

ДОСЛІДЖЕННЯ ВПЛИВУ ЗНАЧЕННЯ ПОРОГУ NON-MAXIMAL SUPPRESSION НА ЗДАТНІСТЬ YOLO ДО РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ЗОБРАЖЕННЯХ НИЗЬКОЇ ЯКОСТІ

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

mariia.ladonia@gmail.com

Вступ

З розвитком сучасних технологій щорічно спостерігається тенденція збільшення попиту на дрони [3] та інтенсивним використанням безпілотних літальних апаратів (БПЛА) у різних галузях, включаючи геодезію, агрокультуру, у військових місцях, як пошукові, розвідувальні, бойові, також для аналізу місцевості для виявлення масштабів пошкодження після стихійних лих [5], транспорт і багато інших, виникає ріст потреби у здатності автоматично аналізувати та обробляти отримані візуальні дані. У цьому контексті, нейронні мережі, які здатні до глибокого навчання та аналізу зображень, стають основним інструментом для вдосконалення якості обробки знімків, отриманих від БПЛА.

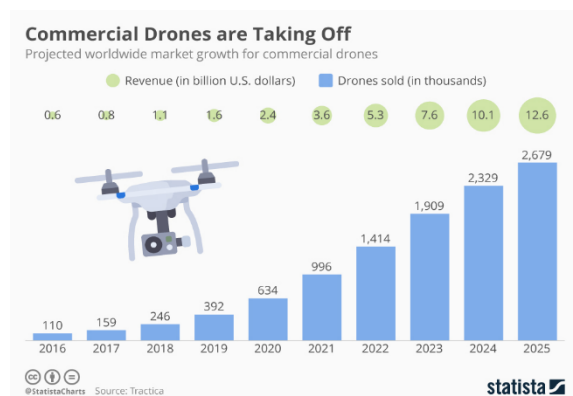


Рис. 1. Статистичні дані щорічного росту продажів дронів [3]

Вивчивши наявне дослідження і порівняння ефективності моделей виявлення об'єктів, включаючи Faster R-CNN, RetinaNet, SSD (Single Shot MultiBox Detector) та YOLO (You Only Look Once), з акцентом на фактори, що впливають на їхню ефективність, такі як швидкість і

точність. Можна виділити такі потенційні переваги YOLO на основі наданої інформації:

1. Обробка в режимі реального часу: Моделі YOLO (You Only Look Once), включаючи YOLO7, відомі своїми можливостями обробки в реальному часі. Вони призначені для застосувань, де швидкість має вирішальне значення, що робить їх придатними для таких завдань, як аналіз відео та виявлення об'єктів у реальному часі.

2. Однокадровий детектор: YOLO - це детектор з одним пострілом, що означає, що він може робити прогнози для декількох об'єктів за один прохід через нейронну мережу. Ця конструкція може бути більш ефективною з точки зору швидкості в порівнянні з двоетапними детекторами, такими як Faster R-CNN, які вимагають окремих етапів пропозиції регіону та класифікації.

3. Компроміс між швидкістю і точністю: YOLO пропонує хороший компроміс між швидкістю і точністю. Хоча він не може бути абсолютним лідером за точністю, він досягає балансу, який часто є достатнім для багатьох практичних застосувань.

4. Багатомасштабне виявлення: YOLO використовує багатомасштабний підхід, що дозволяє ефективно виявляти об'єкти різного розміру. Це важливо для застосувань, де об'єкти можуть бути як великими, так і малими.

5. Простота впровадження: архітектура YOLO є відносно простими, що полегшує його впровадження та розгортання для розробників.

6. Ефективність використання пам'яті: моделі YOLO, як правило, ефективно використовують пам'ять, що може бути важливим при розгортанні моделей на пристроях з обмеженими ресурсами.

Важливо зазначити, що вибір найкращої моделі виявлення об'єктів залежить від конкретних вимог програми. YOLO потенційно ефективною для задач у режимі реального часу, де потрібен хороший баланс між швидкістю і точністю.

