

МЕЛЬНИЧЕНКО Олександр

Хмельницький національний університет

<https://orcid.org/0000-0001-8565-7092>e-mail: [oleksandr.melnychenko@live.com](mailto:oleksandr.melnychenko@live.com)

## МЕТОДИ ЗБОРУ ТА ОБРОБКИ ЗОБРАЖЕНЬ ОТРИМАНИХ З ВИКОРИСТАННЯМ БПЛА ДЛЯ ВИЯВЛЕННЯ ЗАДАНИХ ОБ'ЄКТІВ

У поданій роботі запропоновано новий підхід до виявлення та обчислення кількості заданих структурних об'єктів на прикладі фруктових плодів за зображеннями, що генеруються безпілотними літальними апаратами (БПЛА) в режимі реального часу. Запропонований підхід полягає у послідовному виконанні запропонованих методів, а саме, метод динамічного отримання зображень заданих структурних об'єктів в тривимірному просторі за допомогою групи БПЛА, метод синхронізації відеопотоків з різних БПЛА, метод виявлення заданих структурних об'єктів та спосіб отримання кількості структурних об'єктів. У результаті проведених експериментальних досліджень над розробленими методами було продемонстровано практичну значущість їхнього послідовного застосування до робочого середовища фруктового саду. Так, оцінка синхронізації відеопотоків за індексом SSIM коливається від 0,79 до 0,92, із середнім значенням 0,87, а за індексом PSNR – від 22 до 39, що свідчить про високу ефективність розробленого методу синхронізації та хорошу якість отриманих об'єднаних зображень. Середнє значення показника достовірності виявлення та обчислення кількості фруктових плодів становить 86,75%. Отримані значення свідчать про високий показник ефективності методів збору та обробки зображень отриманих з використанням БПЛА для виявлення заданих об'єктів.

Ключові слова: врожайність яблук, збір зображень, обробка зображень, синхронізація відеопотоків, виявлення об'єктів, YOLOv5, глибоке навчання.

MELNYCHENKO Oleksandr  
Khmelnitskyi National University

## METHODS OF COLLECTING AND PROCESSING IMAGES OBTAINED USING UAVS TO DETECT SPECIFIED OBJECTS

This study introduces a novel approach for identifying and quantifying specific structural objects, using fruit as an example, through images captured by drones in real time. The approach involves a series of steps: dynamically capturing images of the specified objects in 3D space using a fleet of drones, synchronizing video feeds from different drones, and finally, detecting and counting the objects. During the detection phase, a new version of the YOLOv5 model, YOLOv5-v1, was trained on a custom dataset of apple images to learn the features that differentiate apples from their surroundings. The initial size of the source network's binding block was adjusted to avoid misidentifying small objects in the image background, thereby improving counting accuracy. The authors conducted computational experiments on a dataset of apple images they created to assess the approach's effectiveness. The synchronization of video streams was assessed using the SSIM index, which ranged from 0.79 to 0.92, with an average value of 0.87, and the PSNR index, which ranged from 22 to 39. These values indicate the high efficiency of the developed synchronization method and the high quality of the resulting combined images. The results on the test dataset showed that the improved model could effectively identify fruits captured by the drone camera, with recall, precision, mAP, and F1-score of 92.13%, 84.59%, 87.94%, and 89.02%, respectively. The proposed approach was also compared with several other cutting-edge models, YOLOv3, YOLOv4, and YOLOv5, and was found to be superior in accuracy and speed. Our model's average recognition speed was 0.015 seconds per video frame (66.7 frames/s), which was 1.13 and 3.53 times faster than the YOLOv4, and YOLOv3 networks, respectively. The average reliability index for detecting and counting fruit was 86.75%. These results demonstrate the effectiveness of using drone-captured images to detect specified objects. Future research could also explore including additional features, such as fruit shape, by training the model on a larger dataset.

Keywords: apple yield, image collection, image processing, video stream synchronization, object detection, YOLOv5, deep learning.

### Постановка проблеми у загальному вигляді

#### та її зв'язок із важливими науковими чи практичними завданнями

Яблука є однією з найбільших фруктових культур у світі, а Україна має найбільшу площу плантацій, що займає близько двох мільйонів гектарів і дає найвищий урожай у Європі. У центральних та південних регіонах України є чудові кліматичні та ґрунтові умови, сприятливі для вирощування яблук [1]. Крім того, технологічні вдосконалення сільськогосподарської практики в Україні призвели до збільшення врожайності яблук і покращення якості.

Виявлення врожайності яблук за допомогою візуального розпізнавання є ефективним та інтуїтивно зрозумілим методом. Однак, оскільки кожне плодове дерево має різний ріст і кількість плодів, для досягнення більшої точності потрібне індивідуальне визначення врожайності. У садах із великою кількістю дерев ефективність має вирішальне значення. Отже, необхідно мати модель виявлення яблук, яка є швидкою, точною та малою, що дає змогу швидше виявляти врожайність яблук під час роботи на різних вбудованих пристроях.

