

Д. М. КРИЦЬКИЙ, Н. І. ШКУРЕНКО, О. В. ПОПОВ, О. О. КРАВЦОВА

Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут», Харків, Україна

РОЗРОБЛЕННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ СЕГМЕНТАЦІЇ ДАНИХ ПО ФОТО ТА ВІДЕО ІНФОРМАЦІЇ

Об'єктом дослідження є процес сприйняття окремих об'єктів на фото та відео для подальшого аналізу ситуації у місті. *Предметом* дослідження є обробка зображень з метою виділення, класифікації та подальшого використання отриманої інформації про об'єкти на фото. *Мета* – створення інформаційної технології для сегментації окремих сегментів на фото та відео, які були отримані за допомогою безпілотних літальних апаратів (зокрема дронів) у міських сценах. *Основні задачі*: проведення аналізу існуючих методів сегментації даних; реалізація алгоритму, який би виконував сегментацію даних за різними об'єктами місцевості; проведення тестування програмного забезпечення; формування результатів дослідження. Отримані такі *результати*: розглянуто вже існуючі моделі сегментації зображень на предмет обмежень та недоліків. На основі виявлених недоліків сформовано вимоги до розроблюваної системи. Під час проведення аналізу предметної галузі виявлено та описано основні проблеми та визначено задачу, що підлягає автоматизації. У ході проектування алгоритмів вирішення задач було визначено послідовність виконання кожного з процесів. Конструювання програмного забезпечення включало у себе опис інформаційного простору системи та інтерфейсу користувача. Інтерфейс користувача було обрано після аналізу основних видів користувацьких інтерфейсів. Для оцінки якості проведеної обробки даних було використано існуючі метрики – IoU, піксельна точність та матриця невідповідності. У ході тестування було визначено помилки, що виникають під час виконання сегментації даних та причини їх виникнення, після їх усунення був показаний коректний результат сегментації зображень. **Висновки.** Наукова новизна отриманих результатів полягає у тому, що хоча для аналізу категорій зображень, отриманих за допомогою дронів, до сих пір немає методу сегментації, який повністю задовольняє всі потреби, розроблене програмне забезпечення дозволяє завантажувати зображення та шаблони до них, як еталони для штучного інтелекту, проводити сегментацію та перевіряти правильність цієї сегментації візуально та з використанням метрики IoU, матриці невідповідності. Програмне забезпечення можливо вдосконалити, автоматизувавши процес створення еталонного шаблону, провести навчання моделі на більшій кількості даних та інтегрувати створене програмне забезпечення у більшу систему, яка працювала би з дроном напряду.

Ключові слова: сегментація; алгоритм; метрика; БПЛА; шаблон; матриця невідповідності; програмне забезпечення.

Вступ

В останні роки сегментація стала одним із головних напрямків досліджень у галузі комп'ютерного зору. Це основа сприйняття в багатьох галузях, таких як робототехніка та автономне водіння. Багато успішних проєктів базуються на методах глибокого навчання. Вони здатні точно обробляти дані та отримувати високоточні результати. Досягти такого рівня – навчити штучний інтелект, можливо за умови надання великої кількості навчальних зразків і щільно позначених зображень. Це особливо важливо для реальних програм, де потребується точність отриманих даних. Семантична сегментація, як частина розуміння сцени, передбачає призначення міток кожному пікселю зображення.

Рішенням цієї проблеми може бути використання загальнодоступних наборів даних семантичної сегментації, які зосереджуються лише на ознаках об'єктів. Кілька таких наборів сегментації вже використовуються для розпізнавання контурів у більш складних міських сценах. Це набори даних Cityscapes [1] і CamVid [2], де за допомогою фіксованої камери, встановленої на автомобілі за кермом, можна отримати зображення об'єктів збоку. Зокрема набори даних CamVid і Highway Driving [3] надають безперервно позначені кадри водіння, що використовуються для семантичної сегментації відео з оцінкою тимчасової узгодженості.

Існують також набори даних семантичного маркування для супутникових та аерофотознімків, де фіксуються види об'єктів з висоти пташиного

польоту. Але наразі можна відмітити лише декілька наборів даних, що можуть зафіксувати сцени міського типу з точки зору безпілотного літального апарату, які можна було б спостерігати як зверху, так і збоку об'єктів. Безперечно, це надає більше інформації для розпізнавання об'єктів.

Можливість автоматичного отримання інформації про об'єкти на зображеннях може бути корисною у галузях, де важлива насамперед швидкість аналізування фото та відеоматеріалів.

Сьогодні такий інструмент як сегментація є одним із лідерів серед найпотужніших і перспективних інструментів попередньої обробки для інтерпретації зображень, представлених у цифровому вигляді та отриманих різноманітними методами.

1. Проблематика та аналіз досліджень та публікацій

Основна сфера використання БПЛА – збір фото та відео інформації. Коли ж ці дані отримані, то вони потребують аналізування та класифікування об'єктів, що знаходяться на цих матеріалах. Аналіз досить часто зводиться до відбору окремих фрагментів за заданими характеристиками і ознаками, це і є сегментацією. Наразі є безліч методів, що автоматизують даний процес аналізу на простих зображеннях, але вони не дають таких самих вдалих та власне коректних результатів для складніших зображень, де на фото присутні міські вулиці або ж лісні пейзажі. З плином часу було необхідно, щоб з'явився якийсь алгоритм чи метод, який би вирішив цю проблему.

Класичні методи сегментації виконують пошук однорідних областей, тобто схожих ділянок, які б можна було виділити окремо. На сьогоднішній момент їх самостійне використання вважається неефективним. Все більш популярними стають комбіновані методи, частиною яких є використання вже визначених наборів даних – шаблонів, що включають знання про характеристики того чи іншого об'єкта, розуміють його форму, текстуру, розташування, площу тощо. Існуючі методи не завжди дають коректні результати. Це пояснюється фактором унікальності фото та відео за рахунок кута зйомки, висоти, якості освітлення тієї чи іншої сцени. Тож підтверджується потреба у більш продуманому, в якійсь мірі «гнучкому» алгоритмі.

