

УДК 004.8

DOI <https://doi.org/10.32782/EIS/2024-106-15>

ПОРІВНЯННЯ МЕТОДІВ ГЛИБОКОГО НАВЧАННЯ В ЗАДАЧАХ РОЗПІЗНАВАННЯ ЛІТАКІВ ЗА ДАНИМИ АЕРОКОСМІЧНОЇ ЗЙОМКИ

Каштан Віта Юріївна,

кандидат технічних наук, доцент,
доцент кафедри інформаційних технологій та комп'ютерної інженерії
Національного технічного університету «Дніпровська політехніка»
ORCID ID: 0000-0002-0395-5895
SCOPUS AUTHOR ID: 57201902879

Радіонов Євген Дмитрович,

аспірант кафедри інформаційних технологій та комп'ютерної інженерії
Національного технічного університету «Дніпровська політехніка»
ORCID ID: 0009-0002-2839-7161

Мета – дослідження методів глибокого навчання в задачах розпізнавання літаків за даними аерокосмічної зйомки.

Методи. У ході дослідження проведено порівняльний аналіз методів глибокого навчання для розпізнавання літаків, зокрема Faster R-CNN, SSD, SSD Lite, FCOS та RetinaNet, на спеціалізованому наборі даних, що складається з 2120 зображень високої роздільної здатності й містить 14335 анованих літаків. Для кожної з моделей розраховано метрики точності Precision, Recall, mAP50-95 та mAP50.

Результати. Отримані результати дослідження вказують на те, що моделі Faster R-CNN та FCOS забезпечують найвищі значення за метриками mAP50-95 та mAP50. Для моделі Faster R-CNN значення метрики mAP50-95 становило 0,786, а mAP50 – 0,985, для моделі FCOS – 0,789 та 0,984 відповідно. Це свідчить про високу точність у виявленні контурів об'єктів. Моделі RetinaNet і SSD, хоча й показують дещо нижчі результати (0,746 і 0,617 за метрикою mAP50-95 відповідно), зберігають високі показники Precision та Recall, що підтверджує їх ефективність. Для моделі RetinaNet показник Precision становив 0,98, а Recall – 0,91, а для моделі SSD – 0,99 і 0,71 відповідно. Модель SSD Lite продемонструвала найнижчі результати за метриками mAP50-95 (0,275) та mAP50 (0,520), проте її висока швидкість роботи та задовільні значення Precision (0,98) роблять її перспективною для застосування в задачах з обмеженими ресурсами або в режимі реального часу.

Новизна. Виконано порівняльний аналіз продуктивності різних моделей глибокого навчання для задачі розпізнавання літальних апаратів на аерокосмічних зображеннях високої роздільної здатності. Встановлено, що моделі Faster R-CNN та FCOS забезпечують найвищу точність серед протестованих алгоритмів, тоді як модель SSD Lite демонструє оптимальне співвідношення між точністю та швидкістю оброблення, що робить її придатною для застосувань з обмеженими ресурсами або в режимі реального часу.

Цінність. Результати дослідження показали, що моделі Faster R-CNN та FCOS завдяки високим показникам mAP50-95 і mAP50 є найбільш ефективними для задач, де основним пріоритетом є точність виявлення. Це робить їх доцільними для використання у військових або безпекових системах моніторингу, де мінімізація хибних спрацьовувань і підвищення надійності виявлення об'єктів є важливими. Модель SSD Lite, зі своєю вищою швидкістю оброблення, є оптимальним вибором для систем безпілотних літальних апаратів, мобільних платформ або інших додатків, де час реакції є ключовим фактором.

Ключові слова: машинне навчання, розпізнавання образів та контурів, попереднє оброблення оптичних зображень, знімки високої роздільної здатності, розпізнавання літаків.

Kashtan Vita, Radionov Yevhen. Comparison of deep learning methods in aircraft recognition tasks using aerospace imaging data

Purpose. Study of deep learning methods in aircraft recognition tasks using aerospace imaging data.

Methods. During the research, a comparative analysis of deep learning methods for aircraft recognition, specifically Faster R-CNN, SSD, SSD Lite, FCOS, and RetinaNet, was conducted on a specialized dataset consisting of 2,120 high-resolution images containing 14,335 annotated aircraft. Precision, Recall, mAP50-95, and mAP50 metrics were calculated for each model.