### Аналіз досліджень та публікацій

Раніше для виявлення фруктів зазвичай використовували звичайні методи зору, такі як обробка зображень і машинне навчання. Ці методи розпізнавали фрукти за кольором, формою та текстурою. Наприклад, у роботі [2] досягли рівня точності 89,92 % для зеленого лічі та 94,50 % для червоного лічі,

навчивши модель розпізнавати плід лічі за ознаками кольору та текстури. Моделі глибокого навчання (ГН), зокрема згорткові нейронні мережі (CNN), мають можливість витягувати характеристики із зображень, що може призвести до автоматичного розпізнавання цілей і кращої адаптивності. Наприклад, у роботі [3] запропонували архітектуру Faster R-CNN для виявлення багатокатегорійних фруктів, яка досягла середньої точності 90,72 % і часу виявлення 58 мілісекунд (мс) на зображення.

В іншій роботі, як от, You Only Look Once (YOLO) [4], дослідники досягли більш збалансовану продуктивність щодо швидкості точності виявлення. Це призводить до загальної кращої продуктивності моделі. Наприклад, у дослідженні [5] запропонували покращений метод виявлення YOLOv3 для недозрілих яблук у саду та отримали точність 61,6, а в роботі [6] досягли точності розпізнавання яблук 97,13 % у складних садах за допомогою YOLOv4 із середнім часом розпізнавання 16,69 мс на одне зображення на одному графічному процесорі. Нещодавнє дослідження на основі YOLOv5 також продемонструвало відмінні результати в ідентифікації фруктових плодів [7].

Типові фруктові сади містять карликові та веретені подібні дерева, які мають менш щільне листя, тобто менше яблук приховано від огляду. З усім тим, деякі яблука можуть бути закриті гілками або листям, а освітлення на задньому плані може бути нерівномірним або складним. Як наслідок, у фруктовому саду виникає недостатнє освітлення для захоплення зображення на стороні фонового освітлення фруктових дерев, що ускладнює виявлення яблук на зображенні підсвічування, знятому камерою.

Автоматичне виявлення фруктів на деревах відбувається в складному сітчастому середовищі, де фрукти ростуть на деревах, розташованих послідовними рядами. Під час робочої місії фрукти можуть бути зняті відеоканерою БПЛА та подані на передньому плані робочого сегмента (фруктове дерево), на задньому плані або на деревах інших рядів. Як наслідок, фрукти у відеокадри можуть бути занадто малими, через що автоматизована система може їх помилково розпізнати, або вони можуть потрапити в різні відеокадри під час робочої місії, що призведе до кількох підрахунків однакових фруктів. Крім того, мінливі погодні умови, зміна освітлення впродовж дня, падіння листя та гілок у кадр можуть погіршити точність виявлення плодів. Тому, для усунення згаданих вище проблем у цій роботі пропонується новий підхід до виявлення фруктових плодів у вигляді заданих структурних об'єктів. Мета роботи полягає в автоматизації процесу виявлення та обчислення кількості заданих структурних об'єктів за зображеннями, що згенеровані з допомогою БПЛА.

#### **Метод динамічного отримання зображень заданих структурних об'єктів**

Забезпечення мети роботи можливе із впровадженням нового методу динамічного отримання зображень заданих структурних об'єктів використовуючи групу БПЛА. Вхідними даними методу динамічного отримання зображень заданих структурних об'єктів є множина даних: а) координати початкової точки робочого середовища; б) координати кінцевої точки робочого середовища; в) матриця координат початкових точок робочих сегментів; г) матриця координат кінцевих точок робочих сегментів; д) матриця координат початкової траєкторії руху БПЛА.

Метод динамічного отримання зображень структурних об'єктів однієї природи в тривимірному просторі полягає в покроковому виконанні наступних блоків.

**Блок 1.** Зліт групи БПЛА до першої початкової точки. Визначення стану для програмного модуля розташування кожного БПЛА включає перевірку координат та відповідність фізичного розташування всієї групи щодо вказаних програмних координат на початку. Занесення інформації в центральний модуль та всіх підсистем і визначення станів усіх системних модулів.

**Блок 2.** Формування відеоданих, отриманих від групи БПЛА в цільовій зоні. Прийняття рішень про подальшу роботу групи БПЛА, підтвердження цілісності програмної місії та визначення рівня критичності для продовження виконання роботи групи БПЛА, відбувається через визначення станів програмних модулів після виконання блоку 2. Вихідними даними методу є згенерований відеоряд із цільовими об'єктами фруктового саду.

Отримані дані цільового блоку 2 визначають стан та ефективність роботи групи БПЛА в робочій зоні. Якщо програмні модулі набули всі необхідні ознаки для продовження місії, підсистема управління групою БПЛА відправляє сигнал на переміщення групи до наступного цільового фруктового дерева.

