

Я. М. Матвійчук¹
В. П. Яцишин¹
В. М. Тарабан¹

СИСТЕМА КОРЕКЦІЇ ІНЕРЦІЙНОЇ НАВІГАЦІЇ БПЛА НА ОСНОВІ ГЛИБОКОГО НАВЧАННЯ ДЛЯ ЗІСТАВЛЕННЯ ЗОБРАЖЕНЬ З КАМЕРИ ТА СУПУТНИКОВИХ ЗНІМКІВ

¹Національний університет «Львівська Політехніка»

Надійна навігація безпілотних літальних апаратів (БПЛА) є критичною функцією, що значною мірою покладається на сигнали GNSS. В умовах радіоелектронної боротьби (РЕБ) ці сигнали пригнічуються, що призводить до втрати апаратів через неконтрольоване накопичення похибки бортових інерційних навігаційних систем (INS). Наявні візуальні методи локалізації часто є обчислювально-витратними та неоптимальними для замських ландшафтів.

Метою роботи є розробка та експериментальна валідація обчислювально-ефективної гібридної системи автономної геолокалізації, здатної коригувати похибку INS шляхом зіставлення зображень з бортової камери та референсних супутникових знімків, з фокусом на життєздатність на низькоресурсних бортових платформах.

Запропоновано гібридний метод, що поєднує дані INS з дискретною корекцією від модуля глибокого навчання. Проведено збір та розмітку власного набору даних польотів у замській місцевості. Виконано порівняльний аналіз архітектур CNN (VGG11, MobileNetV2, ResNet18) та функцій втрат для задачі крос-модального зіставлення. Точність та продуктивність фінальної моделі валідовано на одноплатному комп'ютері Raspberry Pi.

Експериментально показано, що Triplet Loss, завдяки механізму відбору «негативних» прикладів, змушує модель вивчати диференційовані ознаки, що дозволяє розрізняти візуально схожі ділянки ландшафту. У порівняльному тестуванні архітектур, модель VGG11, навчена з Triplet Loss, показала найкращу точність, перевершивши MobileNetV2 та ResNet18. Під час тестування механізму корекції з симульованим початковим 10-метровим дрейфом INS, система на основі VGG11 досягла середньої похибки 5,4 метри, успішно обчислюючи вектор корекції. На реальному тестовому маршруті довжиною 200 м, де максимальна похибка некерованої INS досягла 15,2 м, запропонована система зменшила максимальне відхилення до 8,5 м. Валідація продуктивності на Raspberry Pi 4 підтвердила обчислювальну придатність: повний цикл корекції для зображення 224×224 становить 2,5 секунди, а для 448×448 — 10 секунд, що є прийнятним для періодичного (непостійного) уточнення координат.

Ключові слова: автономна навігація, БПЛА, GPS-denied, корекція INS, глибоке навчання, згортова нейронна мережа, зіставлення зображень, супутникове зображення, Triplet Loss, Raspberry Pi.

Вступ

БПЛА стали невід'ємним компонентом сучасних технологічних рішень у цивільній та військовій сферах. Їхня ефективність в задачах моніторингу, розвідки, коригування вогню та логістики безпосередньо залежить від здатності до точної навігації та геолокалізації. Основою для абсолютної більшості таких систем є GNSS, зокрема GPS, що забезпечують високу точність визначення абсолютних координат за мінімальних обчислювальних витрат з боку бортового обладнання.

Проте, висока залежність від GNSS створює критичну вразливість. В умовах активного протистояння, зокрема у разі застосування засобів РЕБ, супутникові сигнали можуть бути повністю пригнічені або спотворені (spoofing), що призводить до дезорієнтації та втрати апаратів. Статистичні дані з реальних бойових умов вказують, що до 40 % втрат БПЛА спричинені саме впливом РЕБ. Це зумовлює гостру потребу в розробці надійних методів автономної навігації, здатних функціонувати у GPS-заборонених (GPS-denied) середовищах. Традиційні безсупутникові підходи, такі як INS, хоча і є базовими, страждають від невідворотного накопичення похибки (drift) через

інтегрування показників акселерометрів та гіроскопів, через що вони непридатні для тривалих місій без періодичної корекції. Інші методи, як-от візуальна одометрія або LIDAR, мають суттєві обмеження: висока вартість, обчислювальна складність, або низька ефективність в умовах поганої видимості та на територіях з малою кількістю виразних ознак, наприклад, у полях чи лісових масивах.

