



О. О. Боровий, Р. І. Ілечко

Національний університет "Львівська політехніка", Львів, Україна

СТРУКТУРА СИСТЕМИ АВТОНОМНОГО УПРАВЛІННЯ БЕЗПІЛОТНИМ ЛІТАЛЬНИМ АПАРАТОМ ЗАСОБАМИ КОМП'ЮТЕРНОГО ЗОРУ

Розглянуто особливості побудови фізичної моделі безпілотного літального апарату для завдань комп'ютерного зору. Проаналізовано наявні фізичні моделі, вказано на їх сильні та слабкі особливості. Описано наявні методи та засоби планування маршрутів і навігації. Окреслено попередні досягнення у сфері зменшення розміру нейромережних моделей для завдань комп'ютерного зору, а саме, лінійне квантування нейромереж та обчислення з половинною точністю. Визначено підходи та принципи розроблення системи автономного управління безпілотного літального апарату. Розроблено структуру системи автономного управління безпілотного літального апарату. Зазначено характеристики комп'ютерних засобів системи автономного управління безпілотного літального апарату. Запропоновано концепцію побудови алгоритму корекції геоїдних на підставі зображення з камери, супутникових знімків, інерційно вимірювальних блоків і системи глобального позиціонування GPS (англ. *Global Positioning System*). Вдосконалено метод нейромережового розпізнавання об'єктів шляхом розпаралелення процесу оброблення відеоданих з використанням декількох цифрових пришвидшувачів. Запропоновано сценарій використання системи з великою кількістю апаратних пришвидшувачів. Описано продуктивність моделей до та після квантування в контексті точності розпізнавання. Представлено метод квантування частинами для оброблення надвеликих нейромереж комп'ютерного зору. Запропоновані методи та технології демонструють високий потенціал для підвищення ефективності та точності оброблення даних у реальних умовах експлуатації. Покращені підходи до квантування нейромереж та їх оптимізації дають змогу значно зменшити розмір моделей та підвищити швидкість їх роботи, зберігаючи при цьому високу точність розпізнавання. Інтеграція кількох цифрових пришвидшувачів дає змогу ефективно розпаралелити процес оброблення відеоданих, що особливо важливо для оперативного аналізу інформації в режимі реального часу. Результати цього дослідження можна використати для покращення можливостей навігації та оброблення відеоданих, що сприятиме розширенню сфер застосування безпілотних літальних апаратів.

Ключові слова: візуальна геолокація; нейромережі; квантування; апаратні пришвидшувачі.

Вступ / Introduction

Розвиток технологій безпілотних літальних апаратів (БПЛА) або дронів надає нові можливості для різноманітних сфер людської діяльності, в т.ч. й військової галузь, сільське господарство, логістику, рятувальні операції, моніторинг навколишнього середовища та інші. Зі збільшенням складності завдань, які покладуються на БПЛА, зростає потреба у підвищенні їх автономності. Одним із ключових напрямків удосконалення є впровадження систем автономного управління, що здатні приймати рішення на підставі аналізу оточення у реальному часі.

Комп'ютерний зір, як галузь штучного інтелекту, відкриває широкі можливості для підвищення рівня автономності БПЛА. Використання алгоритмів комп'ютер-

ного зору дає змогу дронам не тільки виявляти та розпізнавати об'єкти, а й аналізувати їхні просторові характеристики, планувати траєкторії польоту та уникати перешкоди. Отже, інтеграція засобів комп'ютерного зору в системи управління БПЛА забезпечує більш точну і надійну автономну навігацію.

Технології БПЛА та комп'ютерного зору досягли значного прогресу, виконуючи широкий спектр завдань. Значний інтерес до цих технологій спонукав науковців розробляти новітні підходи та рішення, які б могли ефективно використовувати потужності сучасної обчислювальної техніки на борту БПЛА. Однак, незважаючи на активні дослідження у цій сфері, на наш погляд, наявні апаратно-програмні рішення не є оптимальними.

Інформація про авторів:

Боровий Орест Олегович, магістр, аспірант, кафедра автоматизованих систем управління.

Email: orest.o.borovyi@lpnu.ua; <https://orcid.org/0009-0008-2069-5336>

Ілечко Роман Ігорович, магістр, аспірант, кафедра автоматизованих систем управління.

Email: orest.o.borovyi@lpnu.ua; <https://orcid.org/0009-0006-8208-5109>

Цитування за ДСТУ: Боровий О. О., Ілечко Р. І. Структура системи автономного управління безпілотним літальним апаратом засобами комп'ютерного зору. Науковий вісник НЛТУ України. 2024, т. 34, № 5. С. 69–77.

Citation APA: Borovyi, O. O., & Ilchko, R. I. (2024). Structure of autonomous control system for unmanned aerial vehicle using computer vision tools. *Scientific Bulletin of UNFU*, 34(5), 69–77. <https://doi.org/10.36930/40340509>

Об'єкт дослідження – збирання, опрацювання та прийняття рішень у системі автономного управління безпілотним літальним апаратом.

Предмет дослідження – методи, алгоритми та засоби збирання, опрацювання та прийняття рішень у системі автономного управління безпілотним літальним апаратом з використанням засобів комп'ютерного зору, що забезпечить стабільну і точну навігацію, уникнення перешкод та автономне виконання місії.

Мета роботи – розробити структуру системи автономного управління безпілотним літальним апаратом з використанням засобів комп'ютерного зору що дасть змогу виконувати точну навігацію, уникати перешкоди, автономно виконувати місії та збирати дані в реальному часі.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- визначити підходи та принципи розроблення системи автономного управління БПЛА, що дасть змогу сформулювати вимоги до конкретних структур та реалізації систем автономного управління БПЛА на підставі наявних апаратних та програмних засобів;
- розробити структуру системи автономного управління БПЛА, що дасть змогу виконувати комплексні завдання з високою автономністю, зокрема навігацію в складних умовах, автономне прийняття рішень та інтеграцію з наземними системами для обміну даними в реальному часі;
- проаналізувати наявні апаратні рішення для побудови системи автономного управління БПЛА, що забезпечить стабільну і точну навігацію, уникнення перешкод, автономне виконання місії та збір даних у реальному часі;
- запропонувати концепцію побудови алгоритму візуальної геолокації, яка допоможе покращити автономність БПЛА в умовах дії завад;
- вдосконалити метод нейромережевого розпізнавання об'єктів унаслідок розпаралелення процесу оброблення відеоданих, що уможливить ефективно використовувати декілька пришвидшувачів нейромережових обчислень одночасно.