Під час перельоту групи БПЛА між цільовими робочими зонами фруктових дерев відеоканери БПЛА тимчасово деактивуються для отримання неспотворених цільових відеоданих. У результаті, система переходить у стан «завершення» виконання програмної місії групою БПЛА в робочій зоні в разі фізичного переміщення групи в кінцеву точку даного робочого сегменту. Отже, робота в зоні вважається завершеною.

Метод динамічного отримання зображень заданих структурних об'єктів у тривимірному просторі за допомогою декількох БПЛА забезпечує в процесі активації кожного БПЛА генерацію відеоряду в центральну систему та прийняття рішень про подальшу роботу групи БПЛА, що покращує узгодження між різними БПЛА та досягнення переміщення всієї групи БПЛА від заданих початкових до кінцевих точок програмної місії автономно.

#### **Метод синхронізації відеопотоків з різних БПЛА**

Процес синхронізації відеопотоків із різних БПЛА може бути ускладнений багатьма чинниками, наприклад, БПЛА можуть мати різні типи відеоканер, швидкість отримання відеорядів може різнитись, а самі відеоряди можуть спотворюватись або не отримуватись взагалі. Тому, щоб не допустити подібні проблеми з отриманням відеорядів, запропоновано новий метод синхронізації відеопотоків у режимі реального часу. Метод складається з трьох основних блоків.

**Блок 1.** Корегування швидкості передачі відеорядів. У процесі виконання програмної місії в робочих зонах група БПЛА утворює відеоряди, які транслюються через мережу в програмний модуль виявлення та обчислення кількості заданих структурних об'єктів. Результатом виконання блоку корегування швидкості є перехід блоку в стан "виконання" та передача утвореної множини програмних структур на наступний блок виконання.

**Блок 2.** Синхронізації відеопотоків у часі. Загальне функціонування блоку зумовлене кількістю множин програмних структур, отриманих від попереднього блоку щодо корегування швидкості передачі відеорядів. Першочергово блок синхронізації відеопотоків у часі здійснює перевірку кількості множини програмних структур, що призводить до оброблення наступних сценаріїв. Результатом виконання даного блоку є множина програмних структур, що слугують входними даними для наступного блока злиття відеокадрів.

**Блок 3.** Злиття відеокадрів. Спершу блок злиття відеокадрів перевіряє еквівалентність часу утворення множини програмних структур даних. Виконання перевірки часових характеристичних забезпечує цілісність сформованих даних під час утворення однієї програмної структури даних. До отриманих із БПЛА відеокадрів застосовується алгоритм алгебраїчного перетворення зображень, що дає змогу утворити одне зображення через злиття множини відкоригованих зображень. Як наслідок, такого фільтрування, злиття відеокадрів в одне ціле зображення відбувається без помітних переходів та із чіткими краями. Відфільтровані відеокадри надсилаються в програмний механізм стиснення. Процедура стиснення відеокадру відбувається на основі афінного перетворення. Результатом виконання такої процедури є підготовлена до злиття множина відеокадрів.

Далі множина підготовлених відеокадрів надсилається до програмної функції злиття. Злиття декількох відеокадрів в одне цілісне зображення полягає в утворенні об'єктної програмної структури даних матричного вигляду. Кожен елемент матриці відповідає значенню коду кольору одного графічного пікселя. Утворена в такий спосіб матриця даних відображає суцільне плодове дерево, де всі відокремлені відеокадри, отримані від різних БПЛА, зливаються в одне ціле. Отримані дані заносяться до внутрішньої бази злитих відеокадрів. Насамкінець, матрична програмна структура даних надсилається до програмного модуля виявлення та обчислення кількості структурних об'єктів однієї природи.

Результатом успішного виконання блоку злиття відеокадрів є отримання на виході програмної структури даних у вигляді матриці кодів кольору. Розроблений елемент функціональності блока злиття щодо занесення даних у внутрішню базу злитих відеокадрів забезпечує цілісність системи у випадку її критичного збою.

### Метод виявлення заданих структурних об'єктів

Метою розроблення та впровадження методу виявлення заданих структурних об'єктів є підвищення точності розпізнавання та обчислення кількості заданих структурних об'єктів, зокрема плодів на деревах фруктових садів. Метод виявлення заданих структурних об'єктів полягає у формальному відображенні плодів на фруктових деревах.

Вхідними даними запропонованого в роботі підходу є цифрове зображення структурних об'єктів, представлене у вигляді матриці чисел, кожне з яких відображає інтенсивність відповідного пікселя на зображенні. Модель цифрового зображення структурних об'єктів представимо у вигляді лінійного представлення:

$$f: I \rightarrow \langle C_{xy}, P_{h \times w}^i, B_{u \times v}^i, B_{h \times w}^i \rangle, \quad (1)$$

де  $C_{xy} = (c_x, c_y)$  – ширина  $c_x$  і висота  $c_y$  координатної сітки  $C_{xy}$  координатної площини  $XY$ , сформованої для виявлення структурних об'єктів;  $P_{h \times w}^i = (p_h^i, p_w^i)$  – висота  $p_h^i$  та ширина  $p_w^i$  рамки прив'язки  $i$ -го цільового структурного об'єкта в координатній площині  $XY$ ;  $B_{u \times v}^i = (b_u^i, b_v^i)$  – координати центру рамки прив'язки  $i$ -го цільового структурного об'єкта в координатній площині  $XY$ ;  $B_{h \times w}^i = (b_h^i, b_w^i)$  – висота  $b_h^i$  та ширина  $b_w^i$  обмежувальної рамки, що окреслює  $i$ -й цільовий структурний об'єкт у координатній площині  $XY$ .