Штучний інтелект доволі часто використовується системами, що сегментують зображення. Штучна нейронна мережа, тобто штучний інтелект – є спрощеною моделлю біологічної нейронної мережі, яка являє собою сукупність штучних нейронів, взаємодіючих між собою. Наразі нейронні мережі є

корисними у багатьох областях машинного навчання, тим самим вирішуючи завдання різної складності [4].

Кожне зображення має дефекти, у тій чи іншій мірі. У роботі «Алгоритмічно-програмний метод колоризації зображень» [5] проаналізовані види недоліків та дефектів на зображеннях, серед них розповсюдженими є: хроматична аберация – явище викликане дисперсією світла (розкладання променя світла на складові), що проходить через об'єктив, дисторсія – геометричне спотворення прямих ліній. Дисторсія спостерігається у зум-об'єктивів при крайніх значеннях [6]. Виньетування – це падіння яскравості від центру до країв зображення. Також існують інші проблеми: розмиття, поява кольорового відтінку, шумів тощо.

У статті «Інтелектуальна система фотореалістичного перенесення стилів між зображеннями» [7] були розглянуті деякі цікаві алгоритми, що можуть бути у нагоді, наприклад, семантична сегментація – задача встановлення до якого класу об'єктів належить кожний піксель зображення, це один з кроків розуміння сцени. Також тема семантичної сегментації піднята у роботі «Новий підхід до семантичної сегментації автоматичних виймань дорожньої мережі з зображень дистанційного зондування модифікованим UNet» [8]. У цій дослідницькій статті пропонується алгоритм на основі глибокого навчання для сегментації дорожніх мереж із зображень дистанційного зондування. Цей алгоритм семантичної сегментації був розроблений з використанням модифікованого UNet. Результати показують, що модифікований UNet може ефективно сегментувати дорожні мережі за зображеннями дистанційного зондування з однаковим фоном.

У роботі «Метод автоматичної кластеризації даних дистанційного зондування» [9] був розроблений метод, що дозволяє автоматично виділяти на зображенні області зі статистично однорідними властивостями класифікаційних ознак. Цей метод заснований на поданні багатомодального емпіричного розподілу спостережень у вигляді суміші ненормованих функцій Гауса. Уточнення списку класів виконується шляхом виключення свідомо нерозрізнених об'єктів на підставі оцінок ймовірностей помилок розпізнавання; процедура виключення класів зводиться до аналізу безлічі порогів прийняття рішення, визначальних меж між класами.

2. Аналіз існуючих алгоритмів та методів сегментації

Порогова сегментація зображення являє собою попиксельне порівняння значень яскравості із заданим пороговим значенням для зміни вхідних

даних. Методи порогової обробки не можуть змінювати значення порогу відповідно до отриманого зображення. В основному даний метод використовується, коли існують стабільні умови – це, наприклад, обробка світлин інтернет-магазинів, задля яких зазвичай використовують штучне освітлення у спеціальних студіях. Порогові методи можуть використовувати складні та прості методи сегментації, але ключовим моментом залишається вибір порогового значення. Недоліком таких методів є те, що бувають випадки, коли через складні умови майже неможливо точно визначити поріг. У таких ситуаціях правильне значення можна вибрати лише після детального аналізу вхідних даних. Порогові методи мають перевагу над іншими, оскільки вони швидкі, не трудомісткі [10] та дають задовільні результати для простих зображень.

Методи розбиття зображення на області групують пікселі, що належать до певного об'єкта та виконують їх класифікацію по створеним областям. Визначено такі методи серед алгоритмів розбиття: центроїдне зв'язування, виділення меж та контурів, злиття-розщеплювання, нарощування. Алгоритм нарощування групує сусідні елементи за приблизними рівнями яскравості. Якщо ж значення мають наближений характер, то вони об'єднуються у однорідні області. Нарощування є корисним та ефективним, коли існують стійкі зв'язки всередині контурів окремих об'єктів.

Методи виділення меж являють собою процес визначення точок, що лежать на межах областей. Ці області можна відзначити за умови різкої зміни яскравості на зображенні.

Центроїдне зв'язування – це генерація точок, позначених як такі, що належать до множини, де кількість точок визначається інформацією про об'єкт. Якщо вхідне зображення складне, ці мітки можна визначити ітеративно.

Сегментація зображень за допомогою злиття-розщеплювання розділяє вхідну інформацію на окремі області. Однорідність кожної ділянки постійно перевіряється алгоритмом, він порівнює сусідні ділянки на можливість злиття. Якщо ділянка неоднорідна, то вона поділяється на чотири області. Якщо сусідні області схожі, вони об'єднуються.

Метод виділення контурів має можливість використовувати алгоритм, що порівнює сегмент зображення з еталонами об'єктів, а також диференційно-градієнтний метод. Ці алгоритми виконують аналіз коливань градієнтів яскравості, тим самим знаходячи контури цих об'єктів. Методи виділення контурів несуть менші витрати пам'яті, отримують правильні та інтуїтивно зрозумілі результати. Недоліком цього методу можна назвати велику обчислювану складність, що унеможливує

використання цього алгоритму у системах реального часу.

Крім перерахованих методів, можна також відзначити семантичну сегментацію, що в свою чергу являє собою поділ зображення на окремі ділянки інформації, що належать до того чи іншого класу об'єктів. Попередні методи створюють угруповання пікселів за загальними ознаками окремих об'єктів. Тому несемантичні методи не можуть бути точно визначеними, оскільки на зображенні можна знайти декілька сегментів. Екземпляри одного й того самого об'єкта повинні бути визначеними за допомогою методів виділення об'єктів. До недоліку можна віднести те, що сусідні пікселі одного класу можуть належати до різних екземплярів об'єктів, а непов'язані області – одному екземпляру об'єкта. До таких ситуацій належить приклад, коли дерево, що розміщене перед автомобілем, візуально поділяє його на дві частини.