Запропонований у роботі підхід базується на гібридній архітектурі, де інерційна навігаційна система контролера польоту (на базі Ardupilot) використовується для отримання апріорної, хоча й схильної до накопичення похибки, позиції. Ця позиція використовується для вибірки відповідної ділянки (патчу) з референсної супутникової карти. Надалі, CNN застосовується для екстракції карт ознак з пари зображень: поточного кадру з бортової камери та підготовленого супутникового патчу. Шляхом крос-кореляції цих карт ознак обчислюється вектор корекції (зміщення у пікселях), який нівелює похибку інерційної системи та уточнює позицію БПЛА. В ході дослідження проведено експериментальне тренування та порівняння трьох архітектур CNN (VGG11, ResNet18, MobileNetV2) на власноруч зібраному наборі даних, що імітує замську місцевість. Окрему увагу приділено підтвердженню ефективності функцій втрат Triplet Loss для задачі навчання зіставленню крос-модальних зображень (візуальний спектр проти супутникового). Фінальну валідацію системи проведено на одноплатному комп'ютері Raspberry Pi 4 для підтвердження її обчислювальної ефективності.

Метою роботи є розробка та експериментальна валідація гібридної системи автономної геолокалізації БПЛА з подальшим уточнення позиції через зіставлення фото з камери та супутника.

Об'єктом дослідження є процес визначення координат БПЛА в умовах відсутності сигналу GNSS.

Предметом дослідження є методи глибокого навчання та алгоритми комп'ютерного зору для зіставлення крос-модальних зображень (візуальних з камери та супутникових) з метою корекції інерційної навігації.

Огляд літературних джерел

Проблема автономної навігації БПЛА в умовах відсутності сигналу GNSS є предметом активних наукових досліджень. Аналіз сучасних публікацій демонструє стійку тенденцію до застосування методів глибокого навчання для розв'язання задачі візуальної геолокалізації шляхом зіставлення зображень з бортової камери та референсних супутникових знімків. Найявні підходи можна умовно класифікувати за архітектурою, цільовою задачею та обчислювальними вимогами.

Значна частина досліджень зосереджена на створенні гібридних систем, що поєднують інкрементальні методи (візуальна одометрія або інерційна навігація) з абсолютною корекцією на основі візуального зіставлення. Такий підхід запропоновано у роботі [1], де автори комбінують візуальну одометрію з модулем корекції на основі VGG16 та алгоритму ICLK. Концептуальна цінність цієї роботи полягає у підтвердженні життєздатності архітектури, де система не покладається виключно на візуальне зіставлення, а використовує його для періодичного нівелювання накопиченої похибки візуальної одометрії. Проте, автори зауважують на критичний недолік: з переходом від міських ландшафтів (середня похибка 7 м) до замських, польових територій, точність системи стрімко деградує (похибка до 25 м). Це демонструє нездатність стандартних, навчених на міських датасетах, моделей ефективно працювати з однорідними природними ландшафтами, що є ключовим викликом для нашого дослідження.

Інша група підходів фокусується на методах виділення та зіставлення ключових ознак. У [2] пропонується комплексний підхід, що використовує моделі SuperPoint для детекції ознак та SuperGlue для їх зіставлення. Цінним у цій роботі є введення механізму оцінки впевненості, що дозволяє системі ігнорувати кадри, які можуть призвести до значної похибки, досягаючи середньої точності 9,5...13 метрів. Проте такий підхід, що базується на дескрипторах, відрізняється від вибраного нами методу крос-кореляції карт ознак і може мати вищі обчислювальні вимоги на етапі зіставлення.

Альтернативний метод, схожий за кінцевим результатом на наш, запропонований у [3]. Автори використовують архітектуру SSPT, яка на виході генерує теплову карту (heatmap), що вказує на найімовірнішу позицію БПЛА в межах ширшого супутникового знімка. Такий підхід демонструє високу точність (до 10,5 м) навіть за несприятливих умов. Проте, він покладається на архітектуру трансформера, обчислювальна складність якої може бути надмірною для цільових низькоресурсних платформ, таких як Raspberry Pi.

Окрему увагу в дослідженнях приділено проблемі доменного розриву (domain gap) між візуальними та супутниковими знімками, а також методам навчання ефективних представлень (feature extraction). У [4] автори не лише використовують Vision Transformer, але й впроваджують меха-

нізм трансферу колірної палітри для нормалізації зображень. Ключовим для нашої роботи є підтвердження високої ефективності функції втрат Triplet Loss для задачі крос-візуального пошуку (image retrieval), що обґрунтовує вибір цієї метрики замість Cosine Similarity. В роботі [5] пропонується оригінальний підхід з використанням автоенкодера, навченого виключно на супутникових знімках, для генерації компактних ознак. Цінною ідеєю з цієї роботи є передобчислення векторів ознак для всієї референсної карти перед початком місії, що значно знижує навантаження в реальному часі. Разом з тим, дане дослідження покладалося на високопродуктивне та дороге обладнання (NVIDIA Tegra TX2), що суперечить меті нашої роботи зі створення доступного рішення.