Для досягнення високої точності виявлення структурних об'єктів у режимі реального часу фізичний розмір оригінальної нейромережевої моделі YOLOv5 має бути мінімальним, а структура нейронної мережі – відповідно максимально стиснутою. Тому, магістральний блок нейронної мережі, який вилучає ознаки цільового об'єкта на вхідних зображеннях, було модифіковано через зменшення кількості вагових параметрів нейронної мережі. Вдосконалення магістрального блоку полягає в наступних змінах.

**Блок 1.** Модуль фокусування нейронної мережі модифіковано для зменшення часу навчання нейронної мережі, а саме: 1) вхідне зображення, що подане трьома каналами RGB та розміром  $3 \times 640 \times 640$ , було розбито на 4 частини розмірами  $3 \times 320 \times 320$  кожна; 2) 4 частини зображення об'єднано в одну карту ознак за RGB каналом із результируючим розміром  $12 \times 320 \times 320$ ; 3) до отриманої карти ознак розміром  $12 \times 320 \times 320$  застосовано операцію згортки із 32 ядрами, яка генерує карту ознак розміром  $32 \times 320 \times 320$ ; 4) до отриманої карти ознак розміром  $32 \times 320 \times 320$  застосовано операцію пакетної нормалізації та активаційну

функцію Hardwish; вихідна карта ознак розміром  $32 \times 320 \times 320$  передається на наступний етап оброблення нейронної мережі.

**Блок 2.** Модуль BottleneckCSPv1. Вдосконалення модуля BottleneckCSPv1 в архітектурі YOLOv5-v1 полягає у видаленні згорткового шару, що в оригінальному модулі BottleneckCSP поєднував вхідну карту ознак з операцією конкатенації.

**Блок 3.** Об'єднання карт ознак. Щоби підвищити точність виявлення структурних об'єктів на відеокадри, до архітектури YOLOv5-v1 застосовано операцію об'єднання карт ознак різних розмірів в одну карту ознак із вищою дискримінаційною здатністю. З огляду на спрощену структуру модифікованого модуля BottleneckCSPv1 та об'єднання карт ознак, що виводяться зі згорткових шарів вдосконаленої архітектури YOLOv5-v1, об'єднання шарів 4 і 15, 6 і 11, 10 і 21 оригінальної архітектури YOLOv5 було замінено на об'єднання шарів 5 і 18, 8 і 14, 13 і 24 вдосконаленої архітектури YOLOv5-v1.

У результаті виконаних модифікацій отримано вдосконалений метод, який полягає в автоматизованому перетворенні вхідної інформації, що подана як цифрове зображення структурних об'єктів однієї природи (плодів фруктових дерев), у результуючу інформацію у вигляді множини координат, що відповідають розташуванню цих об'єктів на зображенні.

#### Спосіб отримання кількості заданих об'єктів

Для успішного обчислення кількості плодів за відеопотоком необхідним є встановлення унікальності кожного плоду на двох послідовних відеокадрах, отриманих від БПЛА. Визначення унікальності ґрунтується на характерних відмінних ознаках кожного об'єкта, їхньому положенні на послідовних кадрах через зміну траєкторії кожного БПЛА в групі та непередбачувані чинники довкілля. Тому особливістю способу отримання кількості структурних об'єктів є встановлення залежностей між візуальними ознаками виявлених структурних об'єктів, що відображаються на послідовних відеокадрах. Спосіб отримання кількості заданих об'єктів полягає в послідовному виконанні наступних блоків.

**Блок 1.** Попередню обробку вхідних даних виконано, відповідно до впроваджених та модернізованих алгоритмів у методі виявлення, де необхідним є сегментація малих, середніх та великих структурних об'єктів. Малі об'єкти вважаються нецільовими для відстеження, тому вони не підлягають обчисленню. Поведінковою сигнатурою поточного блоку є фільтрування даних через вилучення обмежувальних рамок виявлених малих структурних об'єктів. Утворена нова структура даних передається на наступний блок.