Results. The study results indicate that the Faster R-CNN and FCOS models achieve the highest values for the metrics mAP50-95 and mAP50. The Faster R-CNN model's mAP50-95 metric value was 0.786, and mAP50 was 0.985. For the FCOS model, these values were 0.789 and 0.984, respectively. This indicates high accuracy in object boundary detection. While the RetinaNet and SSD models show slightly lower results (0.746 and 0.617 for mAP50-95, respectively), they maintain high Precision and Recall scores, confirming their effectiveness. For

RetinaNet, Precision was 0.98 and Recall was 0.91, while for SSD, these values were 0.99 and 0.71, respectively. The SSD Lite model demonstrated the lowest performance for mAP50-95 (0.275) and mAP50 (0.520); however, its high processing speed and satisfactory Precision (0.98) make it promising for use in resource-constrained or real-time applications.

Originality. *A comparative analysis of various deep learning models was conducted to recognize aircraft in high-resolution aerospace imagery. The study found that the Faster R-CNN and FCOS models exhibited the highest accuracy among those tested, while the SSD Lite model achieved the best balance between accuracy and speed.*

Practicality. *The research results showed that the Faster R-CNN and FCOS models, due to their high mAP50-95 and mAP50 scores, are the most effective for tasks where detection accuracy is the main priority. This makes them suitable for use in military or security monitoring systems, where minimizing false alarms and increasing object detection reliability are critical. The SSD Lite model, with its higher processing speed, is the optimal choice for unmanned aerial systems, mobile platforms, or other applications where response time is a key factor.*

Key words: *machine learning, image and contour recognition, optical image preprocessing, high-resolution imagery, aircraft detection.*

Актуальність. У сучасному світі розпізнавання літаків за допомогою глибокого навчання набуває дедалі більшого значення в різних галузях, включаючи цивільну авіацію, військову оборону, моніторинг повітряного простору, екологічні дослідження та системи безпеки. Із щораз більшою кількістю даних із супутникових та аерофотознімків автоматизація процесу виявлення та класифікації літальних апаратів стає важливою для підвищення точності, оперативності та ефективності аналізу.

Наявні методи розпізнавання літаків на основі глибоких нейронних мереж пропонують різні підходи, кожен з яких має свої переваги та обмеження. Традиційні підходи, як-от використання згорткових нейронних мереж (CNN), продемонстрували суттєві досягнення в розв'язанні задач розпізнавання об'єктів. Однак задачі, пов'язані з високою варіативністю фонових сцен, нечіткими контурами об'єктів та змінними умовами освітлення, залишаються значними викликами для наявних методів розпізнавання. Такі моделі, як Faster R-CNN, RetinaNet, SSD та їх модифікації, демонструють різний рівень ефективності, залежно від специфіки поставлених завдань і типу вхідних даних. При цьому актуальною є проблема оптимізації співвідношення між точністю розпізнавання контурів літаків та швидкістю оброблення зображень.

Важливим аспектом, що підкреслює актуальність цього дослідження, є необхідність систематичного огляду та порівняння наявних методів для виявлення їхніх слабких місць та потенційних напрямків для подальшого розвитку. Це дозволить не лише підвищити ефективність розпізнавання літаків, але й сприятиме розвитку нових підходів у сфері глибокого навчання, що зможуть адаптуватися до змінних умов даних. Тому проведення науково-технічного огляду та порівняння наявних методів розпізнавання літаків з використанням глибокого навчання є актуальним і важливим етапом у напрямі створення більш точних і надійних

систем автоматизованого виявлення літальних апаратів у реальних умовах.

Аналіз літературних джерел. Згідно з дослідженням [1, с. 7778] використання Faster R-CNN для аерокосмічних зображень дозволяє досягати високих показників точності, особливо в разі виявлення літаків на зображеннях із високим рівнем шуму та варіативним фоном.

Автори в дослідженні [2, с. 458] демонструють обмеження моделей Faster R-CNN та SSD за виявлення літаків на великомасштабних зображеннях. Незважаючи на високу точність, зафіксовану авторами на загальних наборах даних, моделі демонструють зниження ефективності під час роботи з даними, що містять великомасштабні зображення літаків.

У статті [3, с. 52] автори покращили модель FCOS для розпізнавання дрібних об'єктів з даних безпілотних літальних апаратів, що дало приріст у точності на 6%, а у швидкості – на 3 кадри за секунду.