YOLO був розроблений та вдосконалився протягом багатьох років, щоб забезпечити задовільний продуктивність для виявлення та класифікації кількох цілей у реальному часі. У контексті того, що перехресне злиття технологій стає новим напрямом, дослідники запропонували технологію БПЛА на основі YOLO (YBUT), об'єднавши дві вищезгадані технології. Пропонована інтеграція сприяє посиленню застосування нових технологій та розширенню ідеї розробки алгоритмів YOLO та технологій дронів. Сучасні дослідження показали, що технології БПЛА на базі YOLO (YOLO-based UAV technology - YBUT) мають високу потенціал.[2] Тому у цій статті розглядатиметься технологія YOLO для задач візуального аналізу двовимірних зображень та детекції об'єктів у контексті дослідження їх ефективності в залежності від параметра порогу Non-Maximal Suppression.

Згадане вище дослідження із порівнянням аргументує вибір YOLO для поточного дослідження. Наразі наявні оновлені версії моделі, тому проаналізовано наявні дослідження із порівняння ефективності від YOLOv3 до YOLOv7 на наборі даних COCO128, що підтверджує доцільність використання YOLOv7 для дослідження результативності на

Model	Size (Pixel)	Date	Framework	Backbone	mAP@0.50	mAP@0.50-0.95
YOLOv3	640	2018	Darknet	Darknet53	83.4	51.9
YOLOv4	640	2020	Darknet	CSPD Darknet53	81.6	56.5
YOLOv5s	640	2020	PyTorch	Modified CSP v7	83.7	56.7
YOLOv6s	640	2022	PyTorch	EfficientRep	87.7	58.4
YOLOv7	640	2022	PyTorch	RepConvN	92.0	62.8

Рис. 2. Порівняння результатів дослідження YOLOv3-YOLOv7 на наборі даних COCO128.[4]

Наявні дослідження говорять про те, що оснащення БПЛА вбудованим

процесором, із використанням алгоритма YOLO, розпізнавання та ідентифікація об'єктів може бути реалізована на основі відеозаписи с камери БПЛА в реальному часі, перетворюючи два етапи виявлення об'єктів БПЛА та комп'ютерного виявлення на одночасне виявлення самим БПЛА, що значно економить кошти, час роботи та підвищує ефективність роботи. Підвищення рівня автономного розпізнавання цілей дронами може сильно сприяти автоматизації або безпілотної експлуатації дронів у більшості галузей.[2]

У сучасних системах виявлення об'єктів, таких як YOLOv7, параметр порогу Non-Maximal Suppression (NMS) відіграє важливу роль в обробці виходів моделі та покращенні точності виявлення. Цей дослід присвячений вивченню впливу значення порогу NMS на здатність YOLOv7 розпізнавати об'єкти на зображеннях, які отримані від безпілотної літальних апаратів (дронів), зокрема, на зображеннях із низькою якістю.

Мета

Головною метою цього дослідження є вивчення та оцінка оптимального значення порогу NMS для моделі YOLOv7 на зображення отриманих за допомогою БПЛА, особливістю яких є: наявність шуму, низька якість, низька роздільна здатність зображень, різні кути об'єктів, їх малі розміри, виявити її потенціал у області розпізнавання двовимірних зображень, що зроблені за допомогою БПЛА у специфічних умовах, як при задимленні, тумані. Дослідження спрямоване на визначення того, як варіювання параметра NMS може вплинути на точність та частоту хибної детекції об'єктів.

Основна частина

Розглянемо особливості даних на яких сфокусоване дослідження та потенційні проблеми при застосуванні нейромереж для обробки знімків з низькою якістю, які необхідно враховувати:

Непередбачувана якість вхідних даних: знімки, отримані від БПЛА, можуть бути вплинуті різними факторами, такими як погодні умови, вібрації апарату,

дефекти обладнання і т. д. Ці фактори можуть призводити до непередбачуваних артефактів та спотворень на зображеннях, що ускладнює їх аналіз і обробку.

Низька роздільна здатність: деякі БПЛА мають обмежену роздільну здатність камер, що може призводити до низької якості отриманих знімків. Це може обмежити здатність нейромережі розпізнавати об'єкти або деталі на зображеннях.

Шум та спотворення: знімки з низькою якістю часто містять шуми та спотворення, які можуть ускладнювати аналіз. Нейромережі можуть бути чутливими до цих шумів і виробляти неточні результати.