**Блок 2.** Першочергове порівняння споріднених координат виявлених об'єктів між двома сусідніми кадрами. На основі отриманої чисельної інформації з блоку 1 формується стан виявлених структурних об'єктів на поточному відео кадрі, кожен із яких представляється у вигляді:

$$(x, y, \gamma, h, v_x, v_y, v_\gamma, v_h), \quad (2)$$

де  $(x, y)$  – координати центра виявленого об'єкта,  $\gamma$  – відношення ширини до висоти обмежувальної рамки виявленого об'єкта,  $v_x, v_y, v_\gamma, v_h$  – швидкість зміни відстежувальної рамки між послідовними відеокадрами.

Результатом поточного блоку є матриця станів відстежених об'єктів, що передається на наступний блок.

**Блок 3.** Присвоєння унікальних ідентифікаторів нововиявлених об'єктів. У цьому блоці відбувається додавання нового характерного стану унікальності кожного виявленого об'єкта до стану розташування виявлених об'єктів, отриманого з блоку 2.

**Блок 4.** Обчислення кількості унікальних структурних об'єктів. Блок отримує на вхід реєстр структурних об'єктів, яка містить тимчасові дані про відстежувальний об'єкт: його тип, рух, унікальний ідентифікатор, стан активності та час запису до реєстру. Допоки об'єкт перебуває в системі відстеження, він вважається "активним". Якщо об'єкт зникає з наступного відеокадру, система формує подію, за якої його стан змінюється на "неактивний" у системному реєстрі.

**Блок 5.** Обчислення прогнозних значень координат розташування об'єктів у наступного кадру. У поточному блоці формується множина прогнозних значень майбутнього розташування структурних об'єктів у наступних послідовних відеокадрах. Якщо об'єкт є присутнім на 100 наступних послідовних відеокадрах, у цьому часовому проміжку між початковим та кінцевим відеокадром об'єкт вважається ідентичним та унікальним. Прогнозні значення обраховуються за лінійною модифікацією фільтра Калмана з дискретним часом.

#### Підготовка робочого середовища та навчальних даних

Для підготовки та формування навчальних даних було використано робоче середовище сучасного фруктового саду, яке знаходиться в Хмельницькій області, Городецького району за координатами 49.265011, 26.548304. Відстань між рядами яблунь робочого середовища становить 4 м, відстань між плодовими деревами – майже 1 м, а висота дерев – майже 2 м. Відповідно до оціночної інформації Фермерського господарства "САМ+", кількість плодів, що росло на деревах фруктового саду на момент проведення експериментів, становила ~160 тисяч штук. Оригінальні зображення яблуневих плодів було отримано в

сонячні та похмурі дні. Для отримання зображень було використано камеру смартфона iPhone 13 Pro Max за різних погодних умов та на різних відстанях щодо дерев (0,5-1,5 м).

Перших 200 зображень (100 сонячних днів і 100 похмурих днів) були випадковим чином відібрані з 1 214 зображень, як тестовий набір даних, а решта 1 014 зображень – як навчальний набір. Також були додані зображення яблукових плодів інших кольорів та форм. Фінальний набір даних мав 16 224 зображень плодів, які використовуються як дані остаточного навчального набору для навчання моделі розпізнавання, включно з 15 210 покращеними зображеннями і 1014 необробленими зображеннями.

**Створення детектора.** За метрики оцінювання побудованого детектора взято: 1) precision; 2) recall; 3) середня точність класифікації (mAP); 4) F1-міра. Результати навчання нейронної мережі за вдосконаленою архітектурою YOLOv5-v1 подано на рис. 1.

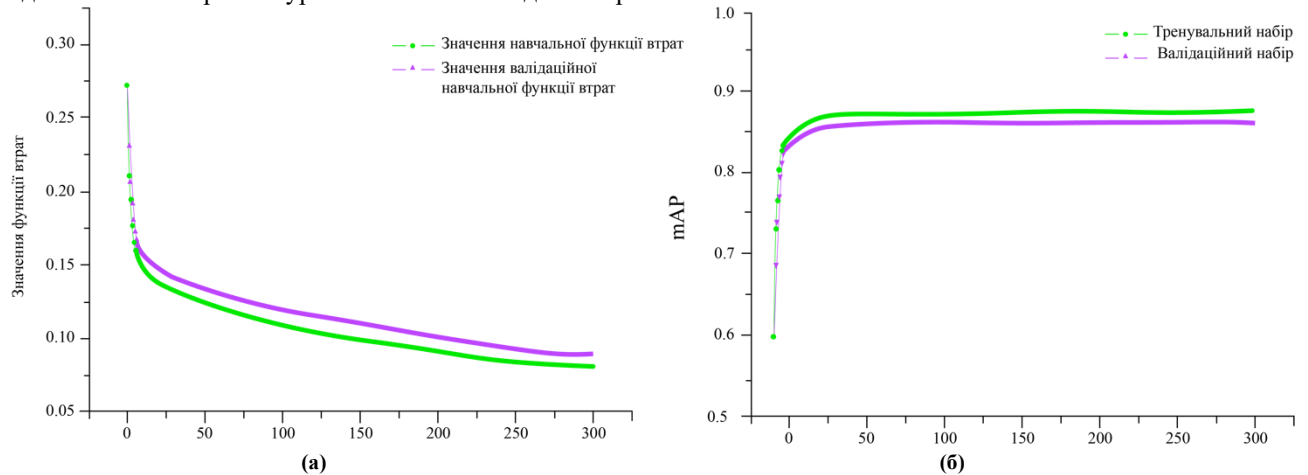


Рис. 1. Результати навчання мережі: (а) втрати під час навчання та перевірки; (б) середня точність класифікації навчальних наборів

