed some shortcomings in certain classes of the earth's surface, further research will be devoted to the creation of a catalog of training samples to support successful differentiation of objects. In general, the obtained accuracy estimates demonstrate the modern characteristics of the developed model, as well as the effectiveness of this combination of network and datasets for test regions. The results show that the proposed algorithm can effectively improve the overall accuracy of the semantic segmentation of high spatial resolution remote sensing images and reduce the training time and segmentation time.*

Гнатушенко Володимир Володимирович - д.т.н., професор, завідувач кафедри інформаційних технологій та комп'ютерної інженерії національного технічного університету «Дніпровська політехніка» (м. Дніпро).

Гнений Ігор Олексійович - аспірант кафедри програмного забезпечення комп'ютерних систем національного технічного університету «Дніпровська політехніка» (м. Дніпро).

Удовик Ірина Михайлівна - к.т.н., доцент, завідувач кафедри програмного забезпечення комп'ютерних систем національного технічного університету «Дніпровська політехніка» (м. Дніпро).

Шевцова Ольга Сергіївна - аспірант кафедри програмного забезпечення комп'ютерних систем національного технічного університету «Дніпровська політехніка» (м. Дніпро).

Гнатушенко Владимир Владимирович - д.т.н., профессор, заведующий кафедрой информационных технологий и компьютерной инженерии национального технического университета «Днепропетровская политехника» (г. Днепр).

Гненний Игорь Алексеевич - аспирант кафедры программного обеспечения компьютерных систем национального технического университета "Днепро-вская политехника" (г. Днепр).

Удовик Ирина Михайловна - к.т.н., доцент, заведующий кафедрой программного обеспечения компьютерных систем национального технического университета "Днепро-вская политехника" (г. Днепр).

Шевцова Ольга Сергеевна - аспирант кафедры программного обеспечения компьютерных систем национального технического университета "Днепро-вская политехника" (г. Днепр).

Hnatushenko Volodymyr - doctor of technical science, professor, head of department of information technologies and computer engineering, Dnipro University of Technology, Dnipro, Ukraine.

Gnennyi Igor - post-graduate student of the department of software for computer systems, Dnipro University of Technology, Dnipro, Ukraine.

Udovyk Iryna - candidate of technical sciences, associate professor, head of the department of software for computer systems, Dnipro University of Technology, Dnipro, Ukraine.

Shevtsova Olga - post-graduate student of the department of software for computer systems, Dnipro University of Technology, Dnipro, Ukraine.