У роботі [4, с. 53] авторами запропоновано оптимальну модель з погляду співвідношення точності розпізнавання швидкодії на основі RetinaNet фреймворку в задачах виявлення пускових шахт ракет на аерокосмічних зображеннях.

Аналіз літературних джерел свідчить про активний розвиток методів розпізнавання літаків за допомогою глибокого навчання. Моделі Faster R-CNN, SSD, FCOS та RetinaNet мають свої переваги та обмеження, що робить їх вибір залежним від конкретних умов застосування. Faster R-CNN є оптимальним для досягнення високої точності, тоді як SSD та FCOS можуть забезпечити високу продуктивність для задач реального часу. RetinaNet є ефективною для розпізнавання малих об'єктів на складних фонах.

Мета роботи – дослідження методів глибокого навчання в задачах розпізнавання літаків за даними аерокосмічної зйомки. Для досягнення мети поставлено декілька завдань:

– дослідження і вибір моделей, що підходять для розпізнавання літаків за даними аерокосмічної зйомки;

– вибір належного набору даних для порівняння роботи моделей;

– оцінювання й аналіз показників точності вибраних моделей.

Матеріали та методи. У ході дослідження проведено порівняльний аналіз методів глибокого навчання для розпізнавання літаків, зокрема Faster R-CNN, SSD, SSD Lite, FCOS та RetinaNet, на спеціалізованому наборі даних HRPlanesv2 [5]. Цей набір даних складається з 2120 зображень високої роздільної здатності, які містять загалом 14 335 анотованих літаків. Для забезпечення належного навчання та оцінювання моделей набір даних було поділено на три категорії: навчальний набір, що включає 1490 зображень, валідаційний набір із 420 зображень та тестовий набір, який складається з 210 зображень. Така структура набору даних дозволила адекватно оцінити продуктивність моделей, забезпечивши незалежну перевірку як на етапі валідації, так і на фінальному тестуванні.

Як початкові було вибрано попередньо натреновані моделі на дата-сеті COCO [6].

Для порівняння точності отриманих результатів вибрано такі метрики: *Precision*, *Recall*, *mAP* та середню точність.

Precision – це частка виявлень успішних передбачень серед усіх зроблених передбачень.

$$Precision = \frac{TP}{TP + FP}, \quad (1)$$

де TP – true positive, FP – false positive.

Recall – це частка успішних передбачень серед усіх можливих передбачень.

$$Recall = \frac{TP}{TP + FN}, \quad (2)$$

де TP – true positive, FN – false negative.

mAP – це комбінація значень *Precision* та *Recall*, обчислених за кількома пороговими значеннями достовірності, які також називають порогом перетину через об'єднання (Intersection over Union). Зміна порогового значення *IoU* призведе до різних істинних і хибних позитивних результатів.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i, \quad (3)$$

де N – кількість класів об'єктів, AP_i – середня точність для i -го класу. Прийнято позначати, які порогові значення *IoU* використовуються цифрами після *mAP*, наприклад, *mAP50-95*

використовує діапазон порогових значень *IoU* від 0,5 до 0,95 з розміром кроку 0,05, тоді як *mAP95* використовує поріг *IoU* 0,95.

Середня точність (*AP*) визначається шляхом спочатку сортування всіх виявлених об'єктів у порядку спадання рівня впевненості *conf*, потім – розрахунку *Precision* та *Recall* за кожним порогом і знаходження області під кривою точності – запам'ятовування шляхом інтегрування через усі можливі порогові значення.

$$AP = \int_{r=0}^1 p(r) dr \quad (4)$$

Результати дослідження. Результати точності розпізнавання літаків на аерокосмічних зображеннях високої роздільної здатності для досліджених моделей наведено в таблиці 1. Показники в таблиці 1 відібрані для передбачень моделей з рівнем впевненості $conf \geq 75\%$

Отримані результати дослідження свідчать про те, що моделі Faster R-CNN і FCOS демонструють найвищі показники за метриками *mAP50-95* та *mAP50*, що вказує на високу точність у розпізнаванні контурів об'єктів. Ці моделі також відзначаються високими значеннями *Precision* і *Recall*, що підкреслює їх ефективність як у правильному виявленні літаків, так і у мінімізації хибних спрацьовувань.

