

DOI 10.31891/2307-5732-2024-345-6-9
УДК 004.8, 004.67, 004.42

РОЗУМОВСЬКИЙ МИКОЛА

Національний лісотехнічний університет України
<https://orcid.org/0009-0003-7291-811X>
e-mail: mykolarozumovskyy@nltnu.edu.ua

КРОШНИЙ ІГОР

Національний лісотехнічний університет України
<https://orcid.org/0000-0003-0018-7998>
e-mail: kroshny.igor@nltnu.edu.ua

СЕГМЕНТАЦІЯ ТА КЛАСИФІКАЦІЯ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ SLIC SUPERPIXEL У ЛІСОВОМУ СЕРЕДОВИЩІ

Безпілотні літальні апарати (БПЛА) стали незамінним інструментом для збору високоточних геопросторових даних завдяки своїй доступності, мобільності та здатності забезпечувати надзвичайно високу деталізацію. На відміну від традиційних систем дистанційного зондування, таких як супутники або повітряні платформи, БПЛА дозволяють спостерігати за локалізованими територіями з високою частотою повторного збору даних. Це відкриває нові можливості для моніторингу змін навколишнього середовища, аналізу землекористування та ідентифікації окремих об'єктів, таких як дерева, завдяки надзвичайно високій роздільній здатності зображень.

Дослідження фокусується на автоматизованій ідентифікації дерев у лісовому масиві, представленому у вигляді ортофотоплану, створеного з 255 знімків високої роздільної здатності. Основним методом обробки даних є цифрова суперпиксельна сегментація з використанням алгоритму простої лінійної ітеративної кластеризації (SLIC). Цей підхід дозволяє групувати пікселі в компактні області (суперпикселі), що характеризуються однорідною текстурою або колірними властивостями, що значно спрощує подальшу класифікацію. Для дослідження було обрано три різні конфігурації сегментації: 2000, 3000 та 4000 суперпикселів. Параметри алгоритму, такі як масштаб згладжування $\sigma = 5$ та компактність $= 10$, були підібрані таким чином, щоб забезпечити оптимальну якість сегментації.

Класифікацію суперпикселів було виконано з використанням архітектури глибокої згорткової нейронної мережі ResNet-50. Ця модель була попередньо навчена на великому наборі зображень для розпізнавання загальних текстур і форм, після чого її ваги були адаптовані для класифікації дерев і фону на основі нового набору даних суперпиксельних зображень. Такий підхід, що поєднує попереднє навчання моделі та передачу знань, значно підвищив точність класифікації.

Під час експериментів було визначено, що найкращих результатів було досягнуто при сегментації 3000 суперпикселів, де точність класифікації склала 87%. Це свідчить про оптимальний баланс між деталізацією сегментації та здатністю моделі точно розпізнавати об'єкти. Використання меншої кількості суперпикселів призводило до втрати дрібних деталей, тоді як надмірна деталізація збільшувала обчислювальні витрати і погіршувала результати через недостатню чіткість об'єктів.

Ключові слова: безпілотні літальні апарати (БПЛА), згорткові нейронні мережі (CNN), зображення високої просторової роздільної здатності, сегментація лісових масивів.

RAZUMOVSKYY MYKOLA, KROSHNYI IHOR
National Forestry University of Ukraine

IMAGE SEGMENTATION AND CLASSIFICATION USING SLIC SUPERPIXEL IN A FOREST ENVIRONMENT

Unmanned aerial vehicles (UAVs) have become an indispensable tool for collecting high-precision geospatial data due to their affordability, mobility, and ability to provide extremely high detail. Unlike traditional remote sensing systems, such as satellites or airborne platforms, UAVs allow for observation of localized areas with a high frequency of repeated data collection. This opens up new possibilities for monitoring environmental changes, analyzing land use, and identifying individual objects such as trees due to the extremely high resolution of the images.