Недостатня кількість навчальних даних: для навчання нейромереж потрібна велика кількість якісних навчальних даних. Однак у випадку з низькоякісними знімками може бути важко зібрати достатню кількість навчальних прикладів, що може вплинути на якість моделі.

Витрати обчислювальних ресурсів: обробка низькоякісних зображень може вимагати більше обчислювальних ресурсів і тривалого часу для навчання та інференсу нейромережі. Це може бути обмежувальним фактором для застосувань в реальному часі.

Неоднорідність даних: Знімки можуть бути надіслані з різних джерел і мати різну якість та характеристики. Розробка нейромережі, яка може ефективно обробляти таку неоднорідність, може бути викликом.

Відсутність посилань на дані: у разі низькоякісних знімків може бути важко визначити точне місцезнаходження та орієнтацію об'єктів на зображеннях, що може ускладнити завдання сегментації та розпізнавання.

Для подолання цих проблем важливо використовувати методи покращення якості зображень, попередню обробку даних, вдосконалені архітектури нейромереж та враховувати специфіку задачі. Також, ретельне планування та експерименти допоможуть визначити оптимальний підхід до вирішення конкретних завдань з низькоякісними знімками.

Non-Maximal Suppression (NMS) техніка, яка використовується для підсумовування багатьох виявлень об'єктів, які можуть бути дубльовані на зображенні. Вона допомагає видалити непотрібні бокові області та зберегти лише найбільш важливі об'єкти. У межах цього дослідження було досліджено різні версії моделі YOLOv7 із різними параметрами.

Для дослідження було виконано наступний алгоритм дій:

Досліджено різне попередньо навчені моделі (YOLOv7, YOLOv7-X, YOLOv7-W6, YOLOv7-E6E) на тестовій вибірці із параметром NMS рівним 0.5, отримані дані порівнянні у таблиці (табл. 1), де

- batch 32 average time – вимірювальна метрика, яка характеризує середній час обчислювальної операції за яку виконувалась обробка моделлю пакета розміром 32.
- Test Size – розмір вхідних зображень використовувався для тестування моделей.
- APtest – вимірювальна метрика, що характеризує середню точність.

Таблиця 1

Model	Test Size	APtest	batch 32 average time
YOLOv7	640	52.3%	2.9 ms
YOLOv7-X	640	53.2%	4.5 ms
YOLOv7-D6	640	52.8%	13.4 ms
YOLOv7-E6E	640	53.5%	15.9 ms

Виконавши аналіз отриманих даних, найбільш ефективною та швидкою для даної задачі показала себе модель yolov7.pt, тому для подальшого дослідження було встановлено наступні параметричними дані:

- weights = 'yolov7.pt';
- conf = 0.5;
- img-size = 640.

Створено тестовий набір зображень низької якості у який увійшло 200 зображень: сформовано на базі набору даних [6], виявлено наявність шуму, перекриття об'єктів, знімки з дуже малими об'єктами.

Досліджено результат роботи YOLOv7 з різними значеннями NMS Threshold на одних і тих самих зображеннях: виконала тестування YOLOv7 на своєму тестовому наборі зображень починаючи від 0.1 до 0.75 із кроком 0.05.

Оцінено результати: виконано порівняння результатів виявлення об'єктів для різних значень порогу NMS. Проведлося спостереження за кількістю об'єктів, які виявляються, та за точністю цих виявлень, а також хибне виявлення.

Вибрано оптимальне значення: знаючи, як значення порогу NMS впливає на виявлення об'єктів на зображеннях низької якості, визначасмо оптимальне значення для задачі розпізнавання об'єктів на зображеннях низької якості.

Отримані результати наведено у таблиці залежності кількості розпізнаних об'єктів, пропущених та хибно ідентифікованих від параметру (табл. 2) NMS, а також продемонстровано на рис. 3-8, для виконання порівняння результатів детекції.

На рис. 3-5 продемонстровано як зменшення параметру збільшує кількість об'єктів, що ідентифіковані моделлю YOLOv7, а також рис. 3 демонструє, що замалий параметр може викликати хибну детекцію.