Семантична сегментація є частиною задач класифікації, саме тому класи, що навчають штучний інтелект, є основною складовою. Більшість алгоритмів працюють з фіксованими наборами класів. Також існують бінарні класи, що часто можуть бути використаними у випадках, де дозволена більш груба класифікація об'єктів на зображеннях.

Необроблені піксельні дані подаються на вхід типового алгоритму семантичної сегментації. Потім виконуються такі процеси, як масштабування та вилучення функцій. Для навчання можна використовувати інші необхідні методи нарощування даних (наприклад, обернення зображення). Кожному об'єкту присвоюється окрема область зображення. Після того, як вікно вибрано, воно класифікується. Для уточнення результатів семантичної сегментації можна використати морфологічні операції або складніші підходи, наприклад, марковські випадкові поля.

Алгоритми семантичної сегментації можуть приймати такі вхідні дані: відтінки сірого у порівнянні з кольоровими зображеннями, виключення або включення даних про глибину, одиночні та стереозображення, 2D та 3D. Велика кількість методів використовуються для вирішення задачі семантичної сегментації. Порівнюючи класичні методи сегментації з алгоритмами, що засновані на згорткових нейронних мережах з використанням різноманітних наборів обробки зображень до та після сегментації, перевага надається останнім. Такий висновок можна зробити через точніші результати.

Семантична сегментація та штучний інтелект дуже тісно пов'язані між собою. На сьогоднішній день сам штучний інтелект є однією з популярних

галузей науки та техніки. Це пояснюється тим, що при вирішенні більшості задач алгоритми невідомі заздалегідь. Тому і є системи, які власне займаються розробкою ефективних алгоритмів для вирішення поставлених задач. Отримані дані можуть бути оброблені, застосовані або ж вдосконалені самим штучним інтелектом. Але для отримання початкових даних, він потребує навчання.

Наразі є перелік методів, спрямованих на навчання цього штучного інтелекту. Нижче описаний один з них – навчання з учителем, воно є основним методом, що використовується для побудови системи, яка проводить сегментацію зображень.

Графічне відображення процесу навчання з учителем зображено на рисунку 1.

Вчителем у даному методі виступають знання про середовище. Штучний інтелект про них нічого не знає. Тож вчитель та мережа, що повинна навчатися, отримують вектор навчання з навколишнього середовища. На основі вбудованих знань вчитель може сформулювати бажаний відгук та його передати нейронній мережі, що навчається.

Штучний інтелект у свою чергу повинен виконати цей бажаний результат, що є оптимальними діями. На основі вектору навчання та сигналу помилки і відбувається налаштування параметрів мережі. Сигнал помилки являє собою різницю між бажаним результатом та поточною реакцією нейронної мережі. Щоб імітувати поведінку вчителя, параметри налаштовуються поетапно. Ця емуляція повинна бути оптимальною у певному статистичному сенсі. Тож у процесі навчання знання вчителя повністю передаються в мережу. Після завершення навчання зв'язок з вчителем можна розірвати та дати можливість нейромережі самостійно працювати із середовищем.

Припустимо, що навчальні приклади генеруються певною функцією. Маючи скінченну кількість навчальних прикладів, можна побудувати систему з нульовою помилкою навчання. Для цього необхідно взяти мережу з більшою кількістю ваг, ніж кількість прикладів. Для відтворення кожного прикладу, є рівняння для невідомих. У разі якщо кількість невідомих менша за кількість рівнянь, то така система є недостатньо визначеною. Вона допускає нескінченну кількість рішень. Це і є головною проблемою – не вистачає інформації, щоб вибрати єдине правильне рішення, функцію-вчителя. У результаті випадково вибрана функція буде робити неадекватні прогнози для нових прикладів, яких немає в навчальній вибірці. Замість того, щоб узагальнювати відомі приклади, мережа запам'ятовує їх. Цей ефект відомий як перенавчання.

Після отримання результатів сегментації необхідно визначити якість проведеної обробки. Для цього і був створений показник якості семантичної сегментації – метрика IoU (Intersection over Union) або індекс Жаккара.

Метрика IoU визначає якість сегментації за допомогою наступної формули:

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

де TP (true positives) – кількість саме правильно розпізнаних пікселів;

FP (false positives) – кількість пікселів, не відносяться до класу, але вони включені до нього;

FN (false negatives) – кількість пікселів, які, навпаки, належать до класу, але не були розпізнані такими.

Приклад використання метрики IoU при порівнянні фігур показаний на рисунку 2.