The study focuses on automated identification of trees within a forest area represented as an orthophoto map created from 255 high-resolution images. The main data processing method is digital superpixel segmentation using the Simple Linear Iterative Clustering (SLIC) algorithm. This approach allows grouping pixels into compact regions (superpixels) characterized by homogeneous texture or color properties, which greatly simplifies further classification. Three different segmentation configurations were chosen for the study: 2000, 3000, and 4000 superpixels. The algorithm parameters, such as smoothing scale $\sigma = 5$ and compactness $= 10$, were selected to ensure optimal segmentation quality.

Superpixel classification was performed using the architecture of the ResNet-50 deep convolutional neural network. This model was pre-trained on a large set of images to recognize common textures and shapes, after which its weights were adapted to classify trees and backgrounds based on the new superpixel image dataset. This approach, which combines model pre-training and knowledge transfer, significantly improved classification accuracy.

During the experiments, it was determined that the best performance was achieved when segmenting 3000 superpixels, where the classification accuracy was 87%. This indicates an optimal balance between segmentation detail and the model's ability to accurately recognize objects. Using a smaller number of superpixels resulted in the loss of fine details, while excessive detail increased computational costs and degraded the results due to insufficient object clarity.

Keywords: unmanned aerial vehicles (UAVs), convolutional neural networks (CNNs), high spatial resolution images, forest segmentation.

1. Постановка проблеми

Безпілотні літальні апарати (БПЛА) є універсальною та відносно недорогою платформою, яка дозволяє швидко й зручно отримувати дані з обраної території або об'єкта відповідно до потреб користувача. Вони забезпечують можливість картографування та моніторингу невеликих і середніх територій (деякі моделі БПЛА можуть легко охоплювати площі до 350 гектарів) з надзвичайно високою деталізацією. Це також дає змогу багаторазово збирати дані з однієї і тієї ж ділянки в різні періоди часу. Високий рівень просторових

деталей, який забезпечують БПЛА, перевершує традиційні системи дистанційного зондування, такі як супутники або авіаційні системи, дозволяючи виявляти специфічні особливості в різноманітних застосуваннях [1][2][3][4].

Останнім часом глибокі згорткові нейронні мережі (CNN) демонструють значні досягнення в таких задачах, як класифікація зображень [5][6], виявлення об'єктів [7] і сегментація зображень [8].

Методи суперпікселів, уперше запропоновані у [9], дають змогу групувати пікселі у сприйнятливо значущі області. Такі області містять більше інформації, ніж окремі пікселі, оскільки представляють собою релевантні групи пікселів [10]. Програмне забезпечення Pynovisão [11] поєднує техніку комп'ютерного зору Simple Linear Iterative Clustering (SLIC) Superpixel [12] із методами опорних векторів (SVM) [13] та конволюційних нейронних мереж (CNN) [14].

Це дослідження аналізує здатність програмного забезпечення Pynovisão класифікувати лісові ортофотознімки високої просторової роздільної здатності.

2. Методи

2.1 Опис досліджуваної території та польоти БПЛА

Дослідження проводилися весною 2024 року в автоматичному режимі з використанням програмного забезпечення Pix4DCapture та БПЛА DJI Mavic 3 Pro. Цей дрон оснащений стандартним 1-дюймовим CMOS-сенсором із роздільною здатністю 20 МП (5280x3956 пікселів), який забезпечує просторову роздільну здатність Ground Sample Distance (GSD) приблизно 27 мм/піксель при висоті польоту 50 м над рівнем землі (AGL).

Досліджувана територія — лісова ділянка з різною щільністю крон дерев і значною варіативністю розмірів деревних крон. Це робить її ідеальним прикладом для демонстрації методів сегментації, реалізованих у програмному забезпеченні Pynovisão.

2.2 Генерація ортогонального зображення лісової території

Для цього дослідження ортогональне зображення було створено за допомогою програмного забезпечення Pix4Dmapper [15], використовуючи 255 зображення, отримані з БПЛА. Перекриття під час зйомки становило 90% як у поздовжньому, так і в поперечному напрямках. У результаті було сформовано композицію зображень у форматі GeoTIFF з роздільною здатністю 11283x8671x4 пікселів і об'ємом 284 МБ для заданої області.