Аналіз останніх досліджень та публікацій. У роботі [25] автор досліджує використання згорткових нейронних мереж (CNN) для виявлення об'єктів на зображеннях з безпілотних літальних апаратів. Дослідники зосередилися на оцінюванні ефективності популярних методів SSD [13], Faster R-CNN [17] і RetinaNet [12] на Stanford Drone Dataset [18]. Хоча цей підхід забезпечує ефективний моніторинг у реальному часі, він можливий завдяки запропонованому Fog/Edge вузлу, який виступає в ролі наземної станції. Незважаючи на низку переваг, на практиці перед пілотами часто постають проблеми через затримку передачі даних, неможливість роботи в умовах радіозвад.

Використовуючи ідею наземної станції в ролі обчислювального центру автори роботи [5] пропонують новий багатошаровий ML/AI фреймворк, для виконання завдань комп'ютерного зору у реальному часі на системах БПЛА. Платформа спрощує інтеграцію AI-додатків на базі БПЛА. Перевагою використання DJI Phantom 4 та Matrice-100 є пришвидшення досліджень, ціною цього – втрата гнучкості апаратного середовища БПЛА. Наземна платформа функціонує на пристроях Samsung Galaxy Note 4 та ноутбуках, обладнаних графічними процесорами NVIDIA GeForce GTX 1070.

На противагу попереднім підходам, що активно використовували готові БПЛА масового вжитку, в дослід-

жені у роботі [26] науковці сконструювали власний гексокоптер на підставі польотного контролера Pixhawk та Raspberry Pi 3B+ у ролі процесора цифрового оброблення. Незважаючи на підвищену гнучкість у підборі компонент, вузьким місцем цієї конфігурації є саме Raspberry Pi 3B+. Хоча використання згаданого SBC (англ. *Single Board Computer*) у комбінації з методами комп'ютерного зору можливе, проте є сильно обмежене через низьку обчислювальну здатність та відсутність апаратних рішень, оптимізованих для нейрообчислень. Звісно, можливим підходом є використання зовнішніх TPU [20], NPU [24] через USB інтерфейс, але штатні виходи USB2.0 мають швидкість передачі сигналів 480 Mbit/s, що на практиці недостатньо для завдань реального часу.

У роботі [15] представлено недорогу вбудовану систему на підставі комп'ютерного зору для БПЛА, спрямовану на покращення автономної навігації за допомогою інтегрованої системи посадки та уникнення перешкод. Науковці використовують Raspberry Pi 4 для оброблення зображень, LIDAR Lite v3 для вимірювання відстані та камеру NoIR для візуалізації. Зв'язок між БПЛА та наземною станцією встановлюється через протокол MAVLink, що дає змогу передавати дані в режимі реального часу. Систему було протестовано на квадрокоптері, оснащеному Pixhawk, вона продемонструвала ефективну роботу в контрольованих умовах. Так, модель пріоритезує енергоефективність та низькі обчислювальні вимоги, що робить її прийнятним рішенням для дистанційного моніторингу та спостереження, наявність комплексного USB3.0, зі швидкістю передачі сигналу 5 Гбіт/с, дає змогу використовувати енергоефективні NPU та TPU. Очевидними недоліками такого підходу є максимально можлива кількість цифрових пришвидшувачів, гомісткість USB версії TPU та дорожнеча порівняно з акселератором на вхідних інтерфейсах PCIe чи M.2.

У роботі [14] автори пропонують вдосконалену обчислювальну модель для підвищення можливостей візуальної посадки БПЛА, використовуючи об'єктне розпізнавання (OD) на підставі YOLO [16]. Модель інтегрує пришвидшення TensorRT для швидшого оброблення та багатопотокове програмування для ефективної передачі зображень на наземні станції. Вона використовує автоматизовану систему стеження за нахилом камери, що покращує продуктивність розпізнавання об'єктів на бортовому комп'ютері NVIDIA Jetson Xavier NX [21]. Основні компоненти моделі містять різні архітектури OD, які порівнюють за точністю та продуктивністю у FPS, і котрі демонструють високу точність у розпізнаванні посадкових точок у різних середовищах. Такий підхід значно скорочує час передачі даних, забезпечуючи виявлення та прийняття рішень в реальному часі для операцій посадки БПЛА. Слабким аспектом таких систем є активне використання згаданих бортових комп'ютерів. Незважаючи на те, що пришвидшувачі сімейства Jetson надають можливість запуску нейромережових алгоритмів без попередніх адаптацій, а конвертація моделі в TensorRT збільшує швидкість роботи, недоліком є низька енергоефективність порівняно з аналогами, висока вартість пришвидшувачів.

Отже, проаналізувавши наявні публікації, можна дійти висновку, що попри широке різноманіття, запропоновані конфігурації БПЛА платформ мають істотні

недоліки, що унеможливило оптимальне покриття широкого кола проблем комп'ютерного зору.

Результати дослідження та їх обговорення / Research results and their discussion

Вибір підходів і принципів розроблення системи автономного управління БПЛА. Система автономного управління БПЛА з використанням засобів комп'ютерного зору повинна забезпечувати можливість орієнтуватися в просторі, виявляти та уникати перешкоди, а також виконувати інші завдання без втручання людини. Розроблення системи пропонуємо здійснювати на підставі таких підходів і принципів:

- **обчислювальна потужність:** система повинна мати достатню обчислювальну потужність для оброблення даних в режимі реального часу, охоплюючи оброблення сигналів з сенсорів, планування траєкторії та прийняття рішень;
- **енергоефективність:** вимоги до енергоспоживання мають бути оптимізовані для забезпечення тривалого часу автономного польоту за обмеженої місткості батарей;
- **система навігації:** БПЛА має бути оснащений високоточними навігаційними системами, такими як GPS та інерціальні навігаційні системи (ІНС), для забезпечення точного позиціонування та руху за заданими маршрутами;
- **оброблення даних з сенсорів:** система має забезпечувати інтеграцію даних з різних сенсорів (камери, лідари, радарми) для створення комплексної картини навколишнього середовища для того, щоб виступити резервним джерелом інформації про позиціонування БПЛА у разі дії перешкод, або у випадку втрати сигналу GPS;
- **система автопілота:** БПЛА має бути оснащений відповідним програмно-апаратним комплексом, який дає змогу програмувати маршрути руху дрона та порядок його дій. Така система повинна приймати інформацію з навігаційної системи та інших сенсорів, щоб виконувати запрограмований маршрут автономно без участі оператора.