Таблиця 1

Результати точності розпізнавання літаків для різних моделей

Модель	mAP50-95	mAP50	Precision	Recall
Faster R-CNN	0.786	0.985	0.92	0.98
SSD	0.617	0.925	0.99	0.71
SSD Lite	0.275	0.520	0.98	0.31
FCOS	0.789	0.984	0.98	0.31
RetinaNet	0.746	0.970	0.98	0.91

До другої групи можна віднести моделі RetinaNet та SSD, які, хоча і демонструють дещо нижчі результати, все ж зберігають досить високі показники *Precision* і *Recall*. Однак варто зазначити, що модель SSD значно гірше справляється з точним виявленням меж об'єкта, що може вплинути на її ефективність у задачах, де критично важливе точне визначення форми літака.

Найгірші показники за метрикою *mAP* спостерігаються у моделі SSD Lite. Проте варто підкреслити, що її ключова перевага полягає у високій швидкості оброблення та прийнятному значенні *Precision*. Це робить SSD Lite придатною для застосувань, де точне визначення меж об'єкта не є критичним фактором, а перевага надається швидкій і точній класифікації об'єктів.

Таким чином, цю модель можна розглядати як опцію для реальних застосувань, які мають обмежені обчислювальні ресурси або вимагають швидкого прийняття рішень, наприклад, як моніторинг повітряного простору в режимі реального часу.

На рис. 1 представлено результати розпізнавання літаків для моделей Faster R-CNN (а) і FCOS (б), а на рис. 2 – для моделей RetinaNet (а) і SSD (б). Синім кольором показано еталонні межі об'єктів, а червоним – межі, передбачені відповідною моделлю. Цифри в дужках вказують на рівень упевненості моделі в класифікації об'єкта в межах від 0 до 1.

Ці результати дозволяють зробити висновок про те, що вибір моделі для розпізнавання літаків залежить від конкретних вимог завдання: точність контурів або швидкість класифікації, що дозволяє оптимізувати процес залежно від умов і ресурсів.

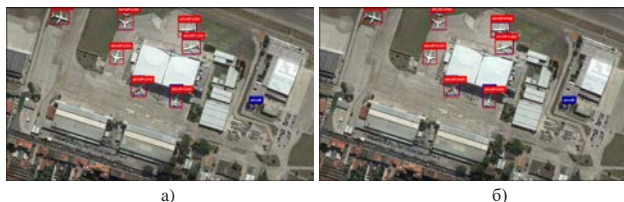


Рис. 1. Результати розпізнавання літаків моделей: а) Faster R-CNN; б) FCOS



Рис. 2. Результати розпізнавання літаків моделей: а) RetinaNet; б) SSD

Дискусія.

Отримані результати свідчать про те, що моделі Faster R-CNN і FCOS є найбільш перспективними для задач розпізнавання літаків завдяки їх високим показникам за метриками mAP_{50-95} та mAP_{50} . Це свідчить про те, що ці моделі здатні точно ідентифікувати об'єкти різних розмірів та форм, забезпечуючи високу точність виявлення контурів. Високі значення *Precision* і *Recall* для цих моделей також підтверджують їх ефективність у мінімізації кількості хибних позитивних і негативних спрацьовувань, що є важливим у задачах, де необхідна

максимальна точність виявлення літальних апаратів.

Моделі RetinaNet і SSD продемонстрували стабільно високі показники, проте поступаються лідерам дослідження за точністю визначення контурів об'єктів, особливо SSD. Це може свідчити про обмежену здатність SSD до роботи з об'єктами складної форми або наявністю перешкод на зображенні. Незважаючи на це, збереження високих показників *Precision* і *Recall* для RetinaNet та SSD вказує на їх потенційну застосовність у випадках, коли точність локалізації менш важлива, а основний акцент ставиться на швидкість і надійність виявлення.

SSD Lite, хоч і продемонструвала найнижчі показники за метрикою mAP , але є найшвидшою серед усіх розглянутих вище моделей, що робить її оптимальною для застосувань з обмеженими ресурсами або в умовах реального часу. Її задовільні значення *Precision* дозволяють припустити, що вона може ефективно використовуватися у випадках, коли важлива швидка класифікація об'єктів, але точність локалізації не є суттєвою. Це відкриває можливості для використання SSD Lite у задачах моніторингу повітряного простору або інших реальних сценаріях, де швидкість є ключовим фактором, наприклад, у системах попередження або в невеликих безпілотних літальних апаратах.