Важливим аспектом є практична реалізація подібних систем на бортових комп'ютерах з обмеженими ресурсами. У цьому контексті, робота [6] надає важливе обґрунтування. Хоча її мета полягала у детекції GPS-спуфінгу, а не в навігації, автори успішно продемонстрували можливість запуску згорткових нейронних мереж (ResNet, SqueezeNet) на платформі Raspberry Pi 3+ для розв'язання схожої задачі зіставлення зображень. Це підтверджує принципову можливість реалізації вибраної нами архітектури на низькоресурсних SoC.

Отже, аналіз літератури показує, що наявні рішення або демонструють високу точність на дорогому обладнанні, або втрачають ефективність поза межами міських ландшафтів. Хоча окремі роботи підтверджують корисність Triplet Loss та життєздатність CNN на платформах Raspberry Pi, бракує комплексного дослідження, яке б поєднало гібридний підхід (INS + корекція) з оптимізованою легкою CNN, навченою спеціально для замських територій. Ця робота спрямована на заповнення цієї прогалини шляхом розробки та валідації саме такої ресурсоефективної системи.

Основний матеріал

В основі цього дослідження лежить розробка гібридної системи автономної геолокалізації, що поєднує переваги інкрементальної INS та абсолютної візуальної корекції. На відміну від підходів, що намагаються повністю замінити GNSS візуальними методами, запропонована архітектура використовує глибоке навчання не для безперервного визначення координат, а для періодичного обчислення та застосування вектора корекції до даних, отриманих від бортових інерційних сенсорів. Такий підхід дозволяє суттєво знизити обчислювальне навантаження на бортовий комп'ютер, оскільки ресурсоемісні операції зіставлення зображень виконуються дискретно, а не в реальному часі з високою частотою.

Принцип роботи системи, зображений на блок-схемі (рис. 1), полягає у такому. Система ініціалізується із завантаженою референсною супутниковою картою всієї операційної зони та її географічними координатами. Під час польоту система безперервно отримує дані про приблизну позицію (широта, довгота, висота) та орієнтацію (кут повороту відносно півночі) від політного контролера БПЛА через протокол MAVLink. Ці дані, отримані від INS, невідворотно накопичують похибку. З заданим інтервалом (наприклад, кожні 10 секунд) активується механізм корекції. Він використовує поточну (неточну) позицію для вибірки відповідної ділянки з супутникової карти, отримує поточний кадр з бортової камери, спрямованої перпендикулярно до землі, та передає обидва зображення до модуля зіставлення на основі CNN. Модуль обчислює точне зміщення між двома зображеннями, що відповідає вектору похибки INS. Цей вектор корекції застосовується для уточнення поточної позиції БПЛА, після чого оновлені, скореговані координати можуть бути передані назад до політного контролера для подальшої навігації.

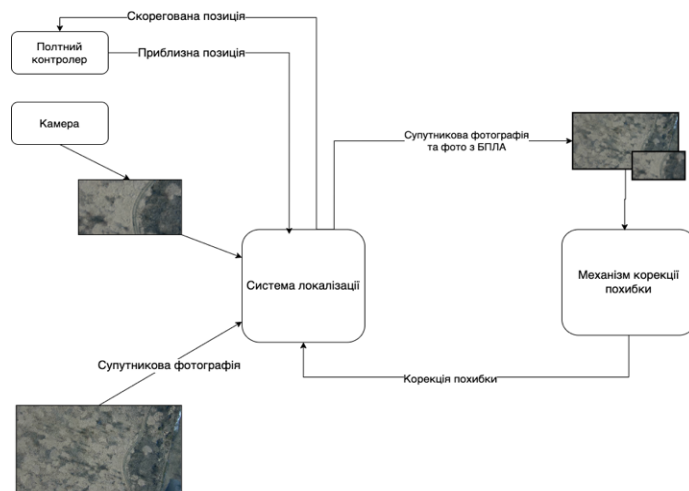


Рис. 1. Блок-схема принципу роботи запропонованої системи

Ключовим елементом системи є блок корекції, який математично та алгоритмічно забезпечує коректне зіставлення зображень, отриманих з різних джерел (бортова камера та супутник) та в різних умовах (масштаб, орієнтація).

Першим етапом є переведення приблизної GPS-позиції, отриманої від INS, у піксельні коор-

динати на завантаженій супутниковій карті. Це досягається лінійною інтерполяцією на основі відомих координат кутів.

$$X_p = W_{px} \cdot \frac{lonBR - lonTL}{lon - lonTL}; \quad Y_p = H_{px} \cdot \frac{latTL - latBR}{latBR - lat}. \quad (1)$$

Далі необхідно визначити, яка ділянка супутникової карти відповідає поточному полю зору камери. Фізичні розміри видимої області на землі обчислюються на основі поточної висоти польоту h та кутів огляду камери (θ_w, θ_h)

$$w_m = 2h \cdot \tan\left(\frac{\theta_w}{2} \cdot \frac{\pi}{180}\right); \quad h_m = 2h \cdot \tan\left(\frac{\theta_h}{2} \cdot \frac{\pi}{180}\right). \quad (2)$$