Формулювання вимог до системи автономного управління БПЛА є критичним етапом у забезпеченні її ефективної та безпечної роботи. Дотримання зазначених вимог дасть змогу створити надійну систему, здатну виконувати завдання різної складності в широкому спектрі умов.

Розроблення структури системи автономного управління БПЛА. Структура системи автономного безпілотного літального апарату (рис. 1) складається з кількох частин. У центрі знаходиться система автопілота, яка отримує сигнали з інших підсистем і приймає рішення з управління двигунами БПЛА для виконання плану польоту. Підсистеми, які дають сигнали автопілота, складаються з GPS-приймача, IMU, системи радіозв'язку з наземною станцією, а також з комп'ютера-компаньйона, який водночас отримує відеопотік з системи камери. Ця фізична модель призначена для того, щоб у випадку постанови радіозавад БПЛА міг автономно продовжувати план польоту, або повернутись на точку зльоту.

Важливими компонентами цієї системи задля досягнення високого ступеня автономності є давачі інерціальної системи IMU (англ. *Inertial Measurement Unit*) та система денної камери з комп'ютером компаньйоном. Використовуючи дані з гіроскопа, акселерометра, барометра і магнітометра, можна побудувати інерціальну навігаційну систему, яка б, використо-

вуючи ці дані, обчислювала вектор і швидкість руху, в такий спосіб могла б розрахувати місцезнаходження БПЛА просторі. Проте не існує таких давачів, які б не давали похибки, особливо доступні MEMS давачі, які присутні на масових моделях БПЛА, тому така система не може бути основною і потребує постійної корекції. У звичайному режимі таким модулем корекції виступає GPS-приймач, який здатен отримати позицію об'єкта у просторі. Отже, похибка інерціальної системи постійно коригується й автопілот БПЛА здатен автономно виконувати наперед запрограмований план польоту без безпосередньої участі оператора. В умовах завад GPS корекція неможлива, бо сигнал або глушиться шумом, або справжній сигнал GPS супутників підміняється сигналом від РЕБ системи, в такий спосіб вводячи в оману GPS приймач про його справжнє розташування. У таких умовах потрібен альтернативний метод геолокації, який би спирався на альтернативні джерела інформації. Йдеться про систему візуальної геолокації, побудована на підставі денної камери та комп'ютера, здатного ефективно працювати з системами комп'ютерного зору, побудованими на базі глибоких нейронних мереж. У такій системі інформація з камери обробляється комп'ютером задля отримання координат місцезнаходження БПЛА, який використовується як альтернатива сигналу GPS приймача.

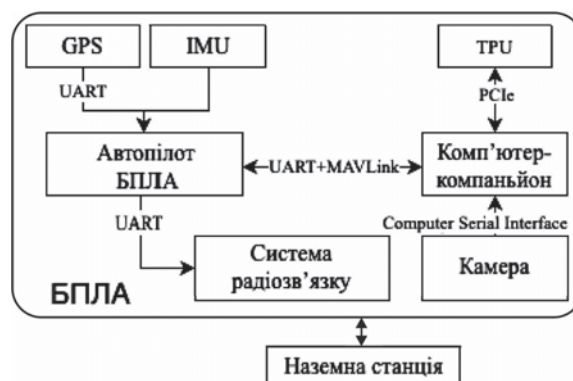


Рис. 1. Структура системи автономного управління безпілотного літального апарату / Structure of autonomous control system for an unmanned aerial vehicle

Характеристики комп'ютерних засобів системи автономного управління БПЛА. Сучасні методи комп'ютерного зору використовують методи глибокого навчання, такі як згорткові нейронні мережі. Такі методи вимагають значних обчислювальних ресурсів, тому для застосування у сфері БПЛА це надзвичайно важливо, адже існують суворі обмеження щодо ваги та розмірів такого комп'ютера, а також обмеження у споживанні електроенергії. Перспективними є одноплатні комп'ютери. Одноплатний комп'ютер (SBC) – це повноцінний комп'ютер, побудований на одній друкованій платі, з усіма необхідними компонентами, такими як мікропроцесор, пам'ять, введення/виведення (I/O) та іншими важливими функціями, необхідними для функціонування комп'ютера. На відміну від традиційних настільних комп'ютерів, які зазвичай використовують материнську плату з окремими компонентами, SBC інтегрує всі ці елементи на одну плату. Вони призначені для вбудованих додатків і специфічних використаннях, де критичними є компактний розмір, низьке споживання енергії та економічність. Сучасні SBC оснащені більш потужними процесорами, в т.ч. й ба-

гатовані ARM-процесори, процесори архітектури x86 та навіть деякі з інтегрованим графічним процесором для кращої графічної продуктивності.

Проте потужного мікропроцесора (CPU) недостатньо задля ефективного обчислення нейронних мереж. TPU, NPU та GPU є кращими для машинного навчання, ніж CPU, завдяки їхній спеціалізованій архітектурі, що дає змогу ефективно виконувати паралельні обчислення, характерні для нейронних мереж. GPU [17] мають тисячі малих ядер, що забезпечує високу обчислювальну потужність для тренування та використання моделей, тоді як TPU розроблені Google для оптимізації тензорних операцій, особливо в екосистемі TensorFlow, забезпечують високу продуктивність і енергоефективність. NPU, з іншого боку, створені для ефективного виконання завдань нейронних мереж у мобільних та вбудованих системах, забезпечуючи високу ефективність за мінімального споживання енергії. Ці процесори перевершують CPU у виконанні завдань машинного навчання завдяки кращому паралельному обробленню, спеціалізованим оптимізаціям і значно вищій продуктивності.