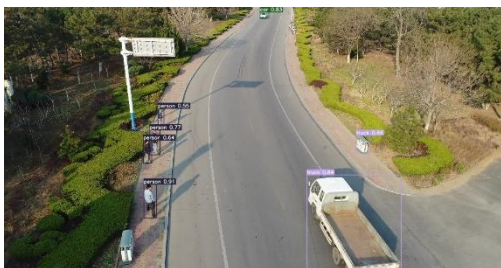


Рис. 3. Результат детекції YOLOv7, $\text{conf_thres}=0.25$

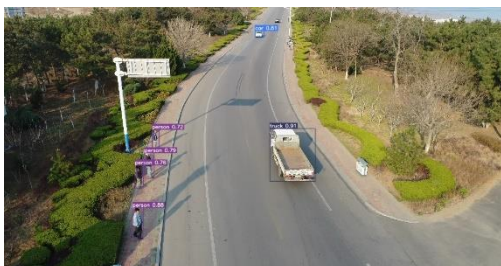


Рис. 4. Результат детекції YOLOv7, $\text{conf_thres}=0.5$



Рис. 5. Результат детекції YOLOv7, $\text{conf_thres}=0.75$

На однакових об'єктах (машини) найкраще спостерігається низька здатність моделі YOLOv7 до детекції об'єктів перекритих іншими, наприклад гілками дерев, зображених під специфічним кутом, занадто малих. На рис. 6-8 наведено результати детекції одного і того ж зображення моделлю, зі зміною параметра, при значення NMS рівному 0.75 модель здатна знайти менше ніж 50% об'єктів, знайденої при значенні NMS рівному 0.25.



Рис. 6. Результат детекції YOLOv7, $\text{conf_thres}=0.25$



Рис. 7. Результат детекції YOLOv7, $\text{conf_thres}=0.5$



Рис. 8. Результат детекції YOLOv7, $\text{conf_thres}=0.75$

Таблиця 2

Номер зображення	conf_thres	ідентифіковані об'єкти	не ідентифіковані	хибно ідентифіковані	відносні показники	
					упущених	хибних
1	0.25	8	0	1	0	0,125
1	0.5	6	1	0	0,142	0
1	0.75	5	2	0	0,286	0
2	0.25	10	1	0	0,09	0
2	0.5	5	6	0	0,545	0
2	0.75	4	7	0	0,636	0
3	0.25	9	0	2	0	0,223
3	0.5	7	0	0	0	0
3	0.75	6	1	0	0,142	0
4	0.25	25	3	0	0,107	0
4	0.5	19	9	0	0,321	0
4	0.75	8	20	0	0,714	0
5	0.25	6	0	1	0	0,167
5	0.5	4	1	0	0,2	0
5	0.75	3	2	0	0,4	0
6	0.25	1	7	0	0,875	0
6	0.5	6	4	1	0,4	0,1
6	0.75	8	2	1	0,2	0,1

Висновки

Застосування моделі YOLOv7 із варіюванням параметра показало можливе досягнення її ефективності при застосуванні для задачі візуального аналізу та розпізнавання об'єктів на зображеннях, що зроблені БПЛА і характеризуються низькою роздільною здатністю, наявністю шуму, як перекриття об'єктів іншими, наприклад, машини, що закриті гілками дерев. Проте, спостерігається наявність обмежень ефективності, при зменшенні параметра, збільшується ймовірність розпізнавання об'єктів, але також збільшується ймовірність хибного виділення об'єктів.

Отриманні дані свідчать, що регуляція значення параметру Non-Maximal Suppression може покращувати результат ідентифікації зображення за допомогою моделі YOLOv7, але для зображень із низькою роздільною здатністю, низькою якістю, наявністю шуму чи пошкодження не може дати очікуваного результату, адже при одному значенні параметра на одних зображеннях показник упущених може

сягати 87%, і одночасно 22% хибних ідентифікацій на інших.

Отже, можна розглядати задачу ідентифікування зображень з БПЛА з низькою роздільною здатністю, як задачу із малим набором даних і у подальшому досліджувати: попереднє покращення якості зображень, пошук індивідуальних підходів та розробок для роботи із такими даними.