Значення функцій втрат за валідаційним набором даних і значення функцій втрат за навчальним набором даних показані кривими на рис. 1а), вказують на те, що значення функції втрат швидко зменшуються на перших 100 ітераціях навчання нейронної мережі, водночас більш стабільна тенденція отримується після 250 ітерацій навчання. Таким чином, вихідна модель нейронної мережі утворюється після 300 ітерацій навчання. Показані вище значення діаграми на рис. 1б) показують, що модель нейронної мережі добре навчена без перенавчання.

**Результати експериментальних досліджень**

За оцінку ефективності розробленого методу синхронізації відеопотоків у режимі реального часу взято індекс SSIM та індекс PSNR. Досягнення ефективності синхронізації відеопотоків за індексом SSIM полягає в отриманні значення від 0,5 до 1, що вважається високим значення ефективності. Значення індексу PSNR обраховується як співвідношення між максимально можливою потужністю візуального сигналу та шуму, який присутній у зображенні; чим вище значення, тим краща якість отриманого зображення.

У таблиці 1 наведено результати виконання модуля синхронізації відеопотоків автоматизованої системи, що реалізує відповідний метод, для потоку із 12 послідовних груп відеокадрів, який випадковим чином відібраний для тестування; кожна група містить три відеокадри, що отримані від трьох БПЛА, які надалі об'єднуються в одне зображення.

Таблиця 1

**Результати дослідження ефективності методу синхронізації відеопотоків, отриманих групою БПЛА**

Об'єднане зображення	SSIM	PSNR	Об'єднане зображення	SSIM	PSNR
1	0,90	35,20	6	0,91	39,10
2	0,45	27,22	7	0,86	31,43
3	0,72	37,50	8	0,47	27,87
4	0,85	29,11	9	0,90	30,12
5	0,87	36,90	10	0,83	31,54

Відповідно до таблиці 1, ефективність за індексом SSIM коливається від 0,79 до 0,92, із середнім значенням 0,87. Водночас зображення, для яких значення індексу SSIM менше, ніж 0,50, вважаються системою спотвореними. Заразом ті об'єднані зображення, для яких значення індексу PSNR перевищує 30, вважаються високої якості; водночас значення PSNR менше, ніж 30 свідчать про низьку якість зображення, що може бути зумовлено зовнішніми чинниками робочого середовища (сильні пориви вітру, опади тощо).

Далі було проведено експерименти із встановленням кількості фруктових плодів. Всього є 2336 зображень з фруктами, серед яких цільова кількість, які можна захопити, становить 1007, а цільова кількість фруктів, які неможливо захопити, – 1329. Результати виявлення яблук за нейромережевою моделлю YOLOv5-v1 подано в таблиці 2.

Таблиця 2

**Результати виявлення яблук за нейромережевою моделлю YOLOv5-v1**

Дані	Кількість об'єктів	Precision (%)	Recall (%)	mAP (%)	F1 (%)
К-сть, яку можна захопити	1007	85,51	94,33	89,23	89,70
К-сть, яку не можна захопити	1329	82,56	89,32	84,87	85,81
Загалом	2336	83,83	91,48	86,75	87,49

Специфічні результати розпізнавання методу, запропонованого в дослідженні, показано в таблиці 2, яка вказує на те, що для фруктів, які можна захопити, точність, запам'ятовування, значення mAP і оцінка F1 запропонованої моделі становили 85,51%, 94,33%, 89,23% і 89,70%, відповідно; для незрозумілих плодів, результати розпізнавання становили 82,56%, 89,32%, 84,87% і 85,81% відповідно. Загальне значення влучності виявлення структурних об'єктів, повноти, mAP і F1 становили 83,83%, 91,48%, 86,75% і 87,49% відповідно.

Приклади результатів розпізнавання запропонованої моделі нейронної мережі фруктів, які можна захопити, і фруктів, які не можна захопити, за різних погодних умов і умов освітлення показані на рис. 2.