Прикладами одноплатних комп'ютерів, які мають TPU, NPU або GPU, є Radxa Rock5b з восьм'ядерним ARM процесором RK3588, який складається з чотирьох ядер Cortex-A76 та чотирьох Cortex-A55 ядер з NPU з потужністю шість TOPs(int8) або 0.5 TFLOPs(float16) і піковим споживанням до 10 Вт. Також NVIDIA Jetson TX2, який містить GPU з потужністю 1.33 TFLOPs зі споживанням до 15 Вт. Окрім цього, компанія Google розробила TPU Coral, які існують у версії плати розширення з потужністю 8 TOPs(int8) на один TPU за споживання одного модуля у 2 Вт. Таку плату розширення можна додати до SBC, в такий спосіб отримуючи додаткові обчислювальні можливості до одноплатного комп'ютера.

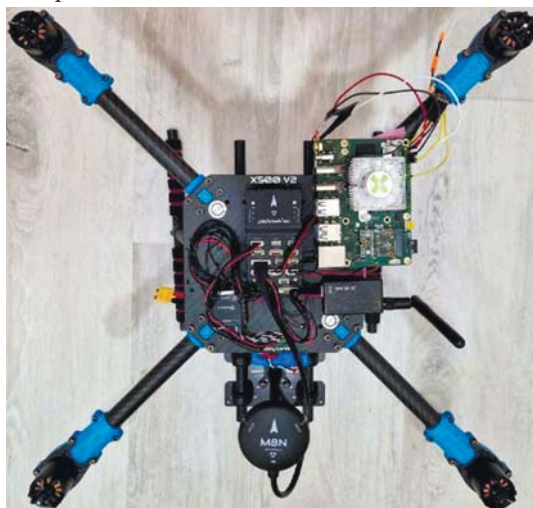


Рис. 2. Фото експериментальної платформи / Photo of the experimental platform

Для експериментальної платформи (рис. 2) було обрано комбінацію з процесора RK3588 на платі Radxa Rock5b та три модулі Google Edge TPU підключених через інтерфейс PCIe. Така система дає змогу розділити обчислення на декілька модулів, саме цим покращити продуктивність.

Засоби навігації та планування маршрутів. Система навігації та планування місії БПЛА є ключовими компонентами, які забезпечують автономну роботу та

виконання завдань. Вони використовують GPS, IMU та алгоритм візуальної геолокації. Навігаційна система БПЛА містить GPS-модуль, який надає дані про географічне положення апарату, що дає змогу слідувати заздалегідь заданим маршрутом. Це особливо важливо для завдань, які потребують точного позиціонування, таких як картографування, моніторинг об'єктів і доставлення вантажів. GPS-модуль постійно отримує сигнали від супутників, що дає змогу визначати місцезнаходження з точністю до кількох метрів.

Інерційний вимірювальний блок доповнює GPS-модуль, надаючи дані про кутову швидкість, пришивидження та орієнтацію БПЛА. IMU використовує акселерометри та гіроскопи для вимірювання рухів апарату у трьох осях. Це дає змогу визначати положення та орієнтацію БПЛА навіть у випадках, коли GPS-сигнал відсутній або недоступний, наприклад, під час польоту в тунелях або під густими деревами. Інформація з IMU постійно обробляється для корекції курсу та стабілізації польоту. Візуальна геолокація допомагає коригувати покази IMU з часом, забезпечуючи стабільне та точне позиціонування.

Система планування місії (англ. *Mission Planner*) дає змогу операторам БПЛА створювати, змінювати та керувати маршрутами польоту. Вона інтегрується з навігаційною системою, використовуючи дані GPS, IMU та візуальної геолокації для розроблення оптимальних маршрутів і завдань. Оператори можуть визначати точки маршруту, висоту, швидкість та інші параметри польоту, а також завантажувати ці дані на бортовий комп'ютер БПЛА. Mission Planner також надає можливість моніторингу та корекції польоту в реальному часі. Оператори можуть бачити поточне знаходження БПЛА, відстежувати його маршрут та вносити корективи за потреби. Це особливо важливо для виконання складних місій у змінних умовах.

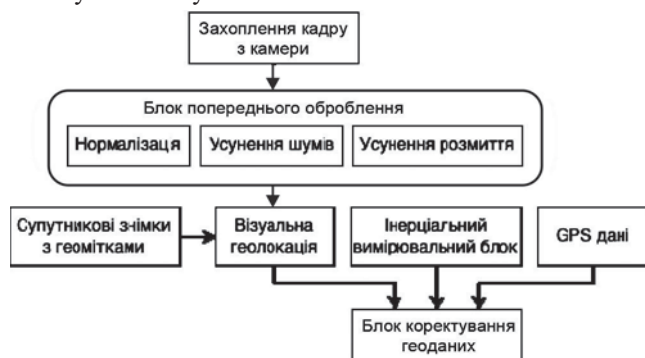


Рис. 3. Схема роботи алгоритму корекції координат / Workflow diagram of the coordinate correction algorithm

Запропонована концепція побудови алгоритму корекції геоданих базується на тому, що до системи автопілота інтегрується альтернативний метод геолокації, побудований на засобах комп'ютерного зору. На схемі (рис. 3) видно, що система захоплює зображення з камери. Після чого зображення піддається попередньому обробленню, а саме нормалізації зображення, усуненню шумів і розмиття. Підготовлене зображення потрапляє на вхід алгоритму візуальної геолокації разом з еталонним знімком з супутника для порівняння. Внаслідок роботи алгоритму отримуємо географічні координати, які подаються як резервні до блоку корекції геоданих. Блок отримує координати з GPS приймача та IMU і порівнює GPS дані з очікуваними, порахованими за допомогою

інерціального вимірювального блоку. Якщо очікувані виміри відрізняються від наявних, то дані беруть з резервного джерела геолокації. У нашому випадку це алгоритм візуальної геолокації.

Вдосконалення методу нейромережевого розпізнавання об'єктів за рахунок розпаралелення процесу оброблення відеоданих. У сучасних завданнях комп'ютерного зору, які виконують на БПЛА, ефективність апаратної платформи є ключовим чинником для забезпечення високої продуктивності та швидкості оброблення даних. Стандартні конфігурації пришвидшувачів часто виявляються недостатньо оптимальними, що обмежує потенціал застосованих алгоритмів. У цьому розділі розглянуто можливість покращення обчислювальних процесів шляхом використання інноваційних апаратних рішень, які дають змогу реалізувати більш ефективні паралельні обчислення. Зокрема, запропоновано конфігурацію, що містить сучасні цифрові пришвидшувачі та процесори, що значно підвищують ефективність виконання різноманітних завдань комп'ютерного зору на БПЛА.