Для орторектифікації було використано 6 контрольних точок на місцевості (GCP) та 16 точок перевірки (CP). Точність цих 16 CP оцінювалася за середньоквадратичною похибкою (RMSE): 7 см у координаті L (схід), 4 см у координаті N (північ) і 9 см у координаті Z.

2.3 Сегментація зображень

Суперпікселі забезпечують сегментацію зображення шляхом формування компактних і однорідних груп пікселів, які мають схожі характеристики, наприклад, за кольором або геометрією [10]. Для позначення крон дерев на зображеннях використовувалося програмне забезпечення LabelMe [16], як показано на Рис. 1.

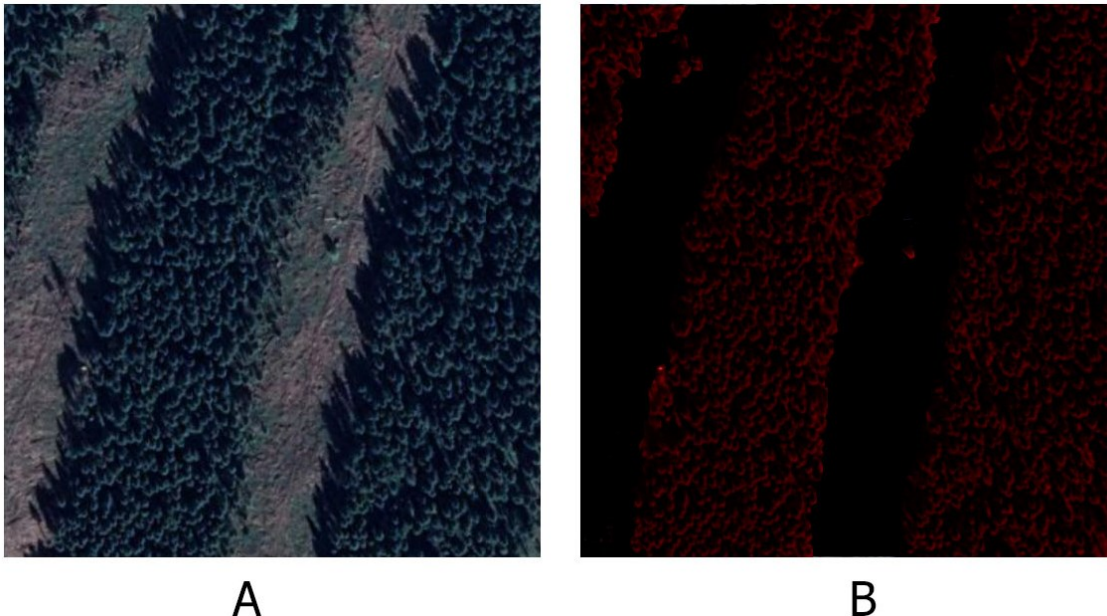


Рис. 1. Ортофотоплан потрібної ділянки дослідження.
Згенерована за допомогою програми LabelMe, для подальшого аналізу точності

Програмне забезпечення LabelMe дозволяє вручну анотувати об'єкти на зображеннях, створюючи маску, яка має такі самі розміри, як оригінальне зображення. Ця маска розділяє об'єкти заданого класу на окремі кольорові області.

Створені маски підтримуються програмним забезпеченням Pynovisão, яке автоматично співвідносить суперпікселі з відповідними регіонами певного класу. У результаті формується анотований набір даних, у

якому кожен суперпіксель отримує свою мітку.

2.4 Виявлення дерев

Процес виявлення дерев включав використання глибоких конволюційних нейронних мереж у поєднанні з методом SLIC Superpixel для сегментації ортогонального зображення та подальшої класифікації. Розглядалися два класи: дерева і фон.