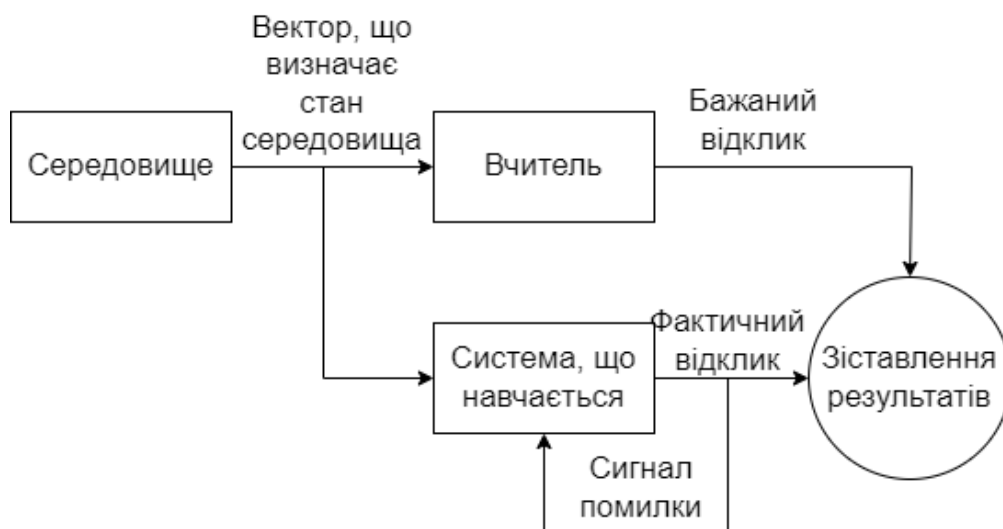


Рис. 1. Графічне відображення процесу навчання штучного інтелекту з учителем

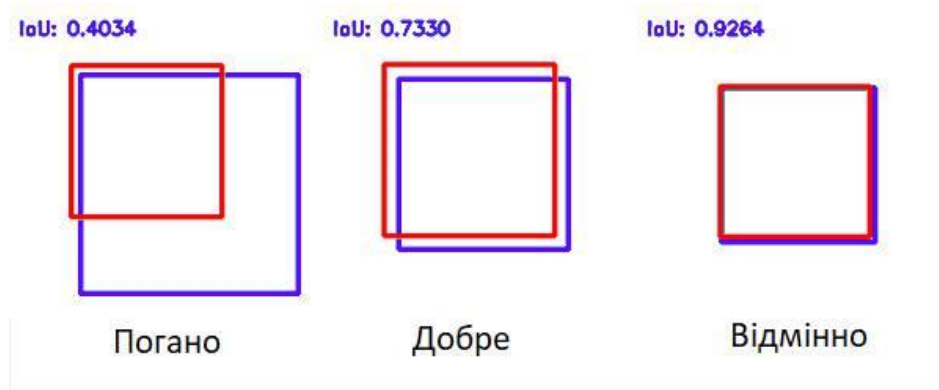


Рис. 2. Приклад використання метрики IoU

Модель FCN підходить для розв'язання задачі сегментації зображень, зроблених безпілотними літальними апаратами. Відрізняється гнучкістю зі сторони вхідних зображень та надає коректні результати. Інші методи не можуть надати таких переваг.

Лише правильно навчивши модель за певних умов, є можливість забезпечити її коректну поведінку. Створюючи навчальні набори даних для кожного конкретного випадку, можна отримати більше контролю над навчальною системою. Це значно спрощує процес розбиття інформації на частини в різних умовах.

Проаналізувавши існуючі методи сегментації даних, можна зробити висновок, що людський контроль необхідний навіть під час використання готових наборів даних. Це означає, що існує необхідність створення розділення на класи та задання їм відповідних ознак. Також слід мати на увазі, що обробка зображень може призвести до помилок, які можуть виникнути через недостатню освітленість, фоновий безлад та неможливість співвіднести шаблон з отриманими даними через розташування будь-якої перешкоди попереду об'єкта.

3. Етапи роботи програмного забезпечення

Програмне забезпечення для сегментації зображень виконується за допомогою мови програмування Python. Вона містить в собі пакети `сіру`, `pandas`, `scikit-learn`, `numpy`, які є корисними для аналізування даних. Також у якості IDE було обрано `PyCharm`. Додаток, що був розроблений, використовується за допомогою комп'ютера, який може зберігати і обробляти великі обсяги інформації, включаючи проміжні дані, та виконувати поставлені задачі з високою швидкістю. Умови використання застосунку дозволяють використовувати його в операційних системах Windows та Unix-подібних

через можливість у них виконання додатків, написаних на Python.

Головним завданням є автоматизація процесу сегментації фото та відео інформації. Збір даних здійснюється за допомогою БПЛА. У результаті обробки отриманої інформації необхідно ідентифікувати об'єкти, які присутні на фото. Далі ця інформація може застосовуватися у різних галузях.

Деталізуючи основну задачу вдалося виділити три основні етапи роботи програмного забезпечення: отримання даних, обробку інформації, виведення результату. В свою чергу деякі з них розділені на підзадачі.

Що стосується першого етапу: основна проблема полягає у тому, що на знімках високої роздільної здатності, зроблених за допомогою дронів під кутом, можлива достатньо сильна відмінність розмірів об'єктів на різних відстанях. Масштабні зміни зображення у свою чергу можуть вплинути на точність прогнозування. У мережі кожен вихідний піксель на кінцевому рівні прогнозування має фіксоване поле сприйняття, сформоване вихідними пікселями зображення, що може вплинути на кінцеве передбачення цього вихідного пікселя. Саме тому необхідно обробити шаблон, який міг би допомогти відрізнити об'єкти на зображенні.

На цьому етапі є дві підзадачі. «Конвертація в одноканальне зображення»: на вхід задачі подається шлях до папки, де знаходяться шаблони. Перед підготовкою даних перевіряється наявність каталогів. У випадку, коли вони відсутні, на екрані з'являється повідомлення, що дані не знайдені, і програма завершує роботу. В іншому випадку, програма продовжує своє виконання.

Після перевірки наявності даних створюються нові каталоги для збереження проміжних результатів. Необхідно забезпечити перевірку для створених директорій задля коректної роботи програми та адекватного виконання алгоритму. У тому разі, якщо вони не були створені, з'являється повідомлення про помилку та завершується виконання процесу.

Перетворення в одноканальне зображення досягається шляхом трансформації кольорових сегментів шаблону у чорний колір – колір безладу. Іншими словами, на зображення шаблону накладається маска. Після створення директорії, проводиться конвертація шаблонів в одноканальні зображення та відбувається збереження у створені раніше папки.

Підзадача «Зворотна конвертація» має такий порядок дій: на вхід подається шлях до даних, що були створені на попередньому етапі. Створюється директорія для збереження проміжних даних. Після цього відбувається перевірка чи були вони створені. За їх наявності процес продовжує своє виконання. Інакше, процес завершується.