Квантування нейронних мереж [10] – це метод трансформації чисел з рухомою комою, таких як float32, у більш компактні цілочисельні формати. Основна мета – це зменшення обчислювальної складності та обсягу пам'яті, необхідних для зберігання та оброблення моделей. Найбільш поширеним підходом є лінійне квантування, яке містить афінні та симетричні методи. При квантуванні значення перетворюються за формулою:

$$X_q = \left\lfloor \frac{X}{S} + Z \right\rfloor,$$

де: X_q – квантоване значення; S – масштабуючий коефіцієнт; Z – нульова точка. У разі потреби, обернену трансформацію, для відновлення початкових значень, здійснюють як:

$$X = S(X_q - Z).$$

У випадку симетричного квантування нульова точка Z дорівнює нулю, а коефіцієнт розраховують як:

$$S = \frac{|r_m|}{2^N - 1},$$

де: r_m – максимальне значення вхідних даних; N – кількість бітів для цілочисельного представлення. Цей метод є особливо ефективним для нейронних мереж із симетричними значеннями вагових коефіцієнтів, оскільки нульова точка знаходиться посередині діапазону значень. У несиметричному методі квантування параметри описують так:

$$S = \frac{r_{\max} - r_{\min}}{q_{\max} - q_{\min}},$$

де: $r_{\min}$, $r_{\max}$ – межі вхідних значень; $q_{\min}$, $q_{\max}$ – межі цілочисельного діапазону. Оскільки позиція нульової точки важлива для нейромережових алгоритмів розглянемо її розташування за цих несиметричних трансформацій. У цьому випадку здійснюється пропорційний зсув точки нуля, що є важливим для нейромереж з несиметричними функціями активації, такими як ReLU, оскільки це дає змогу мінімізувати втрати точності.

Оптимізація констант квантування є критичним етапом і може здійснюватися двома способами: під час навчання (quantization aware training) або після навчання (англ. *Post-Training Quantization*) за допомогою репрезентативного набору з вибірки навчальних даних.

Найпоширенішим варіантом є перехід від типу float32 до int8, що дає змогу зменшити розмір моделі в чотири рази за мінімальних втрат точності, які зазвичай не перевищують 5 % за умови правильної оптимізації констант.

У запропоновану конфігурацію фізичної системи БПЛА, не менш важливу роль відіграє Radxa Rock 5b. Незважаючи на те, що програмно є можливість запуску квантованих моделей, вбудований NPU не оптимізований під цілочисельні обчислення. Проте, для зменшення розміру нейромереж, ми можемо використати обчислення з половинною точністю [2]. Обчислення з половинною точністю забезпечує ефективне оброблення даних за обмежених ресурсів. Використання чисел з половинною точністю (16-бітних) замість традиційних 32-бітних дає змогу значно зменшити вимоги до пам'яті та підвищити швидкість обчислень. Застосування Half-precision Inference на Radxa Rock Pi 5B з NPU процесором дає змогу оптимізувати паралельні обчислення, що є критичним для виконання інтенсивних завдань комп'ютерного зору в режимі реального часу. Такі обчислення знижують споживання пам'яті, зменшують обсяг передачі даних і підвищують пропускну здатність обчислювальних блоків, що сприяє більш ефективному використанню апаратних ресурсів. Завдяки цьому обчислення з половинною точністю сприяють підвищенню продуктивності та енергоефективності, що є ключовим чинником для тривалого автономного функціонування БПЛА.

Використання описаних алгоритмів у комбінації з запропонованою фізичною моделлю дрона, оснащеної декількома апаратними пришвидшувачами, значно підвищує продуктивність завдяки паралельному виконанню завдань. Одним з ключових сценаріїв використання є застосування кількох нейронних мереж, які працюють незалежно одна від одної. Такий підхід дає змогу оптимізувати обчислювальні ресурси та забезпечити більш ефективне оброблення даних в режимі реального часу. Наприклад, під час виконання завдань моніторингу, одна нейронна мережа займатиметься виявленням об'єктів, тоді як інша – прогнозуванням траєкторії руху. Це забезпечує зниження латентності оброблення та підвищення загальної продуктивності системи, адже кожен апаратний пришвидшувач призначається для конкретного завдання, без потреби очікування завершення інших процесів. Згадана архітектура дає змогу значно підвищити ефективність використання енергії, знижуючи при цьому ризик перевантаження окремих компонентів системи.

Наступним варіантом є використання нейромережі з рівнобіжними вітками обчислень, як от, "сіамських". У запропонованій конфігурації дрона, з кількома апаратними пришвидшувачами це надає ще одну важливу перевагу – можливість паралельного оброблення кількох зображень. "Сіамські" нейромережі [11], що часто застосовуються для визначення подібностей між парами зображень, містять етап виокремлення ознак, який може бути істотно пришвидшений завдяки одночасному обробленню на різних апаратних пришвидшувачах. Наприклад, у сценарії розпізнавання та ідентифікації об'єктів дрон може одночасно обробляти декілька зображень з різних камер або різних часових моментів, виділяючи ознаки зображень паралельно. Це значно зни-

жує тривалість оброблення, оскільки різні частини моделі можуть працювати одночасно, не очікуючи завершення попередніх обчислень. Отже, нова модель БПЛА забезпечує значні переваги в ефективності та швидкості оброблення даних, що є критично важливим для завдань моніторингу, навігації та аналізу складних сценаріїв у реальному часі.

Застосування ансамблевих моделей у новій моделі дрона з кількома апаратними пришвидшувачами також може надати значну перевагу. Ансамблеві моделі, що складаються з кількох базових моделей, працюють разом для покращення точності та надійності прогнозів. У цьому сценарії БПЛА одночасно обчислюватиме декілька базових моделей на різних апаратних пришвидшувачах, що дає змогу паралельно обробляти дані та отримувати прогнози від кожної моделі. Це зменшує тривалість оброблення та підвищує точність остаточного рішення, оскільки результати різних моделей можуть бути об'єднані для формування більш точного та надійного прогнозу. Оскільки кожна з моделей має різний ступінь деградації (втрати точності) після квантування, цей чинник потрібно враховувати. Тоді фінальне прогнозування виглядатиме як:

$$Y = \frac{1}{k} \sum_{i=1}^k \varepsilon_i \alpha_i y_i,$$

де: k – кількість базових мереж; ε_i – коефіцієнт деградації; α_i – відповідний чинник впливу; y_i – прогнозування кожної з мереж.