Для переведення цих фізичних розмірів у піксельні (w_{px}, h_{px}) обчислюється коефіцієнт масштабування $meters_per_pixel$, що враховує географічну широту місцевості.

$$\Delta lat = |latTL - latBR|; \quad \Delta lon = |lonTL - lonBR|; \quad lat_m = 111320 \cdot \Delta lat. \quad (3)$$

$$lon_m = \frac{40075000 \cdot \Delta lon}{360} \cdot \cos\left(\frac{latTL + latBR}{2} \cdot \frac{\pi}{180}\right); \quad (4)$$

$$meters_per_pixel_x = \frac{lon_m}{W_{px}}; \quad meters_per_pixel_y = \frac{lat_m}{H_{px}}; \quad (5)$$

$$meters_per_pixel = \frac{meters_per_pixel_x + meters_per_pixel_y}{2}; \quad (6)$$

$$w_{px} = \frac{w_m}{meters_per_pixel}; \quad h_{px} = \frac{h_m}{meters_per_pixel}. \quad (7)$$

Для надійного зіставлення, система вирізає із супутникової карти не лише точну ділянку $w_{px} \cdot h_{px}$ а й розширену область I_{region} (8) з використанням коефіцієнта λ (напр., 1.5). Це створює «вікно пошуку», що компенсує похибку апіорної позиції INS.

$$I_{region} = I_s \left[\left(Y_p - \frac{h_{px} \cdot \lambda}{2} \right); \left(Y_p + \frac{h_{px} \cdot \lambda}{2} \right), \left(X_p - \frac{w_{px} \cdot \lambda}{2} \right); \left(X_p + \frac{w_{px} \cdot \lambda}{2} \right) \right]. \quad (8)$$

$$c_x = \frac{w_{px} \cdot \lambda}{2}; \quad c_y = \frac{h_{px} \cdot \lambda}{2}. \quad (9)$$

До цієї розширеної області застосовується афінне перетворення (обертання) за допомогою ротатійної матриці R , щоб вирівняти орієнтацію супутникового знімка з поточним кутом повороту БПЛА θ .

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta & (1 - \cos \theta) \cdot c_x + \sin \theta \cdot c_y \\ \sin \theta & -\sin \theta & (1 - \cos \theta) \cdot c_x + \sin \theta \cdot c_y \end{bmatrix}. \quad (10)$$

$$\begin{pmatrix} c_x \\ c_y \end{pmatrix} = R_\theta \times \begin{pmatrix} c_x \\ c_y \\ 1 \end{pmatrix}. \quad (11)$$

З центру цієї оберненої області I_r вирізається фінальний патч I_{patch} (13), який за масштабом та орієнтацією точно відповідає зображенню з камери.

$$I_r = R_\theta \times I_{region}. \quad (12)$$

$$I_{patch} = I_r \left[\left(c_y - \frac{H_{px}}{2} \right); \left(c_y + \frac{H_{px}}{2} \right), \left(c_x - \frac{W_{px}}{2} \right); \left(c_x + \frac{W_{px}}{2} \right) \right]. \quad (13)$$

На етапі зіставлення, зображення з БПЛА та підготовлений супутниковий патч Patch подаються на вхід попередньо навченої згорткової нейронної мережі, яка функціонує як екстрактор ознак.

$$\bar{F}_d = \frac{F_d(c, i, j)}{\sqrt{\sum_c F_d(c, i, j)^2}}; \bar{F}_s = \frac{F_s(c, i, j)}{\sqrt{\sum_c F_s(c, i, j)^2}}. \quad (14)$$

Мережа генерує дві матриці ознак: $F_d \in R^{c \cdot h_d \cdot w_d}$ (з БПЛА) та $F_s \in R^{c \cdot h_s \cdot w_s}$ (з супутника). Ці матриці проходять L2-нормалізацію (14) для стабілізації процесу зіставлення. Фінальна корекція позиції знаходиться шляхом обчислення згортки (або крос-кореляції) між двома картами ознак (15).

$$H(u, v) = \sum_{c=1}^C \sum_{i=1}^{h_d} \sum_{j=1}^{w_d} \bar{F}_d(c, i, j) \cdot \bar{F}_s(c, i+u, j+v). \quad (15)$$

Результатом є теплова карта $H(u, v)$, де пік яскравості вказує на найімовірніше зміщення (u, v) зображення БПЛА відносно супутникового патча. Це зміщення (u, v) і є обчисленим вектором похибки INS, який використовується для корекції.

Експерименти

Для валідації запропонованого підходу проведено серію експериментів, що охоплювали підготовку даних, тренування моделей та їх тестування в симульованих і реальних умовах.

Збір та підготовка набору даних. Критичною проблемою для цієї задачі є відсутність публічних наборів даних, які б відповідали цільовим умовам експлуатації — замиським, переважно польовим та лісовим ландшафтам з низькою кількістю виразних орієнтирів. Наявні датасети (наприклад, University-1652 [7]) орієнтовані на міські середовища. Тому вирішили формувати власний набір даних.