Після створення директорії, починається обробка даних, що були передані. Одноканальні зображення перетворюються у триканальні, потім вони зберігаються у раніше створеному каталозі. Трансформація в триканальне зображення досягається шляхом ідентифікації кольорових сегментів на шаблоні, отриманому на попередньому кроці. Іншими словами, накладається маска з відповідними до класів об'єктів кольорів на зображення шаблону. Перелік класів та відповідних кольорів показані на рисунку 3. Дане розбиття важливе для визначення приналежності об'єкта до класу.

Другий етап – це етап обробки інформації, тут відбувається сегментація даних. Вхідними даними є фото, що було передано на етапі отримання даних, та шаблони, що були оброблені раніше. На початку виконання підзадачі створюються директорії для зберігання оброблених даних. Потім перевіряється, чи були вони створені. Якщо папки було створено, процес продовжується, якщо ні – завершується. Починається обробка даних, що було передано, далі відбувається збереження у раніше створені директорії. Зіставлення зображення з шаблоном виконується шляхом призначення ваги кожному з зображень та ваги, що буде застосовано до кожного пікселя зображення. Нижче приведена формула, що забезпечує змішування двох зображень:

$$\text{dst} = \text{src1} * \alpha + \text{src2} * \beta + \gamma, \quad (2)$$

де dst – результат виконання (зображення);

src1 – перше зображення для «змішування»;

α – вага першого зображення;

src2 – друге зображення;

β – вага другого зображення;

γ – вага кожного з пікселів.

Третій етап – отримання результатів. Для отримання результатів потрібно передати на вхід

початкові дані, отримані на першому етапі, та зображення після обробки.



Рис. 3. Розбиття на кольори

Аналіз виконаної сегментації починається зі створення матриці невідповідності. Це дає уявлення про ефективність створеного алгоритму. Кожен з рядків цієї матриці представляє зразки прогнозованого класу, тоді як кожен стовпець – зразки справжнього класу. Матриця невідповідності показана на рисунку 4.

		Справжній клас			
		А	Б	В	Г
Прогнозований клас	А	50	3	0	0
	Б	26	8	0	1
	В	20	2	4	0
	Г	12	0	0	1

Рис. 4. Матриця невідповідності

Матриця невідповідності містить:

– істинно-позитивні результати (визначає кількість правильних розпізнавань певного класу);

– хибно-позитивні результати (визначає кількість розпізнавань певного класу для об'єктів, які не належать до класу);

– хибно-негативні результати (визначає кількість випадків, коли об'єкт було розпізнано неправильно та було віднесено до іншого класу).

Загальну точність проведеної класифікації можна визначити за допомогою значень з матриці невідповідності. Вона проводиться за допомогою формули 3:

$$OvAc = \frac{aA + bB + cC}{N}, \quad (3)$$

де $OvAc$ – загальна точність класифікації;

aA , bB та cC – перетин справжніх та прогнозованих класів (на рисунку 4 вони позначені зелено-блакитним кольором);

N – загальна кількість класів.

Після отримання матриці невідповідності необхідно використати метрики якості для визначення коректності виконаної сегментації. Для початку розраховується значення IOU для кожного класу об'єктів окремо та отримується його середнє значення. Розрахунок IOU для кожного з класів проводиться за допомогою формули 1.

Ще одна метрика якості проведеної обробки даних має назву піксельна точність. Точність пікселів (PA) являє собою показник семантичної сегментації, який показує відсоток правильно класифікованих пікселів. Ця метрика обчислюється співвідношенням між кількістю адекватно класифікованих пікселів і загальною кількістю пікселів у зображенні (4):

$$PA = \frac{\sum_{j=1}^k n_{jj}}{\sum_{j=1}^k t_j}, \quad (4)$$

де n_{jj} – загальна кількість пікселів, класифікованих і позначених як клас j ;

t_j – загальна кількість пікселів, позначених як клас j .

Оскільки в семантичній сегментації присутні декілька класів, середня точність пікселів (mPA) представляє середню точність класу як (5):

$$mPA = \frac{1}{k} \sum_{j=1}^k \frac{n_{jj}}{t_j}, \quad (5)$$

де k – загальна кількість класів.

PA та mPA є інтуїтивно зрозумілими і інтерпретованими показниками. Однак високий PA не є прямим показником кращої продуктивності сегментації, особливо, коли є незбалансовані набори даних класів. У цьому випадку, коли один клас домінує в зображенні, а деякі інші класи складають лише невелику частину зображення, лише правильна класифікація домінантного класу призведе до високого показника PA.

Визначення піксельної точності і є наступним кроком виконання процесу аналізування результатів.

Результати, що були отримані при аналізі виконаної сегментації даних, у будь-якому випадку будуть записані у командний рядок, а окрема візуалізація результатів виконується лише за бажанням користувача. Тому після виведення результатів перевіряється чи потрібно додатково візуалізувати дані аналізу. Якщо так, то в окремих вікнах відтворюються зображення матриці невідповідності та середні значення IOU для кожного класу.

4. Інформаційний простір системи

При конструюванні програмного забезпечення було створено діаграму варіантів використання (рис. 5).

Проаналізувавши поставлене завдання, можна виділити трьох акторів: дрон, оператор системи та система, що використовує оброблені зображення. Система починає свою роботу після отримання фото або відео інформації з дрона. Можливі два типи роботи: з оператором, який контролює вхідні дані та ініціює їх обробку, або без нього. Для даного програмного забезпечення був обраний перший варіант взаємодії. Результати сегментації зображень можливо передати до іншої системи, де це необхідно.