Для коректної роботи алгоритмів комп'ютерного зору на тензорних пришвидшувачах необхідний процес квантування. Розглянемо загальноприйнятий підхід квантування. Нехай G спрямований граф нейромережі, x – вхідне зображення, $x \in H \times W \times N$, де H , W висота та ширина зображень, N розмір вибірки квантування.

$$o_f = G_f(x),$$

де o_f – вихід нейромережі для входу x . Введемо операцію квантування Q , тоді константи квантування Z_f , Z_f для цілого графа G_f визначаємо як:

$$S_f, Z_f = Q(G_f(x)).$$

Незважаючи на це, під час квантування розмір моделі зменшується приблизно в чорити рази, часто цього недостатньо. Непоодинокими є ситуації, коли цифрові пришвидшувачі, як от Coral TPU, не в змозі завантажити нейромережі через їх підвищений розмір. Для вирішення цієї проблеми ми пропонуємо квантування нейромереж частинами. Розглянемо згаданий спрямований граф G_f як кортеж відповідних підграфів G , тоді:

$$G = (g_1, g_2, \dots, g_j), k \leq t, k \in N,$$

де t – кількість TPU.

Тоді вихід підграфа g_1 для вхідного зображення x представимо так:

$$o_1 = g_1(x).$$

Очевидно, що o_1 буде входом для підграфа g_2 , тоді справедливим буде:

$$o_i = g_i(o_{i-1}), i = \overline{2, k}.$$

Виходячи з цього, сформуємо впорядковану множину входів I , де:

$$I = (x, o_1, g_i(o_{i-1}) : i = \overline{2, k}).$$

Аналогічно до прикладу сформуємо впорядковану множину виходів, використовуючи елементи кортежів G та I :

$$O = (G_j(I_j) : j = \overline{1, k}).$$

Сформувавши впорядковані множини підграфів, їх входів і виходів перейдемо до квантування. Насправді операція квантування Q буде ідентичною, як у випадку цілого графа G_f , відпрацьовуючи на кожному з підграфів множини G .

Відповідно, набори констант квантування шукаємо так:

$$S_i, Z_i = Q(G_i(I_i)), i = \overline{1, k}.$$

Варто зазначити, що значення k має бути мінімально необхідним для розділеного розміщення нейромережі у пам'яті пришвидшувачів. Адже потрібно враховувати, що загальна тривалість виконання алгоритму, окрім нейромережових обчислень, також міститиме витрати на послідовне переведення результатів з одних пришвидшувачів на інші. Однак, завдяки сучасним фізичним протоколам, як наприклад, PCIe, згадані витрати є неістотними.

Таблиця. Результати OD моделей до квантування*, після квантування** / Results of OD models before quantization*, after quantization **

Модель	mAP@0.5*	mAP@0.5**
SSD MobileNetV1 [1]	20,1	18,8
SSD MobileNet V2 [19]	22,0	21,5
EfficientDet-Lite0 [23]	30,1	27,3
SSDLite MobileDet [27]	31,8	30,2

У таблиці показано результати роботи популярних нейромережових алгоритмів розпізнавання об'єктів. Як метрику для оцінювання продуктивності моделі було використано mAP з пороговим значенням 0.5. Mean Average Precision – визначає середню точність об'єктів, знайдених моделлю, враховуючи всі категорії об'єктів у наборі даних.

Обговорення результатів дослідження. У роботі [25] досліджено використання БПЛА з алгоритмами комп'ютерного зору для автономної навігації та виявлення об'єктів. Розглянуто застосування БПЛА в різних галузях, таких як військові операції та сільськогосподарське управління. У дослідженні акцентовано увагу на розвитку технологій глибокого навчання, для вирішення традиційних завдань комп'ютерного зору. Зокрема, аналізується продуктивність сучасних алгоритмів OD, що опрацьовуються на наземних станціях. Автори роботи [5] запропонували нову ефективну структуру для інтеграції алгоритмів комп'ютерного зору в наземні станції БПЛА. Для демонстрації можливостей було реалізовано та оцінено OD моделі глибокого навчання.

Дослідники [15] вважають, що економічні системи комп'ютерного зору сприяють автономії БПЛА. У цій роботі представлено бюджетний прототип для навігації БПЛА, який мінімізує вплив на тривалість польоту. Прототип, що використовує систему посадки та уникнення перешкод (LOAS), базується на відкритому ПЗ та комунікації MAVLink, і протестований на квадрокоптері з Pixhawk і Raspberry Pi.

Робота [14] пропонує підхід для покращення автономних можливостей БПЛА шляхом використання OD алгоритмів на підставі YOLO. Автори продемонстрували здатність системи точно розпізнавати різні посадкові точки в нових і невідомих середовищах. Запропонований підхід значно знижує час передачі та оброблення даних до наземних станцій та забезпечує швидке

розпізнавання об'єктів за рахунок використання NVIDIA Jetson Xavier NX. У дослідженні [22] запропоновано навігаційну систему безпечної посадки БПЛА, яка використовує OD алгоритми комп'ютерного зору, здатні працювати на обчислювальних ресурсах NVIDIA Jetson Xavier, наявних на борту типового БПЛА. У праці [3] запропоновано структурну інновацію на підставі моделі YOLOv5-MobileNetv3Small для вирішення проблеми оброблення зображень високої роздільної здатності отриманих з БПЛА. Зокрема, дослідники адаптували топологію мережі MobileNetv3 [9]. Покращений алгоритм використовує датасет Visdrone2019-DET [4] для тренування моделі та пришвидшувач Jetson NX для оброблення відеопотоку на платформі БПЛА.