Для збору даних використано власноруч зібраний БПЛА на базі контролера польоту Ardupilot, оснащений додатковою камерою SkyEye, що знімала відео з роздільною здатністю 1080p перпендикулярно до землі. Польоти проводились на висотах від 40 до 100 метрів у замиській місцевості. Як референсні дані вибрано супутникові знімки з сервісу Google Earth Pro, незважаючи на їхні відомі недоліки, зокрема застарілість зображень для певної території.

Процес розмітки полягав у синхронізації відеопотоку з лог-файлами контролера польоту. За допомогою бібліотеки rtmavlink[8] з логів витягувались точні часові мітки, GPS-координати, висота та кут повороту. Ці дані дозволили для кожного кадру відео, знятого з БПЛА, згенерувати відповідний йому супутниковий патч за описаною вище математичною методологією.

Початковий набір даних дещо обмежений, тому для його розширення та підвищення робастності моделей застосовано аугментацію: дзеркальні перевертання по горизонталі та вертикалі, розмиття та зміна яскравості. Це дозволило штучно розширити набір даних до понад 4000 пар зображень.

Архітектури моделей та методологія тренування. Для експериментів вибрано три поширені архітектури CNN, що відповідають критеріям наявності переднавчених ваг та відносно низької обчислювальної складності: VGG11/16 [9], ResNet18/34 [10] та MobileNetV2 [11]. У всіх моделях останні повноз'язні (класифікаційні) шари видалено, щоб використовувати згорткові шари як екстрактори ознак.

Ключовим аспектом методології тренування було застосування функції втрат Triplet Loss за (16). Цей підхід вимагає трьох типів вхідних даних: «Якір» (Anchor) – зображення з БПЛА, «Позитив» (Positive) — коректний супутниковий патч, та «Негатив» (Negative) — некоректний супутниковий патч [12]. Мета функції — зробити відстань у просторі ознак між Якорем та Позитивом меншою ніж відстань між Якорем та Негативом, на певну величину (margin α). Для підвищення ефективності навчання, негативні приклади вибиралися динамічно, використовуючи часовий інтервал між кадрами (наприклад, супутниковий патч, що відповідає позиції 10 секунд тому), що змушувало модель розрізняти дуже схожі, але просторово зміщені ділянки.

$$L(F_a, F_p, F_n) = \max\left(\left|F_a - F_p\right|_2^2 - \left|F_a - F_n\right|_2^2 + a, 0\right). \quad (16)$$

Обговорення результатів

Тренування моделей. Моделі (VGG11, ResNet18, MobileNetV2) демонстрували стабільні та реалістичні криві навчання протягом 60 епох (рис. 2).

Візуальний аналіз згенерованих карт ознак показав, що моделі навчилися виділяти та зіставляти ключові просторові характеристики ландшафту, такі як дороги, межі полів та окремі групи дерев. (рис. 3).

За результатами цього етапу, моделі VGG11 та MobileNetV2 продемонстрували найкращий баланс між якістю ознак та складністю, і вибрані для подальших тестів.

Валідація механізму корекції. Для кількісної оцінки точності проведено симуляцію роботи механізму корекції. Приклади пари фотографій (супутник та з дрона), що використовувались у симуляції показані на рис. 4. На тестовому наборі даних до реальної позиції кожного зображення штучно додавалася випадкова похибка (симуляція дрейфу INS) величиною 10 та 15 метрів. Задача системи — знайти коректну позицію у створеному вікні пошуку. На рис. 5 подано приклад матриці ключових ознак для пари фотографій (з дрона та супутника). На рис. 6 подано теплову мапу, на якій позначено найкращий збіг, що і є результатом визначення місця розташування дрона. Результати для моделі VGG11 (найкращої) показали, що за початкової похибки в 10 метрів, середня похибка системи становила 5,4 метра (мінімальна — 1,0 м, максимальна — 16,3 м). За початкової похибки 15 метрів, середня похибка системи склала 10,8 метра. Це підтвердило, що система здатна ефективно знаходити корекцію та суттєво зменшувати похибку локалізації.

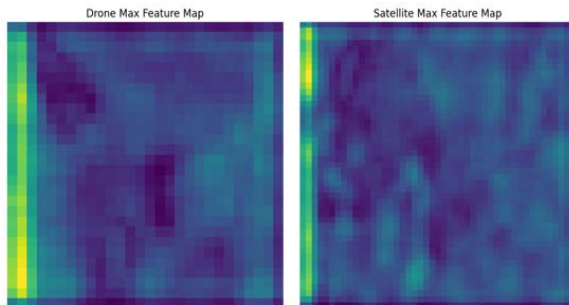


Рис. 5. Матриці ключових ознак для пари фото (зліва — БПЛА, справа — супутникове)

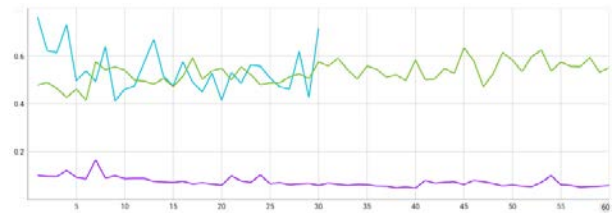


Рис. 2. Графіки функції похибки на валідаційному наборі даних для моделей ResNet18 (блакитна крива), VGG11 (фіолетова крива) та MobileNetV2 (зелена крива).