Для сегментації суперпікселів у методі SLIC використовувалися параметри: складність (σ) — 5, компактність — 10. Кількість суперпікселів (K) варіювалася між 2000, 3000 і 4000. Для вирівнювання кількості прикладів між класами застосовувався метод DownSampling, що зменшував кількість прикладів у класі з переважною кількістю. Для кожного класу було враховано 1179, 1837 і 2538 суперпікселів для 2000, 3000 і 4000 суперпікселів відповідно.

Гіперпараметри глибокої згорткової мережі залишалися за замовчуванням:

- архітектура — ResNet-50;
- оптимізатор Gradient Descent із швидкістю навчання 0,01 та імпульсом 0,9;
- розмір — 32;
- кількість епох — 50.

Для навчання також використовували метод передачі знань (Transfer Learning) зі 100% тонким налаштуванням (fine-tuning), де мережу повторно навчали, використовуючи ваги ResNet-50, натреновані на ImageNet [17]. Вихідний розмір мережі було встановлено на 256x256x3 пікселі.

Набір зображень тренувався алгоритмом CNN через інтерфейс Rynovisão. Для навчання використовувалися залежності Keras, TensorFlow і CUDA. Зображення було випадковим чином розділено на дві частини: 80% для навчання і 20% для валідації. Для оцінки точності моделі використовувався глобальний параметр точності.

3. Результати

Під час використання CNN для класифікації із значеннями K (кількість суперпікселів) — 2000, 3000 і 4000, а також застосуванням 100% тонкого налаштування (fine-tuning) у процесі передачі знань, результати показали, що збільшення кількості суперпікселів із 2000 до 3000 сприяло покращенню точності класифікації. Однак при подальшому збільшенні кількості суперпікселів до 4000 спостерігалось невелике зниження точності.

Таблиця 1

Кількість суперпікселів за класами та точність класифікації на основі маски, згенерованої за допомогою LabelMe

SUPERPIXELS SLIC NUMBER (K)	2000	3000	4000
ДЕРЕВА	1179	1837	2538
ФОН	560	767	987
ТОЧНІСТЬ	81.50%	87.39%	82.01%

На рисунках 2, 3 і 4 представлені результати роботи програмного забезпечення Rynovisão для значень K : 2000, 3000 і 4000 суперпікселів відповідно. По осях X і Y показано кількість пікселів у зображеннях. Червоні області відповідають кронам дерев, а зелені — фону. Метод SLIC формує суперпікселі регулярної форми в нетекстурованих областях і нерегулярної форми в текстурованих областях [10].

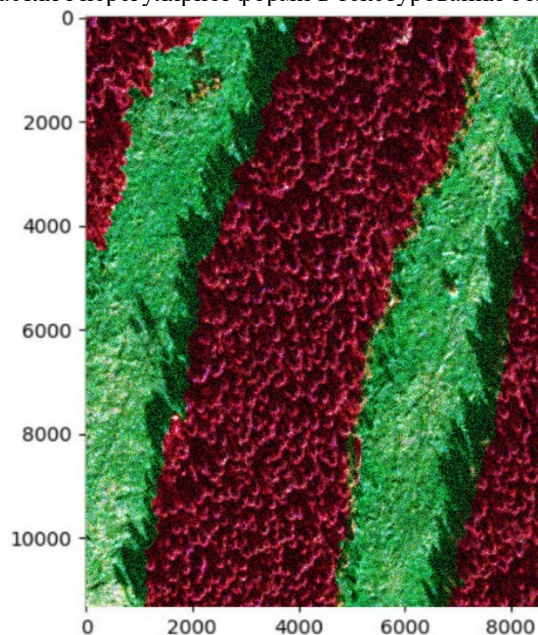


Рис. 2. Вихід Rynovisão з $k=2000$ на SLIC Superpixels накладений на ортознімок лісу з БПЛА з відстанню до землі 0.02 відстанню вибірки GSD