У публікації [7] досліджено необхідність виявлення БПЛА в реальному часі для моніторингу та запобігання можливим загрозам. Використовуючи алгоритми YOLOv5 та YOLOv5 s автори досягнули швидкості детекції 265 кадрів за секунду. За допомогою квантизації моделі дослідники підвищили швидкість до 300 кадрів на секунду, що дало змогу здійснювати виявлення дронів в реальному часі. Дослідники [6] порівняли два методи масштабування згорткових нейронних мереж: рівномірне масштабування та NeuralScale, проаналізували їхній вплив на затримку роботи, використання пам'яті та енергоспоживання на моделях MobileNetV2 та ResNet-50 [8], масштабованих до різних розмірів. Тести проводили на процесорі ARM Cortex-A72, графічному процесорі NVIDIA Jetson AGX Xavier та FPGA Xilinx ZCU104. Висновки показали, що NeuralScale покращує точність, але збільшує обчислювальну складність, тоді як рівномірне масштабування є простішим у застосуванні та ефективніше для зниження затримки і використання пам'яті.

Порівняно з роботами [5, 25] запропонована фізична модель є стійкою до радіозавад, оскільки обчислення проводили безпосередньо на БПЛА без залучення наземної станції. Запропонована обчислювальна система більш продуктивна для нейромережових обчислень, ніж [15] через використання трьох TPU модулів і NPU. Порівняно з дослідженнями [6, 7], запропоноване квантування частинами дає змогу не тільки зменшити розмір моделі, а й розподілити обчислення на декілька апаратних пришвидшувачів.

Отже, за результатами виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – вдосконалено метод нейромережового розпізнавання об'єктів шляхом поділу зображення на фрагменти та їх одночасного оброблення з використанням вагових коефіцієнтів і вхідних даних у форматі з фіксованою комою, що забезпечило зменшення тривалості опрацювання, спростило вимоги до продуктивності та складності комп'ютерних засобів.

Практична значущість результатів дослідження – використання розроблених нейромережових засобів забезпечує пришвидшене розпізнавання об'єктів на комп'ютерних засобах з фіксованою комою.

Висновки / Conclusions

Розроблено структуру системи автономного управління безпілотним літальним апаратом з використанням засобів комп'ютерного зору, що дасть змогу виконувати

точну навігацію, уникати перешкоди, автономно виконувати місії та збирати дані в реальному часі. За результатами проведеного дослідження можна зробити такі основні висновки.

1. Проаналізовано апаратні засоби, вибрано підходи та принципи побудови, розроблено структуру системи автономного управління безпілотним літальним апаратом з використанням комп'ютерного зору, блок-схему алгоритму візуальної геолокації та вдосконалено метод нейромережового розпізнавання об'єктів.
2. Обрано підходи та сформульовано вимоги до побудови системи безпілотного літального апарату з використанням комп'ютерного зору, основними з яких є: забезпечення необхідної обчислювальної потужності, високої енергоефективності, автономної навігації, оброблення даних з сенсорів в реальному часі, використання автопілота для управління БПЛА.
3. Розроблено структуру системи автономного управління БПЛА, основними компонентами якої є: комп'ютер-компаньйон з TPU та камерою, контролер автопілота з GPS та IMU та радіомодуль зі з'єднанням до наземної станції, що забезпечує пришвидшення розпізнавання об'єктів і автономне управління БПЛА.
4. Обрано одноплатні комп'ютери Radxa Rock 5 b та 3 Google Edge TPU, які забезпечують пришвидшене розпізнавання об'єктів і оброблення даних від навігаційних сенсорів і відеокамери в режимі реального часу.
5. Вдосконалено метод штатного планування маршрутів системи автопілота ArduPilot шляхом використання алгоритму візуальної корекції курсу БПЛА, що дало змогу коригувати траєкторію польоту в реальному часі на підставі візуальних даних, отриманих з камери, та покращити точність навігації та уникнення перешкод, забезпечуючи більш ефективне виконання місій та зниження ризику аварійного польоту.
6. Удосконалено метод нейромережового розпізнавання об'єктів шляхом поділу зображення на фрагменти та їх одночасного оброблення з використанням вагових коефіцієнтів і вхідних даних у форматі з фіксованою комою, що забезпечило зменшення тривалості опрацювання, спростило вимоги до продуктивності та складності комп'ютерних засобів.

References

1. Andrew, G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, & Hartwig Adam. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv:1704.04861. <https://doi.org/10.48550/arXiv.1704.04861>
2. Brennan, J., Bonner, S., Atapour-Abarghouei, A., Jackson, P. T., Obara, B., & McGough, A. S. (2020). Not Half Bad: Exploring Half-Precision in Graph Convolutional Neural Networks. International Conference on Big Data (Big Data), 2725–2734. <https://doi.org/10.1109/BigData50022.2020.9378263>
3. Chen, Y., Chen, X., Chen, L., He, D., Zheng, J., Xu, C., & Liu, L. (2022). UAV Lightweight Object Detection Based on the Improved YOLO Algorithm. Proceedings of the 2021 5th International Conference on Electronic Information Technology and Computer Engineering, 1502–1506. Presented at the Xiamen, China. <https://doi.org/10.1145/3501409.3501674>
4. Du, D., Zhu, P., Wen, L., Bian, X., Lin, H., Hu, Q., & Liu, Z. (2019). VisDrone-DET2019: The Vision Meets Drone Object Detection in Image Challenge Results. 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 213–226. <https://doi.org/10.1109/ICCVW.2019.00030>
5. Enkhtogtokh Togotogtokh, Christian Micheloni, Gian Luca Foresti, & Niki Martinel. (2020). An Efficient UAV-based Artificial Intelligence Framework for Real-Time Visual Tasks. arXiv preprint arXiv:2004.06154. <https://doi.org/10.48550/arXiv.2004.06154>