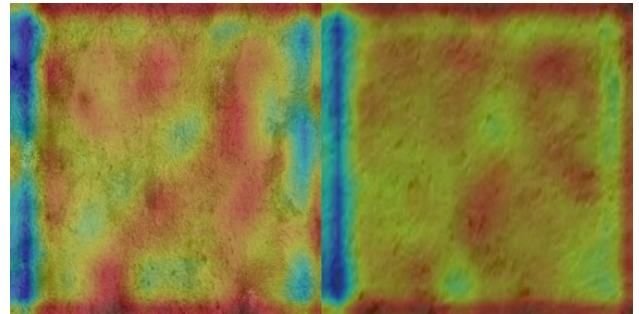


Рис. 3. Матриця ознак для пари фото для VGG11 (зліва — БПЛА, справа — супутник), приклад зіставлення фото дерев



Рис. 4. Приклад пари фото (зліва — супутникове, справа — БПЛА)

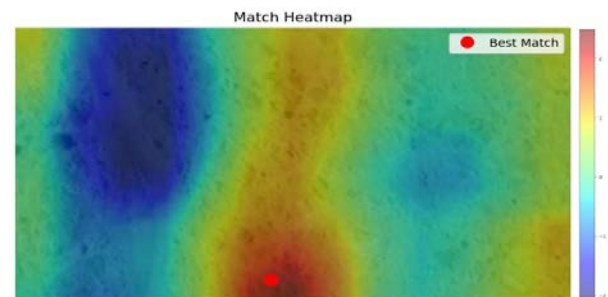


Рис. 6. Теплова карта отримана в результаті згортки двох матриць

Першим експериментом став пошук правильної локації з доданою похибкою 10 метрів у випадковому напрямі. Результати подані у табл. 2.

Таблиця 2

Результати роботи механізму корекції похибки (початкова похибка 10 метрів)

Модель	Похибка, м		
	середня	мінімальна	максимальна
VGG11	5,4	1	16,3
MobileNetV2	9,9	6,4	14,3

Також проведено експеримент зі збільшеною доданою похибкою в 15 метрів відносно правиль-

ної. Результати подані в табл. 3.

Таблиця 3

Результати роботи механізму корекції похибки (початкова похибка 15 метрів)

Модель	Похибка, м		
	середня	мінімальна	максимальна
VGG11	10,8	1,90	33,9
MobileNetV2	16,7	1,76	29,3

Проаналізувавши результати (табл. 2, 3), можемо побачити, що створена система демонструє дуже низькі середні похибки у 5,4 метри за початкової похибки в 10 метрів з використанням моделі VGG11. Модель MobileNetV2 продемонструвала дещо гірші результати, 9,9 метри за початкової похибки 10м. Такої точності недостатньо для коректної роботи в реальних умовах. Для подальших експериментів будемо використовувати модель VGG11.

Тестування на реальному маршруті. Наступним кроком провели валідацію системи в динаміці на реальному тестовому маршруті довжиною 200 метрів. Система отримувала приблизні координати від INS БПЛА та виконувала корекцію кожні 10 секунд. Результати порівнювалися з еталонним GNSS-треком (рис. 7).

З графіка (рис. 7) випливає, що приблизний маршрут, обчислений БПЛА за даними інерційних сенсорів, швидко накопичував похибку, максимальне відхилення якої сягнуло 15,2 метра від реального треку. Скорегований маршрут, отриманий за допомогою запропонованої системи, залишався значно ближчим до еталону, а його максимальне відхилення не перевищило 8,5 метра. Це доводить ефективність гібридного підходу в реалістичному сценарії.

Оцінка продуктивності на Raspberry Pi 4. Ключова вимога до системи — це її здатність працювати на низькоресурсних бортових комп'ютерах. Для цього проведено тестування продуктивності моделі VGG11 на одноплатному комп'ютері Raspberry Pi 4. Вимірювався час, необхідний для виконання одного повного циклу корекції (обробка двох зображень та обчислення згортки). Отримані результати подані у табл. 4. Використовуючи зображення розміром 448×448 пікселів, середня похибка становила 5,4 м, але час виконання був високим — 10 секунд. Зі зменшенням розміру зображень до 224×224 пікселів, час виконання скоротився у 4 рази — до 2,5 секунди, тоді як середня похибка незначно зросла до 5,9 м.