Актор «Дрон» виконує лише «Внесення інформації». Розглядаючи цей прецедент з точки зору дрона, на цій фазі використання інформація зі знятих фото та відео надсилається до системи, яка її обробляє. На цьому взаємодія між актором і системою завершується. Актор «Оператор» має найбільшу кількість варіантів взаємодії з системою. У варіанті використання «Внесення інформації» оператор перевіряє отримані від дрона дані. Якщо вони підлягають та потребують обробки, він вносить їх у систему. Для здійснення обробки даних оператор повинен ініціювати виконання «Сегментація зображення». Цей процес включає в себе функціонал прецедентів «Підготовка зображення», «Зіставлення з шаблоном» та «Аналіз результатів». Це необхідно для коректної роботи алгоритму сегментації даних. Оператор також повинен активувати «Отримання результату» після коректного завершення сегментації для отримання актором «Система з використанням оброблених зображень» результатів. У ході виконання «Отримання результату» обов'язково необхідно зберігати результат. Саме тому прецедент «Отримання результату» використовує також функціонал «Збереження результату».

Також була створена діаграма послідовності для розроблюваного застосунку, вона зображена на рисунку 6.

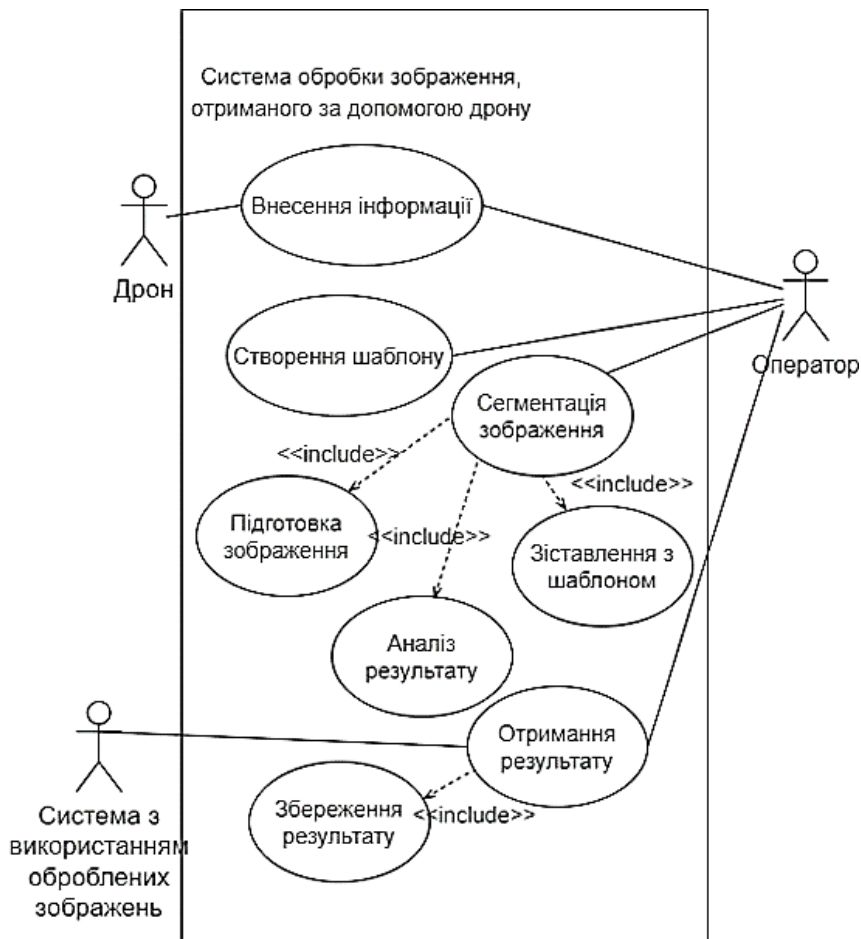


Рис. 5. Діаграма прецедентів

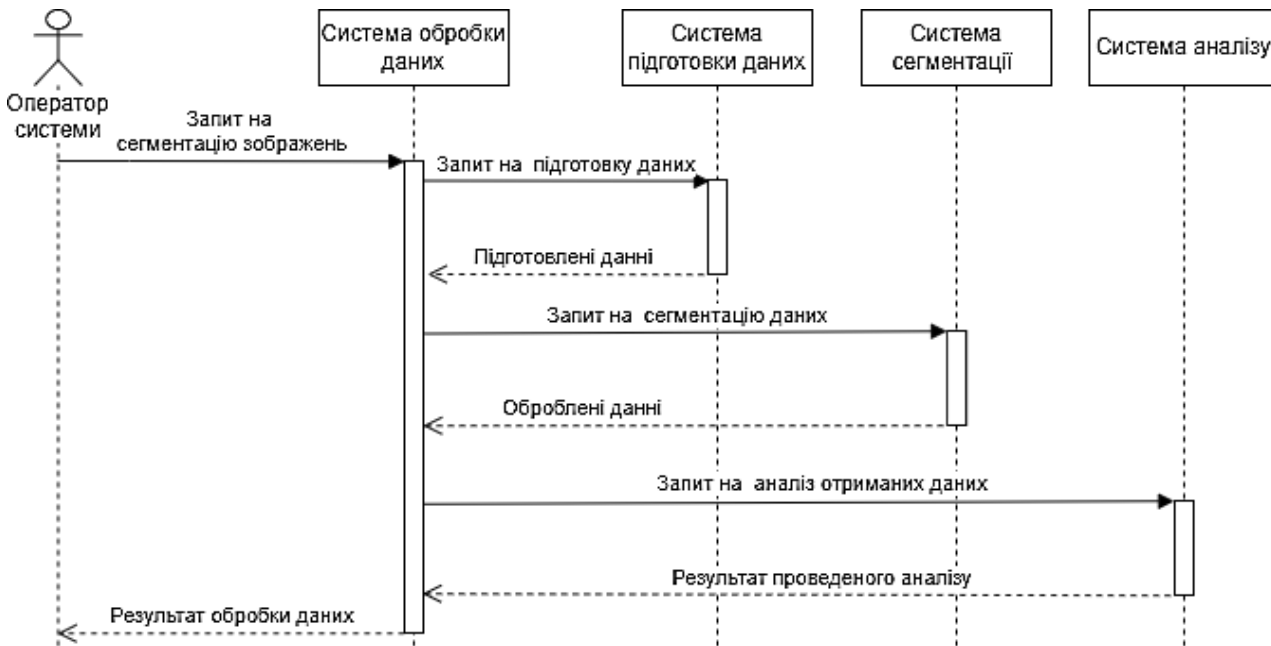


Рис. 6. Діаграма послідовності

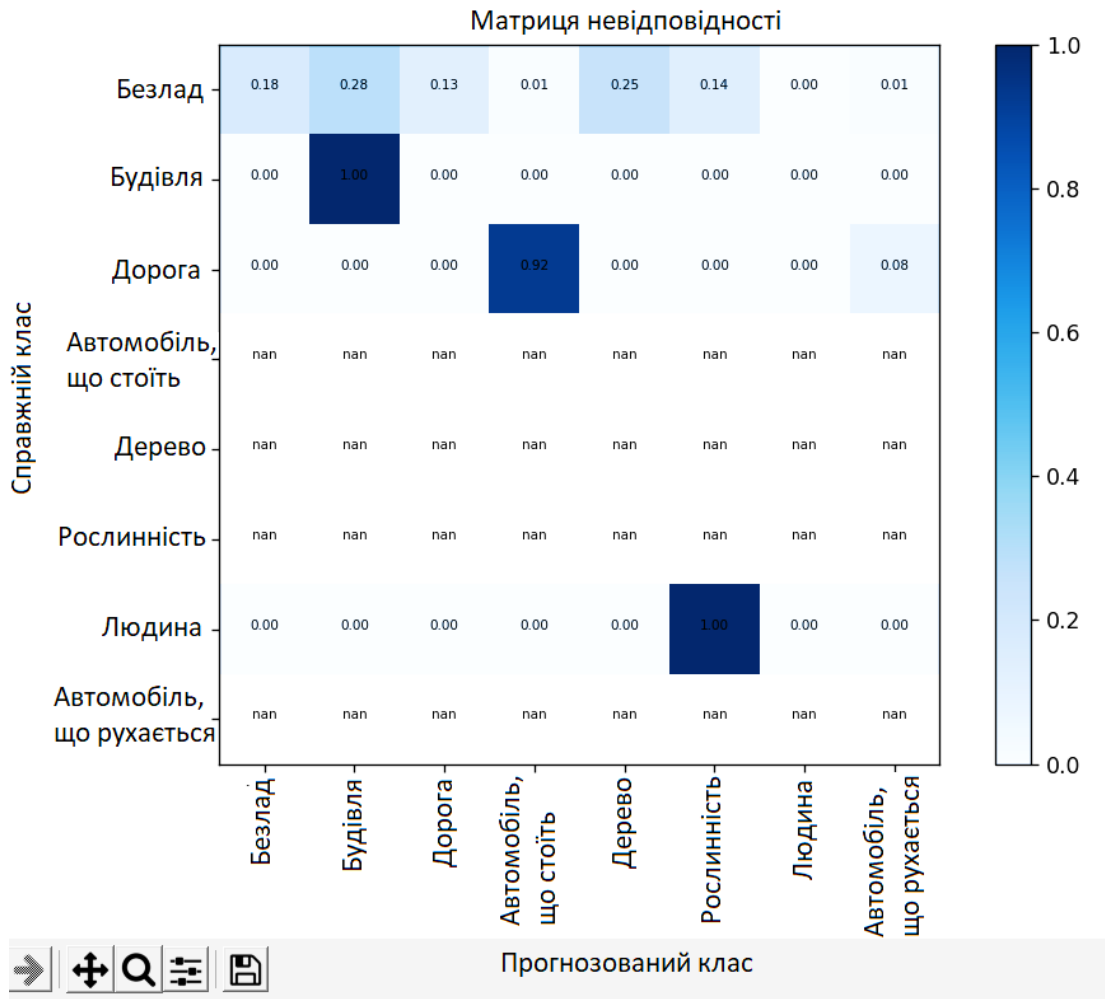


Рис. 9. Матриця невідповідності

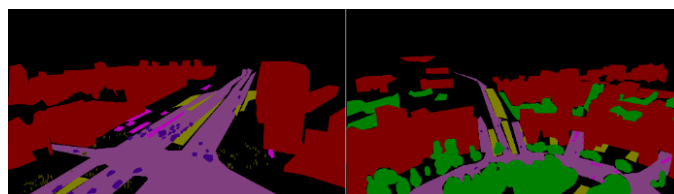
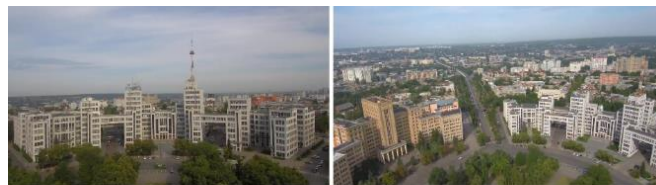
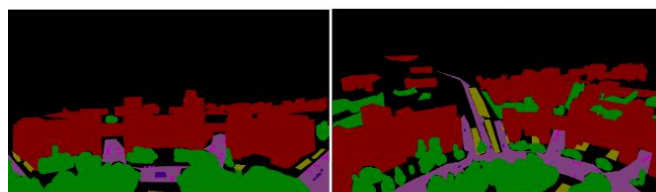


