

Олег Служала, Олександр Овсяк*

¹ НЛТУ України, Львів, Україна, sluzhala.oleh@nltu.edu.ua

² НЛТУ України, Львів, Україна, ovsjak@ukr.net

Порівняння одноетапних та двоетапних моделей комп'ютерного зору для автоматизованого обліку водних біоресурсів

***Анотація.** Здійснено порівняння підходів візуального розпізнання. Було порівняно два основні підходи для аналізу зображення та виявлено їхні переваги та недоліки. Результати цього дослідження можуть бути використані при створенні інформаційних систем, орієнтованих на візуальний аналіз біоресурсів.*

***Ключові слова** – математична модель, розпізнавання об'єктів, OpenCV, YOLO.*

Вступ. Сучасне управління водними біоресурсами вимагає високоточної, адаптивної та масштабованої аналітики, що здатна виявляти закономірності в гетерогенних наборах екологічних, біологічних та ринкових даних. Інтелектуальні обчислення - це сукупність методів штучного інтелекту, машинного навчання, обробки великих даних та візуалізації, яка дозволяє ефективно аналізувати великі обсяги інформації, що надходить з екологічних сенсорів, супутників, риболовних звітів і наукових спостережень.

Традиційні методи аналізу екологічної інформації виявились недостатньо ефективними у роботі з великими обсягами неструктурованих даних, які надходять із датчиків, супутників, гідробіологічних досліджень і автоматизованих риболовних систем. У зв'язку з цим актуальним є впровадження математичних моделей, та систем візуального аналізу біоресурсів природних систем. [1].

Математична модель. YOLO, як представник одноетапного підходу, розбиває зображення на регулярну сітку $S \times S$, де кожна клітинка прогнозує: B обмежувальних рамок (bounding boxes), клас об'єкта та ймовірність наявності об'єкта (objectness score)

Вихідний тензор має вигляд:

$$Output = S \times S \times (B \cdot Z + C) \quad (1)$$

де: Z – координати (x, y, w, h), а C – кількість класів [2].

Переваги одноетапного підходу: висока швидкодія - до 60+ FPS, підходить для реального часу, мала обчислювальна складність - можлива реалізація навіть на вбудованих пристроях, гнучкість - легка адаптація до нових класів чи середовищ.

Недоліки: нижча точність, ніж у двоетапних моделей, проблеми з перекриттям - складно виявити об'єкти, що накладаються, низька деталізація, особливо для дрібних об'єктів у фоні.

Суть двоетапного підходу полягає у послідовному виконанні двох основних процедур: генерація регіонів-претендентів (Region Proposal Network, RPN) - спеціальна мережа прогнозує області з високою ймовірністю наявності об'єктів. RPN повертає обмежувальні рамки (bounding boxes), які називають "пропозиціями" (proposals) та класифікація та регресія меж - для кожної рамки модель визначає клас об'єкта та уточнює координати рамки. Це виконується за допомогою окремого класифікатора та регресійної голови.

Формально, вихідну ймовірність можна представити як:

$$P(c, b | I) = P(b | r, I) \cdot P(c | r, I) \quad (2)$$

де $P(b | r, I)$ - функція уточнення рамки, а $P(c | r, I)$ - ймовірність належності до класу c , залежно від зображення I і регіону r .

Модель Faster R-CNN включає такі компоненти: Backbone (наприклад, ResNet-50) - виділяє ознаки з зображення, RPN - генерує набори рамок, RoI Pooling - перетворює фрагменти зображення в єдину форму та класифікатор + регресор - визначають клас та координати [3].

Переваги двоетапного підходу: надає високу точність Mean Average Precision (mAP), гнучкість - є можливість точної класифікації за багатьма класами та стійкість до складних сцен, адже добре справляється з перекриттями об'єктів і слабкими контурами [4].

Недоліки: висока обчислювальна складність, і як наслідок - потреба у потужному GPU та доволі низька швидкодія в 5–10 кадрів/секунда - недостатньо для відео в реальному часі.