Отриманий час 2,5 секунди є абсолютно прийнятним для системи, що виконує періодичну корекцію (наприклад, раз на 10...30 секунд), підтверджуючи життєздатність рішення на доступних обчислювальних платформах.

Таблиця 4

Результати та час виконання механізму корекції похибки на Raspberry Pi 4

Розширення зображення	Похибка, м			Час виконання
	Середня	Мінімальна	Максимальна	
224×224	5,9	2	16,8	2.5 секунди
448×448	5,4	1	16	10 секунд

Висновки

У роботі розроблено, реалізовано та експериментально валідовано гібридну систему автономної геолокалізації БПЛА, призначену для роботи в умовах відсутності або придушення сигналів GNSS. Запропонована архітектура поєднує інкрементальні дані від бортових інерційних сенсорів (INS) з дискретною абсолютною корекцією, що обчислюється модулем глибокого навчання шля-

хом зіставлення зображень з камери БПЛА та референсних супутникових знімків.

Для тренування та валідації успішно сформовано та використано власний набір даних, що відображає цільові умови експлуатації, що є критичним для робастності системи.

Кількісна оцінка продемонструвала високу точність системи. Під час симуляції похибки INS у 10 метрів, запропонований механізм на основі VGG11 здатний відновити позицію з середньою похибкою 5,4 метра. Валідація на реальному польотному маршруті показала, що система здатна ефективно компенсувати дрейф INS, зменшивши максимальне відхилення від еталонного треку з 15,2 метрів до 8,5 метрів.

Критично важливим результатом є підтвердження обчислювальної ефективності. Доведено, що система здатна виконувати повний цикл корекції на низькоресурсному одноплатному комп'ютері Raspberry Pi 4 за 2,5 секунди (з роздільною здатністю 224×224). Це підтверджує практичну життєздатність запропонованого гібридного підходу, оскільки дозволяє створювати доступні, енергоефективні та стійкі до впливу РЕБ системи навігації для БПЛА.

Перспективи подальших досліджень включають дослідження методів оптимізації та квантизації згорткових нейронних мереж для подальшого зниження часу висновку на бортових пристроях. Перспективним напрямком є розробка мультимодальних архітектур, що інтегрують не лише зображення з камери, але й дані візуальної одометрії для підвищення стійкості до швидких змін орієнтації. Також важливим є дослідження робастності підходу за використання інших спектральних діапазонів, зокрема інфрачервоних камер для нічної навігації, та аналіз впливу якості і сучасності референсних супутникових знімків на точність корекції.

Декларація щодо генеративного ШІ

Під час підготовки цієї роботи автори використовували Gemini 3.0 Pro для виявлення та виправлення граматичних, орфографічних та стилістичних помилок. Після застосування цих інструментів автори провели ретельний перегляд і внесли необхідні зміни, беручи на себе повну відповідальність за остаточний зміст цієї публікації.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] H. Goforth, and S. Lucey, "GPS-Denied UAV Localization using Pre-existing Satellite Imagery," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 2974-2980. <https://doi.org/10.1109/ICRA.2019.8793558>.
- [2] X. Qiu, D. Yang, S. Liao, S. Wang, and Y. Li, "Image moment extraction based aerial photo selection for UAV high-precision geolocation without GPS," *Measurement*, no. 226, pp. 114141, 2024. <https://doi.org/10.1016/j.measurement.2024.114141>.
- [3] J. Fan, et al., "A Cross-View Geo-Localization Algorithm Using UAV Image and Satellite Image," *Sensors*, no. 24, issue 12, Чеп 2024, <https://doi.org/10.3390/s24123719>.
- [4] Z. Cui, et al., "A Novel Geo-Localization Method for UAV and Satellite Images Using Cross-View Consistent Attention," *Remote Sens.*, no.15, issue 19, Беп 2023, <https://doi.org/10.3390/rs15194667>.
- [5] M. Bianchi, and T. D. Barfoot, "UAV Localization Using Autoencoded Satellite Images," *IEEE Robot. Autom. Lett.*, no. 6, issue 2, pp. 1761-1768, 2021, <https://doi.org/10.1109/LRA.2021.3060397>.
- [6] N. Xue, L. Niu, X. Hong, Z. Li, L. Hoffaeller, and C. Pöpper, "DeepSIM: GPS Spoofing Detection on UAVs using Satellite Imagery Matching," in *Proceedings of the 36th Annual Computer Security Applications Conference*, в ACSAC '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 304-319. <https://doi.org/10.1145/3427228.3427254>.
- [7] Z. Zheng, Y. Wei, and Y. Yang, "University-1652: A Multi-view Multi-source Benchmark for Drone-based Geolocalization," in *Proceedings of the 28th ACM International Conference on Multimedia*, в MM '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1395-1403. <https://doi.org/10.1145/3394171.3413896>.
- [8] *ArduPilot/pymavlink*. (17, 01.2026). Python. ArduPilot. Дата звернення: 18, 2026. [Electronic resource]. Available: <https://github.com/ArduPilot/pymavlink>.
- [9] K. Simonyan, and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," 10.04, 2015, *arXiv: arXiv:1409.1556*. <https://doi.org/10.48550/arXiv.1409.1556>.
- [10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770-778. Accessed:18, 01. 2026. [Electronic resource]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html.
- [11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510-4520. [Electronic resource]. Available: https://openaccess.thecvf.com/content_cvpr_2018/html/Sandler_MobileNetV2_Inverted_Residuals_CVPR_2018_paper.html Accessed: 18. 01 2026.
- [12] Triplet loss, *Wikipedia*. 07, 09, 2025. [Electronic resource]. Available: https://en.wikipedia.org/w/index.php?title=Triplet_loss&oldid=1310086175. Accessed: 18. 01 2026.