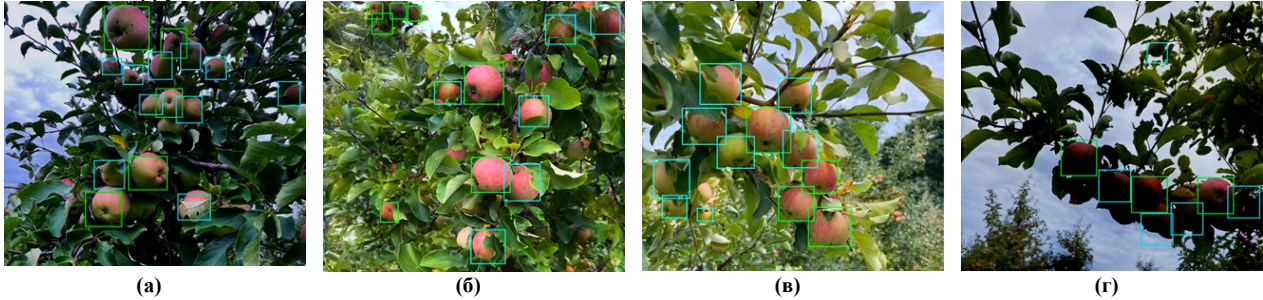


Рис. 2. Результати розпізнавання фруктових плодів за допомогою вдосконаленої мережі YOLOv5-v1. (а) Хмарно (б) Бічне сонячне світло (в) Підсвічування сонцем (г) Сонячний фронт

Обмежувальні рамки зеленого кольору використовувались, як мітки, які можна захопити, тоді як сині використовувались, як мітки, які неможливо захопити. Як видно на рис. 2, запропонована модель нейронної мережі виявляє структурні об'єкти на зображеннях знятих при рівномірному освітленні, похмурому освітленні, а також при сонячному освітленні. Крім цього, поточна модель виявляє структурні об'єкти в умовах переднього, заднього та бічного освітлення при сонячному освітленні.

Результати виявлення та обчислення кількості фруктових плодів наведено в таблиці 3.

Таблиця 3

**Результати експериментів щодо виявлення та обчислення кількості фруктових плодів**

№	Погодні умови	Виявлено та обчислено, шт.	Виявлено, але не обчислено, шт.	Виявлено багатьма камерами, але обчислено одноразово, шт.	Всього, шт.
1	Сонце (сонце в zenіті)	147 382	72 143	5 743	~160 000
2	Похмуро	145 223	69 944	6 611	~160 000
3	Сонце (сильна затіненість)	132 304	69 336	8 442	~160 000

У результаті проведених експериментів було правильно виявлено та обчислено 147 382, 145 223 та 132 304 (таблиця 3) фруктових плодів за трьох різних погодних умов. Водночас встановлено, що певна кількість фруктових плодів була виявлена, але не обчислена, а саме 72 143, 69 944 та 69 336 для трьох погодних умов; ці об'єкти були ідентифіковані детектором автоматизованої системи як фруктові плоди, проте не враховані під час обрахунку, оскільки вони перебували поза поточними робочими зонами. Певну кількість плодів, а саме 5 743, 6 611 та 8 442 за трьох погодних умов, було одночасно виявлено камерами двох або трьох БПЛА та обчислено одноразово, відповідно до логіки запропонованого методу обчислення структурних об'єктів.

Для подальшого аналізу ефективності розпізнавання запропонованого підходу у роботі вдосконалену мережу YOLOv5-v1 було порівняно з оригінальною мережею YOLOv3 [6], YOLOv4 [7] та YOLOv5 [8] на 200 зображеннях тестового набору. В якості оціночних показників були прийняті значення та середня швидкість розпізнавання моделі. Результати виявлення, розмір і кількість параметрів кожної моделі мережі наведені в таблиці 4.

Відповідно до таблиці 4, значення mAP для вдосконаленої моделі розпізнавання YOLOv5-v1, запропонованої в дослідженні, було найвищим, і воно було вищим, ніж у вихідній мережі YOLOv5, і вищим, ніж у мережах YOLOv3, YOLOv 4 і EfficientDet-D0 відповідно, що вказує на те, що запропонований метод є найкращим для задачі розпізнавання структурних об'єктів серед п'яти інших методів. Для швидкості розпізнавання моделі середня швидкість виявлення вдосконаленої моделі YOLOv5-v1 за зображенням (66,7 кадрів/с) (кадр/с, кадрів на секунду), яка становила 2,53 і час мережі EfficientDet-D0, YOLOv4 і YOLOv3 відповідно, вказуючи, що запропонована модель нейронної мережі може задовольнити вимоги групи БПЛА для розпізнавання плодів у режимі реального часу.

Таблиця 4 Результати порівняння моделей виявлення фруктових плодів

Мережі виявлення об'єктів	mAP (%)	Середня швидкість виявлення (s/ріс)	Кількість параметрів	Розмір моделі (Мб)
YOLOv3 [6]	71,80	0,053	6,15×107	235,0
YOLOv4 [7]	82,01	0,017	6,39×107	244,0
YOLOv5 [8]	81,70	0,013	7,25×106	14,0
YOLOv5-v1	86,75	0,015	6,52×106	12,7

У підсумку, побудована модель YOLOv5-v1 у рамках запропонованого підходу краще підходить для розгортання на вбудованих пристроях і може легко адаптуватися до змінних погодних умов у фруктовому саду.

#### Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

У поданій роботі запропоновано новий підхід до виявлення та обчислення кількості заданих структурних об'єктів на прикладі фруктових плодів за зображеннями, що генеруються безпілотними літальними апаратами в режимі реального часу. Підхід полягає в послідовному виконанні трьох розроблених методів та одного способу. У результаті проведених експериментальних досліджень отримано, що оцінка синхронізації відеопотоків за індексом SSIM коливається від 0,79 до 0,92, із середнім значенням 0,87, а за індексом PSNR – від 22 до 39, що свідчить про високу ефективність розробленого методу синхронізації та хорошу якість отриманих об'єднаних зображень. Результати експериментів на тестовому наборі даних демонструють, що вдосконалена модель YOLOv5-v1, як результат впровадження запропонованого підходу, може ефективно розпізнавати фрукти, зняті камерами БПЛА. Модель досягла 92,13%, 84,59%, 87,94% і 89,02% показників запам'ятовування, точності, mAP і класифікації F1 відповідно, а середня швидкість виявлення становила 0,015 секунди на кадр у відеопослідовності. Отримані значення свідчать про високий показник ефективності методів збору та обробки зображень отриманих із використанням БПЛА для виявлення заданих об'єктів.

Подальші дослідження також можуть бути проведені для включення додаткових ознак, таких як форма фрукта, використовуючи більший набір даних для навчання моделі.

#### Література

1. Medvedeva Y., Kucher A., Lipsa J., Heldak M. Human health risk assessment on the consumption of apples growing in urbanized areas: Case of Kharkiv, Ukraine. *International Journal of Environmental Research and Public Health*. – 2021. – Vol. 18, No. 4. – P. 1504.
2. A litchi fruit recognition method in a natural environment using RGB-D images / L. Yu et al. *Biosystems Engineering*. – 2021. – Vol. 204. – P. 50–63.
3. Mai X., Zhang H., Jia X., Meng M.Q.-H. Faster R-CNN with classifier fusion for automatic detection of small fruits. *IEEE Transactions on Automation Science and Engineering*. – 2020. – Vol. 17, No. 3. – P. 1555–1569.
4. Wang S. Research towards YOLO-series algorithms: Comparison and analysis of object detection models for real-time UAV applications. *Journal of Physics: Conference Series*. – 2021. – Vol. 1948, No. 1. – P. 012021.
5. Huang Z., Zhang P., Liu R., Li D. Immature apple detection method based on improved YOLOv3. *ASP Transactions on Internet of Things*. – 2021. – Vol. 1, No. 1. – P. 9–13.
6. Chen W., Zhang J., Guo B., Wei Q., Zhu Z. An apple detection method based on Des-YOLO v4 algorithm for harvesting robots in complex environments. *Mathematical Problems in Engineering*. – 2021. – Vol. 2021. – P. e7351470.
7. Behera S.K., Mishra N., Sethy P.K., Rath A.K. On-tree detection and counting of apple using color thresholding and CHT. *2018 International Conference on Communication and Signal Processing (ICCSP-2018) : Proceedings*. – Vol. 2018. (India, Chennai, 3-5 April 2018). – IEEE Inc., 2018. – P. 0224–0228.

#### References

1. Medvedeva Y., Kucher A., Lipsa J., Heldak M. Human health risk assessment on the consumption of apples growing in urbanized areas: Case of Kharkiv, Ukraine. *International Journal of Environmental Research and Public Health*. – 2021. – Vol. 18, No. 4. – P. 1504.
2. A litchi fruit recognition method in a natural environment using RGB-D images / L. Yu et al. *Biosystems Engineering*. – 2021. – Vol. 204. – P. 50–63.
3. Mai X., Zhang H., Jia X., Meng M.Q.-H. Faster R-CNN with classifier fusion for automatic detection of small fruits. *IEEE Transactions on Automation Science and Engineering*. – 2020. – Vol. 17, No. 3. – P. 1555–1569.
4. Wang S. Research towards YOLO-series algorithms: Comparison and analysis of object detection models for real-time UAV applications. *Journal of Physics: Conference Series*. – 2021. – Vol. 1948, No. 1. – P. 012021.
5. Huang Z., Zhang P., Liu R., Li D. Immature apple detection method based on improved YOLOv3. *ASP Transactions on Internet of Things*. – 2021. – Vol. 1, No. 1. – P. 9–13.
6. Chen W., Zhang J., Guo B., Wei Q., Zhu Z. An apple detection method based on Des-YOLO v4 algorithm for harvesting robots in complex environments. *Mathematical Problems in Engineering*. – 2021. – Vol. 2021. – P. e7351470.
7. Behera S.K., Mishra N., Sethy P.K., Rath A.K. On-tree detection and counting of apple using color thresholding and CHT. *2018 International Conference on Communication and Signal Processing (ICCSP-2018) : Proceedings*. – Vol. 2018. (India, Chennai, 3-5 April 2018). – IEEE Inc., 2018. – P. 0224–0228.