Результати порівняння. Для емпіричного підтвердження ефективності різних підходів до детекції водних біоресурсів, було проведено порівняльний експеримент, у якому тестувалися моделі типу YOLOv8 (одноетапний підхід) та Faster R-CNN (двоетапний підхід). Основною метою дослідження було оцінити якісні та кількісні характеристики роботи моделей на уніфікованому наборі даних, що відображає реальні умови зйомки у водному середовищі: різноманітність освітлення, прозорість води, перекриття об'єктів, зміни масштабу, наявність частково видимих або зашумлених об'єктів. точність класифікації, так і швидкодію моделей у реальному часі. [5]

Таблиця 1. Порівняння одноетапного та двоетапного підходу

Метрика	mAP@0.5	mAP@[.5:.95]	Precision	Recall	Time
YOLOv8	89.3%	69.8%	88.1%	86.5%	29 мс
FR-CNN	92.4%	73.5%	91.2%	90.7%	125 мс

Після завершення навчання результати виявили характерні властивості обох підходів. Згідно з отриманими показниками, YOLOv8 продемонстрував mAP@0.5 на рівні 89.3%, а mAP@[.5:.95] - 69.8%, при середній швидкості обробки в (29 мс) 34 FPS на тестовому наборі. Натомість Faster R-CNN досяг mAP@0.5 = 92.4%, а mAP@[.5:.95] склав 73.5%, проте його швидкодія не перевищувала (125 мс) 8 FPS, що є критичним у контексті систем реального часу, проте прийнятним для офлайн-аналізу.

Висновок. У ході дослідження було розроблено, реалізовано та експериментально протестовано програмно-математичну систему для автоматизованого аналізу та обліку водних біоресурсів на основі комп'ютерного зору й алгоритмів глибокого навчання. Основною технологічною базою стали бібліотека OpenCV для обробки зображень та глибокі нейронні мережі одноетапної (YOLOv8) та двоетапної (Faster R-CNN) архітектур для задач детекції об'єктів у відеопотоці. Порівняння одноетапних і двоетапних моделей засвідчило, що компроміс між швидкістю та точністю залишається ключовим фактором при виборі архітектури для конкретного завдання. Встановлено, що одноетапні мо-

делі є ефективними в реальному часі завдяки прямому підходу до регресії координат об'єктів і класів, тоді як двоетапні архітектури, розділяючи процес генерації регіонів інтересу та класифікації, забезпечують вищу структурну точність за рахунок зменшення швидкості обробки.

Таким чином, у залежності від прикладного контексту – оперативний облік, довготривалий моніторинг, наукові спостереження або екологічна аналітика – рекомендовано застосовувати гібридні рішення, що поєднують переваги обох підходів. Результати даного дослідження можуть бути використані для побудови інтелектуальних систем автоматизації екологічного контролю.

Список використаних літературних джерел

- [1] Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). YOLOv7: Trainable Bag-of-Freebies and Coarse-to-Fine Training for Real-Time Object Detection. arXiv preprint arXiv:2207.02696.
- [2] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In Advances in Neural Information.
- [3] Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2022). A Survey of Modern Deep Learning-Based Object Detection Algorithms. IEEE Access, 10, 798–820
- [4] Kumar, A., & Singh, P. (2023). Performance Comparison of YOLOv5 and Faster R-CNN for Real-Time Object Detection. arXiv preprint arXiv:2305.08745.
- [5] Zhou, F., & Zhang, L. (2022). Attention-Enhanced Faster R-CNN for Small Object Detection. Sensors, 22(11), 4056.

Comparative analysis of one-stage and two-stage computer vision models for automated accounting of aquatic biological resources

Shuzhala Oleh, Oleksandr Ovsiak

A comparison of visual recognition approaches was made. Two main approaches to image analysis were compared and their advantages and disadvantages identified. The results of this study can be used in the creation of information systems focused on the visual analysis of bioresources.