Висновки. У межах дослідження було проведено огляд моделей для розпізнавання об'єктів на наборі даних HRPlanesv2, що містить анотовані зображення літаків високої роздільної здатності. Результати показали, що моделі Faster R-CNN і FCOS забезпечують найвищі показники точності за метриками mAP_{50-95} та mAP_{50} , що робить їх найбільш підходящими для застосувань, де критично важлива точність локалізації об'єктів. Для моделі Faster R-CNN значення метрики mAP_{50-95} становило 0,786, а mAP_{50} – 0,985. Для моделі FCOS – 0,789 та 0,984 відповідно.

Моделі RetinaNet та SSD продемонстрували оптимальні результати, що дозволяє їх розглядати як альтернативу для задач, де швидкість і ефективність важливі, але вимоги до точності дещо нижчі. Для моделі RetinaNet показник mAP_{50-95} становив 0,764, а mAP_{50} – 0,970, водночас для SSD – 0,617 та 0,925 відповідно.

SSD Lite є найбільш підходящою моделлю для застосувань з обмеженими обчислювальними ресурсами або в режимі реального часу, як-от моніторинг повітряного простору.

ЛІТЕРАТУРА:

1. Jian Ding, Nan Xue, Gui-Song Xia, Xiang Bai, Wen Yang, Micheal Ying Yang, Serge Belongie, Jiebo Luo, Mihai Datcu, Marcello Pelillo, Liangpei Zhang. Object Detection in Aerial Images: A Large-Scale Benchmark and Challenges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021. 44, 11, 7778–7796. 10.1109/TPAMI.2021.3117983
2. UgurAlganci, MehmetSoydas, and ElifSertel. Comparative Research on Deep Learning Approaches for Airplane Detection from Very High-Resolution Satellite Images. *Remote Sensing*. 2020, 12, № 3, 458. 10.3390/rs12030458
3. Zhang, P., Yan, P., Zhang, T., Cui, W., Zhu, Q. UAV Aerial Image Detection Based on Improved FCOS Algorithm. Digital TV and Wireless Multimedia Communication. IFTC 2020. *Communications in Computer and Information Science*. 1390, 2021, 52–62. Springer, Singapore.10.1007/978-981-16-1194-0_5
4. Yadav, A., Jain, K., Pandey, A., Ranyal, E., & Majumdar, J. (2022). An Optimal Retinanet Model For Automatic Satellite Image Based Missile Site Detection. *Defence Science Journal*. 72(5), 753–761. 10.14429/dsj.72.18215
5. DilsadUnsal. HRPlanesv2 – High Resolution Satellite Imagery for Aircraft Detection. 2022, 10.5281/ZENODO.7331974
6. Tsung-Yi Lin, Michael Maire, Serge Belongie, LubomirBourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. 2014. Microsoft COCO: Common Objects in Context. 10.48550/ARXIV.1405.0312

REFERENCES:

1. Jian, Ding, Nan, Xue, Gui-Song, Xia, Xiang Bai, Wen, Yang, Micheal, Ying Yang, Serge, Belongie, Jiebo, Luo, Mihai, Datcu, Marcello, Pelillo, & Liangpei, Zhang. (2021). Object Detection in Aerial Images: A Large-Scale Benchmark and Challenges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44, 11, 7778–7796. 10.1109/TPAMI.2021.3117983
2. Ugur, Alganci, Mehmet Soydas, & Elif Sertel. (2020). Comparative Research on Deep Learning Approaches for Airplane Detection from Very High-Resolution Satellite Images. *Remote Sensing*., 12, № 3, 458. 10.3390/rs12030458
3. Zhang, P., Yan, P., Zhang, T., Cui, W., & Zhu, Q. (2021). UAV Aerial Image Detection Based on Improved FCOS Algorithm. Digital TV and Wireless Multimedia Communication. IFTC 2020. *Communications in Computer and Information Science*, 1390, 52–62. Springer, Singapore.10.1007/978-981-16-1194-0_5
4. Yadav, A., Jain, K., Pandey, A., Ranyal, E., & Majumdar, J. (2022). An Optimal Retinanet Model For Automatic Satellite Image Based Missile Site Detection. *Defence Science Journal*, 72(5), 753–761. 10.14429/dsj.72.18215
5. Dilsad Unsal. (2022). HRPlanesv2 – High Resolution Satellite Imagery for Aircraft Detection. 10.5281/ZENODO.7331974
6. Tsung-Yi, Lin, Michael, Maire, Serge, Belongie, LubomirBourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, & Piotr Dollár. (2014). Microsoft COCO: Common Objects in Context. 10.48550/ARXIV.1405.0312