Рекомендована кафедрою інфокомунікаційних систем і технологій ВНТУ

Дата надходження 28.01.2026

Дата прийняття до друку після рецензування 18.03.2026
Дата публікації 6.07.2026

Ця робота ліцензується відповідно до
[Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

Матвійчук Ярослав Миколайович — д-р техн. наук, професор, професор кафедри систем штучного інтелекту, e-mail: yaroslav.m.matviychuk@lpnu.ua . <https://orcid.org/0000-0002-5570-182X> ;

Яцишин Володимир Петрович — аспірант, асистент кафедри систем штучного інтелекту, e-mail: volodymyr.p.yatsyshyn@lpnu.ua . <https://orcid.org/0009-0001-9727-5080> ;

Тарабан Владислав Михайлович — студент Інститут комп'ютерних наук та інформаційних технологій, e-mail: vladyslav.taraban.mknssh.2025@lpnu.ua . <https://orcid.org/0009-0006-8925-8113>.

Національний університет «Львівська політехніка», Львів

Ya. M. Matviychuk¹
V. P. Iatsyshyn¹
V. M. Taraban¹

UAV Inertial Navigation Correction System Based on Deep Learning for Matching Camera Images and Satellite Imagery

¹Lviv Polytechnic National University

Reliable navigation of unmanned aerial vehicles (UAVs) is a critical function that relies heavily on GNSS signals. In electronic warfare (EW) environments, these signals are suppressed, leading to the loss of vehicles due to the uncontrolled accumulation of errors in onboard inertial navigation systems (INS). Existing visual localization methods are often computationally expensive and suboptimal for rural landscapes.

The aim of this work is the development and experimental validation of a computationally efficient hybrid autonomous geolocation system capable of correcting INS errors by matching images from an onboard camera with reference satellite imagery, with a focus on viability for low-resource onboard platforms.

A hybrid method combining INS data with discrete correction from a deep learning module is proposed. A custom dataset of flights in rural areas was collected and labeled. A comparative analysis of CNN architectures (VGG11, MobileNetV2, ResNet18) and loss functions for the cross-modal matching task was performed. The accuracy and performance of the final model were validated on a Raspberry Pi single-board computer.

It was experimentally demonstrated that Triplet Loss, thanks to its mechanism of selecting "negative" examples, forces the model to learn discriminative features, enabling it to distinguish visually similar areas of the landscape. In comparative testing of architectures, the VGG11 model trained with Triplet Loss showed the best accuracy, outperforming MobileNetV2 and ResNet18. When testing the correction mechanism with a simulated initial 10-meter INS drift, the VGG11-based system achieved an average error of 5.4 meters, successfully calculating the correction vector. On a real test route of 200 m, where the maximum error of the unmanaged INS reached 15.2 m, the proposed system reduced the maximum deviation to 8.5 m. Performance validation on a Raspberry Pi 4 confirmed computational suitability: a full correction cycle for a 224x224 image takes 2.5 seconds, and 10 seconds for 448x448, which is acceptable for periodic (non-continuous) coordinate refinement.

Keywords: autonomous navigation, UAV, GPS-denied, INS correction, deep learning, convolutional neural network, image matching, satellite imagery, Triplet Loss, Raspberry Pi.

Matviychuk Yaroslav M. — Dr Sc. (Eng.), Professor, Professor of the Chair of AIS, e-mail: yaroslav.m.matviychuk@lpnu.ua . <https://orcid.org/0000-0002-5570-182X> ;

Iatsyshyn Volodymyr P. — Post-Graduate Student, Assistant of the Chair of AIS, e-mail: volodymyr.p.yatsyshyn@lpnu.ua . <https://orcid.org/0009-0001-9727-5080> ;

Taraban Vladyslav M. — Student of the Institute of Computer Science and Information Technology, e-mail: vladyslav.taraban.mknssh.2025@lpnu.ua . <https://orcid.org/0009-0006-8925-8113>