Література

- Jonathan Hui. Object detection: speed and accuracy comparison (Faster R-CNN, R-FCN, SSD, FPN, RetinaNet and YOLOv3). URL: <https://jonathan-hui.medium.com/object-detection-speed-and-accuracy-comparison-faster-r-cnn-r-fcn-ssd-and-yolo-5425656ae359>.
- Chen C., Zheng Z., Xu T., Guo S., Feng S., Yao W., Lan Y. YOLO-Based UAV Technology: A Review of the Research and Its Applications. *Drones*. 2023. V. 7. 190.
- Buchholz K., Statista. Commercial Drones Are Taking Off. URL: <https://www.statista.com/chart/17201/commercial-drones-projected-growth/>.

4. Yang H., Liu Y., Wang S., Qu H., Li N., Wu J., Yan Y., Zhang H., Wang J., Qiu J. Improved Apple Fruit Target Recognition Method Based on YOLOv7 Model. *Agriculture*. 2023. V. 13. 1278.

5. Weber E., Vedaldi A., Bischof H., Brox T., Frahm J.M. Detecting Natural Disasters, Damage, and Incidents in the Wild. *Computer Vision – ECCV 2020. ECCV 2020*.

Lecture Notes in Computer Science. 2020. V. 12364. P. 331–350.

6. Du D., Qi Y., Yu H.g, Yang Y., Duan K., Li G., Zhang W.g, Huang Q., Tian Q. The Unmanned Aerial Vehicle Benchmark: Object Detection and Tracking. *European Conference on Computer Vision (ECCV)*. 2018. 17 p.

Ладоня М.С.

ДОСЛІДЖЕННЯ ВПЛИВУ ЗНАЧЕННЯ ПОРОГУ NON-MAXIMAL SUPPRESSION НА ЗДАТНІСТЬ YOLO ДО РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ЗОБРАЖЕННЯХ НИЗЬКОЇ ЯКОСТІ

На основі аналізу існуючих підходів щодо виявлення об'єктів на зображенні, було обрано модель YOLO для дослідження залежності ефективності від значення порогу Non-Maximal Suppression. У статті розглядається актуальна проблема використання безпілотних літальних апаратів (БПЛА), зокрема дронів, у різних галузях, де вони використовуються для збору візуальних даних. Основним завданням дослідження є вивчення та оцінка оптимального значення порогу Non-Maximal Suppression (NMS) для моделі YOLOv7 при обробці зображень, отриманих від БПЛА в умовах, характеризованих низькою роздільною здатністю, наявністю шуму та іншими артефактами. Дослідження показує, що модель YOLOv7 може бути ефективною для розпізнавання об'єктів на зображеннях, отриманих з дронів, навіть при наявності низької роздільної здатності та шуму. Проте, виявлене, що зміна параметра NMS впливає на точність та частоту помилкових детекцій об'єктів. Зниження значення параметра може підвищити впізнаваність об'єктів, але при цьому збільшується ймовірність хибних виявлень. Отримані результати вказують на необхідність подальших досліджень у цій області, зокрема, поліпшення якості вихідних зображень, розробка індивідуальних підходів та методів для роботи з візуальними даними, отриманими від БПЛА з низькою роздільною здатністю.

Ключові слова: YOLO, YOLOv7, UAV, NMS, ідентифікація об'єктів.

Ladonia M.S.

RESEARCH ON THE IMPACT OF NON_MAXIMAL SUPPRESSION THRESHOLD VALUE ON YOLO'S ABILITY TO RECOGNIZE OBJECTS IN LOW-QUALITY IMAGES

Based on the analysis of existing approaches to object detection in images, the YOLO model was selected for investigating the dependency of performance on the Non-Maximal Suppression (NMS) threshold value. This article addresses the current challenge of utilizing unmanned aerial vehicles (UAVs), particularly drones, in various fields where they are employed for collecting visual data. The primary objective of this research is to study and evaluate the optimal NMS threshold value for the YOLOv7 model when processing images acquired from UAVs under conditions characterized by low resolution, noise, and other artifacts. The study demonstrates that the YOLOv7 model can be effective in recognizing objects in images obtained from drones, even in the presence of low resolution and noise. However, it was observed that altering the NMS parameter affects the accuracy and frequency of false object detections. Decreasing the parameter value can enhance object recognition but concurrently increases the likelihood of false detections. The obtained results indicate the need for further research in this area, including improving the quality of source images, developing individualized approaches, and methods for working with visual data obtained from UAVs with low resolution.

Keywords: YOLO, YOLOv7, UAV, NMS, objects detection.