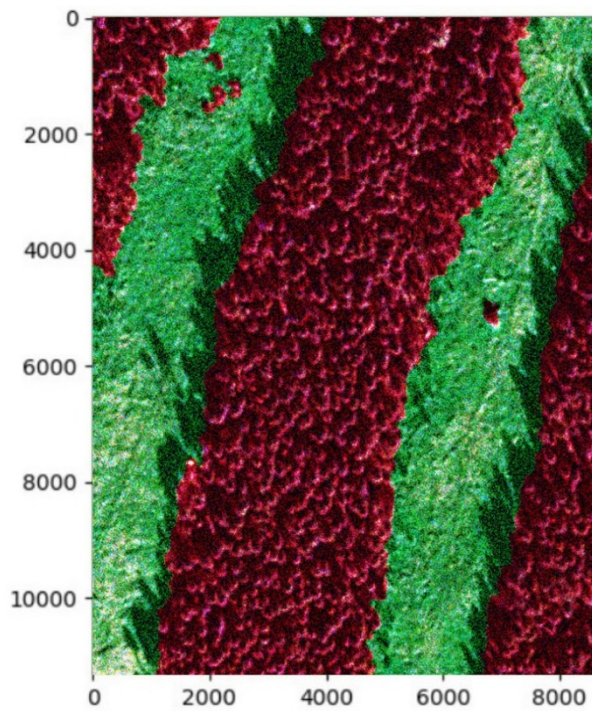


Рис. 3. Вихід Руповишão з $k=3000$ на SLIC Superpixels накладений на ортознімок лісу з БПЛА з відстанню до землі 0.02 відстанню вибірки GSD

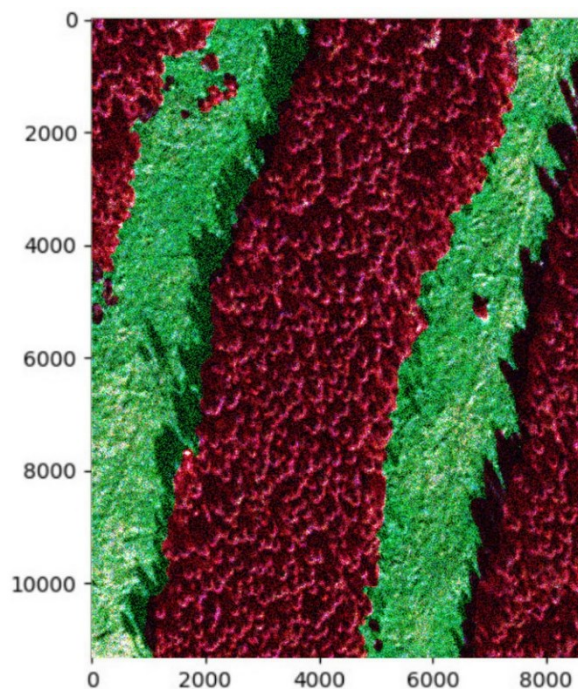


Рис. 4. Вихід Руповишão з $k=4000$ на SLIC Superpixels накладений на ортознімок лісу з БПЛА з відстанню до землі 0.02 відстанню вибірки GSD

У межах цього дослідження було проведено три експерименти на трьох збалансованих наборах даних для виявлення крон дерев на зображеннях, отриманих із БПЛА. Для сегментації зображень застосовувався метод SLIC Superpixel, а для класифікації суперпікселів — CNN.

Результати експериментів були оптимістичними, із мінімальною точністю 81,50% для 2000 суперпікселів при розрізненні крон дерев і фону.

4. Висновки

У цьому дослідженні було підтверджено, що програмне забезпечення Руповишão, яке базується на методах суперпікселів і згорткових нейронних мереж (CNN), здатне досягати високої точності при класифікації дерев, використовуючи високороздільні зображення, отримані з БПЛА.