6. Gealy, C. B., & George, A. D. (2024). Characterizing Parameter Scaling with Quantization for Deployment of CNNs on Real-Time Systems. *ACM Transactions on Embedded Computing Systems*, 23(3), 1–35. <https://doi.org/10.1145/3654799>
7. Guo, D., Zhang, Y., She, W., Liao, Y., Xu, H., Tang, J., & Geng, Y. (2024). Real-time recognition application for UAV based on YOLOv5. *Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering*, 1128–1133. Presented at the Dalian, China. <https://doi.org/10.1145/3652628.3652813>
8. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
9. Howard, A., Sandler, M., Chen, B., Wang, W., Chen, L.-C., Tan, M., & Le, Q. (2019). Searching for MobileNetV3. *CVF International Conference on Computer Vision (ICCV)*, 1314–1324. <https://doi.org/10.1109/ICCV.2019.00140>
10. Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., & Kalenichenko, D. (2018). Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference. *CVF Conference on Computer Vision and Pattern Recognition*, 2704–2713. <https://doi.org/10.1109/CVPR.2018.00286>
11. Koch, G., Zemel, R., & Salakhutdinov, R. (2015). Siamese neural networks for one-shot image recognition. *ICML deep learning workshop*, 2(1), 8 p. URL: <https://www.cs.cmu.edu/~rsalakhu/papers/oneshot1.pdf>
12. Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal Loss for Dense Object Detection. *International Conference on Computer Vision (ICCV)*, 2999–3007. <https://doi.org/10.1109/ICCV.2017.324>
13. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In: *Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, Proceedings, part I* 14, 21–37. Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2
14. Ma, M.-Y., Shen, S.-E., & Huang, Y.-C. (2023). Enhancing UAV Visual Landing Recognition with YOLOs Object Detection by Onboard Edge Computing. *Sensors*, 23(21). <https://doi.org/10.3390/s23218999>
15. Ortega, L. D., Loyaga, E. S., Cruz, P. J., Lema, H. P., Abad, J., & Valencia, E. A. (2023). Low-Cost Computer-Vision-Based Embedded Systems for UAVs. *Robotics*, 12(6), 32–39. <https://doi.org/10.3390/robotics12060145>
16. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. <https://doi.org/10.1109/CVPR.2016.91>
17. Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
18. Robicquet, A., Sadeghian, A., Alahi, A., & Savarese, S. (2016). Learning social etiquette: Human trajectory understanding in crowded scenes. In *Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands. Proceedings, Part VIII* 14, 549–565. Springer International Publishing. https://doi.org/10.1007/978-3-319-46484-8_33
19. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *CVF Conference on Computer Vision and Pattern Recognition*, 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
20. Seshadri, K., Akin, B., Laudon, J., Narayanaswami, R., & Yazdanbakhsh, A. (2022). An Evaluation of Edge TPU Accelerators for Convolutional Neural Networks. *International Symposium on Workload Characterization (IISWC)*, 79–91. <https://doi.org/10.1109/IISWC55918.2022.00017>
21. Süzen, A. A., Duman, B., & Şen, B. (2020). Benchmark Analysis of Jetson TX2, Jetson Nano and Raspberry PI using Deep-CNN. *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 1–5. <https://doi.org/10.1109/HORA49412.2020.9152915>
22. Symeonidis, C., Kakaletsis, E., Mademlis, I., Nikolaidis, N., Tefas, A., & Pitas, I. (2021). Vision-based UAV Safe Landing exploiting Lightweight Deep Neural Networks. *Proceedings of the 2021 4th International Conference on Image and Graphics Processing*, 13–19. Presented at the Sanya, China. <https://doi.org/10.1145/3447587.3447590>
23. Tan, M., Pang, R., & Le, Q. V. (2020). EfficientDet: Scalable and Efficient Object Detection. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10778–10787. <https://doi.org/10.1109/CVPR42600.2020.01079>
24. Tan, T., & Cao, G. (2020). FastVA: Deep Learning Video Analytics Through Edge Processing and NPU in Mobile. *INFOCOM 2020 – Conference on Computer Communications*, 1947–1956. <https://doi.org/10.1109/INFOCOM41043.2020.9155476>
25. Wang, X., Cheng, P., Liu, X., & Uzochukwu, B. (2018). Fast and Accurate, Convolutional Neural Network Based Approach for Object Detection from UAV. *IECON 2018–44th Annual Conference of the IEEE Industrial Electronics Society, DC, USA*, 3171–3175. <https://doi.org/10.1109/IECON.2018.8592805>
26. Wubben, J., Fabra, F., Calafate, C. T., Krzeszowski, T., Marquez-Barja, J. M., Cano, J.-C., & Manzoni, P. (2019). Accurate Landing of Unmanned Aerial Vehicles Using Ground Pattern Recognition. *Electronics*, 8(12), 114–119. <https://doi.org/10.3390/electronics8121532>
27. Xiong, Y., Liu, H., Gupta, S., Akin, B., Bender, G., Wang, Y., & Chen, B. (2021). MobileDets: Searching for Object Detection Architectures for Mobile Accelerators. In: *CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 3824–3833. <https://doi.org/10.1109/CVPR46437.2021.00382>

O. O. Borovyi, R. I. Iechko

Lviv Polytechnic National University, Lviv, Ukraine

STRUCTURE OF AUTONOMOUS CONTROL SYSTEM FOR UNMANNED AERIAL VEHICLE USING COMPUTER VISION TOOLS

This paper examines the features of constructing a physical model of an unmanned aerial vehicle (UAV) for computer vision tasks. The study analyzes existing physical models, highlighting their strengths and weaknesses. Consequently, it describes existing methods and tools for route planning and navigation. Moreover, the research outlines preliminary achievements in reducing the size of neural network models for computer vision tasks, specifically linear quantization of neural networks and half-precision computation. We identify approaches and principles for developing an autonomous control system for UAVs. As a result, we develop the structure of the UAVs autonomous control system and note the characteristics of its computer tools. For example, the concept for constructing a geodata correction algorithm uses images from a camera, satellite images, inertial measurement units, and a global positioning system (GPS). The study improves the method of neural network object recognition by parallelizing video data processing using multiple digital accelerators. The benefits of this study include proposed scenarios for using the system with numerous hardware accelerators. We describe the performance of models before and after quantization in the context of recognition accuracy. Furthermore, we present a method of partial quantization for processing ultra-large neural networks in computer vision. The proposed met-

hods and technologies demonstrate high potential for improving the efficiency and accuracy of data processing under real-world operating conditions. Improved approaches to neural network quantization and optimization significantly reduce model size and increase processing speed while maintaining high recognition accuracy. Consequently, the integration of multiple digital accelerators allows for effective parallelization of video data processing, which is particularly important for real-time information analysis. The results of this study can be used to improve navigation capabilities and video data processing, contributing to the expansion of UAV applications.

Keywords: visual geolocalisation; deep neural network; quantization; hardware accelerator.