Подальші дослідження можуть бути спрямовані на зйомку зображень із більшої території для кращої узагальненості вирішення задачі. Ще одним перспективним напрямком є тестування гіперпараметрів CNN, що може покращити продуктивність моделі для цього типу задач.

References

1. Zarco-tejada. Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods. / Pablo J., et al. // *European journal of agronomy* – 2014 – № 55 – P. 89-99.
2. Mohan. Individual tree detection from unmanned aerial vehicle (UAV) derived canopy height model in an open canopy mixed conifer forest. / Midhun, et al. // *Forests* – 2017 – №8.9 – P. 340.
3. Bazi Yakoub. Convolutional SVM networks for object detection in UAV imagery. / Farid Melgani // *Ieee transactions on geoscience and remote sensing* – 2018 – № 56.6 – P. 3107-3118.
4. LI Dong. Individual tree delineation in windbreaks using airborne-laser-scanning data and unmanned aerial vehicle stereo images. / H., Wang, C., Li, W., Chen, H., & Zuo, Z. // *IEEE Geoscience and Remote Sensing Letters* – 2016 – № 13.9 – P. 1330-1334.
5. He K. Spatial pyramid pooling in deep convolutional networks for visual recognition. / Zhang, X., Ren, S., & Sun, J // *IEEE transactions on pattern analysis and machine intelligence* – 2015 – № 37.9 – P. 1904-1916.
6. Krizhevsky. Imagenet classification with deep convolutional neural networks. / Alex, Ilya Sutskever, and Geoffrey E. Hinton // *Advances in neural information processing systems* – 2012 – № 25.
7. Ren, S. Object detection networks on convolutional feature maps. / He, K., Girshick, R., Zhang, X., & Sun, J // *IEEE transactions on pattern analysis and machine intelligence* – 2016 – № 39.7 – P. 1476-1481.
8. Long Jonathan. Fully convolutional networks for semantic segmentation. / Evan Shelhamer and Trevor Darrell // *In: Proceedings of the IEEE conference on computer vision and pattern recognition* – 2015 – P. 3431-3440.
9. Ren. Learning a classification model for segmentation. / Malik // *In: Proceedings ninth IEEE international conference on computer vision. IEEE* – 2003 – P. 10-17 .
10. Crommelinck, S.. SLIC superpixels for object delineation from UAV data. / Bennett, R., Gerke, M., Koeva, M. N., Yang, M. Y., & Vosselman, G. // *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* – 2017 – № 4 – P. 9-16.
11. dos Santos Ferreira, A.. Weed detection in soybean crops using ConvNets. / Freitas, D. M., da Silva, G. G., Pistori, H., & Folhes, M. T // *Computers and Electronics in Agriculture* – 2017 – № 143 – P. 314-324.
12. Achanta, R. SLIC superpixels compared to state-of-the-art superpixel methods. / Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. // *IEEE transactions on pattern analysis and machine intelligence* – 2012 – № 34.11 – P. 2274-2282.
13. Bonyadi, M. R. Optimal-margin evolutionary classifier. / Reutens, D. C. // *IEEE Transactions on Evolutionary Computation* – 2019 – № 23.5 – P. 885-898.
14. Silva, J. Convolutional neural network with multi-column characteristics extraction for image classification. / Varela, N., Patiño-Saucedo, J. A., & Lezama, O. B. P. // *In: Image Processing and Capsule Networks: ICIPCN 2020. Springer International Publishing* – 2021 – P. 20-30.
15. Strecha, C. Automatic mapping from ultra-light UAV imagery. / Küng, O., & Fua, P. // *In: EuroCOW 2012* – 2012.
16. Russell, B. C. LabelMe: a database and web-based tool for image annotation. / Torralba, A., Murphy, K. P., & Freeman, W. T. // *International journal of computer vision* – 2008 – № 77 – P. 157-173.
17. Shafiq, M. Deep residual learning for image recognition. / Gu, Z. // *A survey. Applied Sciences* – 2022 – № 12.18 – P. 8972.