Рис. 10. Приклад створених шаблонів

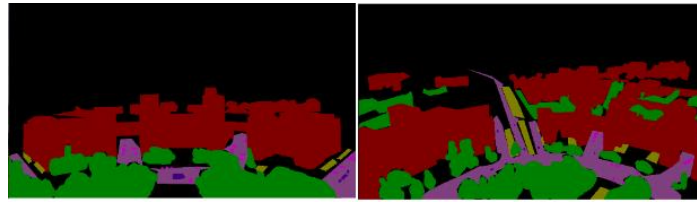


Вхідні зображення



Вхідні шаблони

Рис. 11. Процес обробки даних

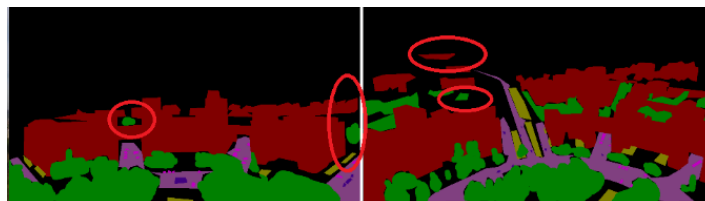


Шаблони, створені програмним забезпеченням

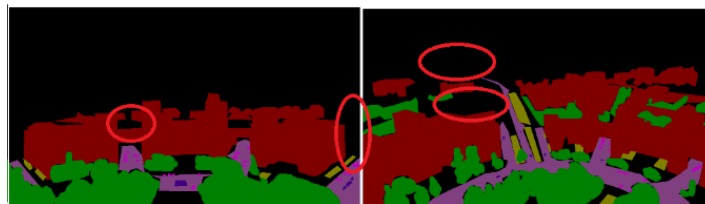


Оброблені зображення

Рис. 12. Результат обробки даних



Вхідні шаблони



Шаблони, створені програмним забезпеченням

Рис. 13. Виділені помилки системи

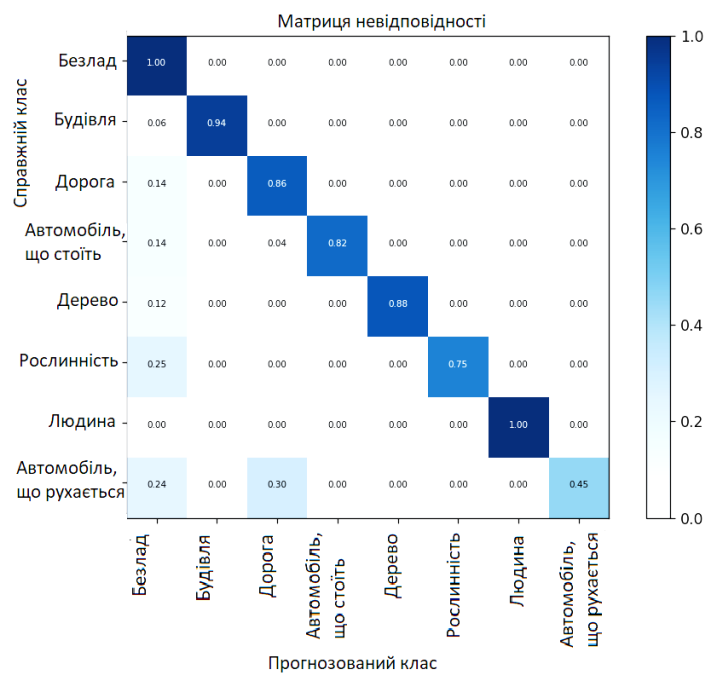


Рис. 14. Матриця невідповідності

```
IOUs: {'Clutter': 0.926536719365698,
'Building': 0.9418007905483611,
'Road': 0.8544821212767469,
'Static_Car': 0.8238462919734243,
'Tree': 0.8776153610544882,
'Vegetation': 0.7515258094548464,
'Human': 1.0,
'Moving_Car': 0.4534565366187543}
mIOU: 0.82865795378654
acc: 0.9573936631944444
```

Рис. 15. Значення IoU та піксельної точності

Якщо переглянути результат аналізу, то можна зробити висновок про виникнення помилок під час обробки інформації.

Помилки можуть бути спричинені такими факторами:

- недостатньо навчальних даних;
- присутні артефакти освітлення;
- погана якість зображення;
- невідомий кут розпізнавання для моделі.

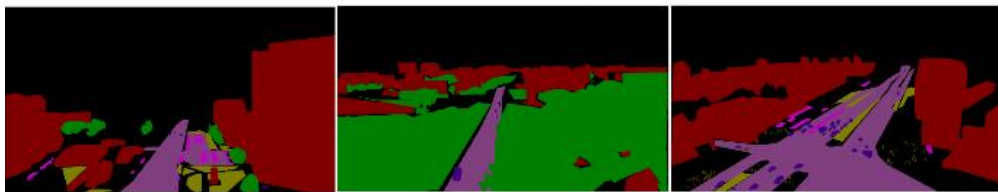
Позбувшись цих чинників, було проведено обробку даних на інших зображеннях. На рисунку 16 зображено вхідні зображення та відповідні їм шаблони.

Підтвердити правильно виконану обробку даних можна виконавши аналіз отриманого результату (рис. 17).

Перевірка якості проведеної сегментації за допомогою IoU та піксельної точності показана на рисунку 18.

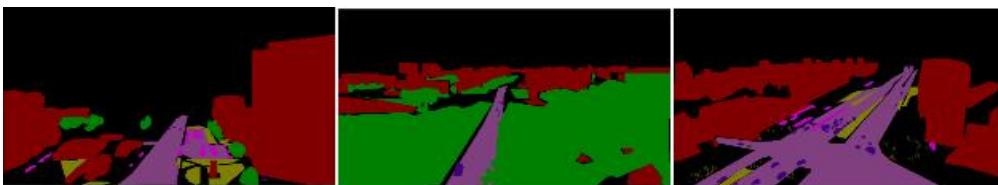


Вхідні зображення



Вхідні шаблони

Рис. 16. Вхідні зображення та відповідні шаблони



Шаблони, створені програмним забезпеченням



Оброблені зображення

Рис. 17 Результат після усунення помилок

```
IOUs: {'Clutter': 1.0, 'Building': 1.0, 'Road': 1.0,
'Static_Car': 1.0, 'Tree': 1.0, 'Vegetation': 1.0,
'Human': 1.0, 'Moving_Car': 1.0}
mIOU: 1.0
acc: 1.0
```

Рис. 18. Перевірка якості за допомогою IoU та піксельної точності

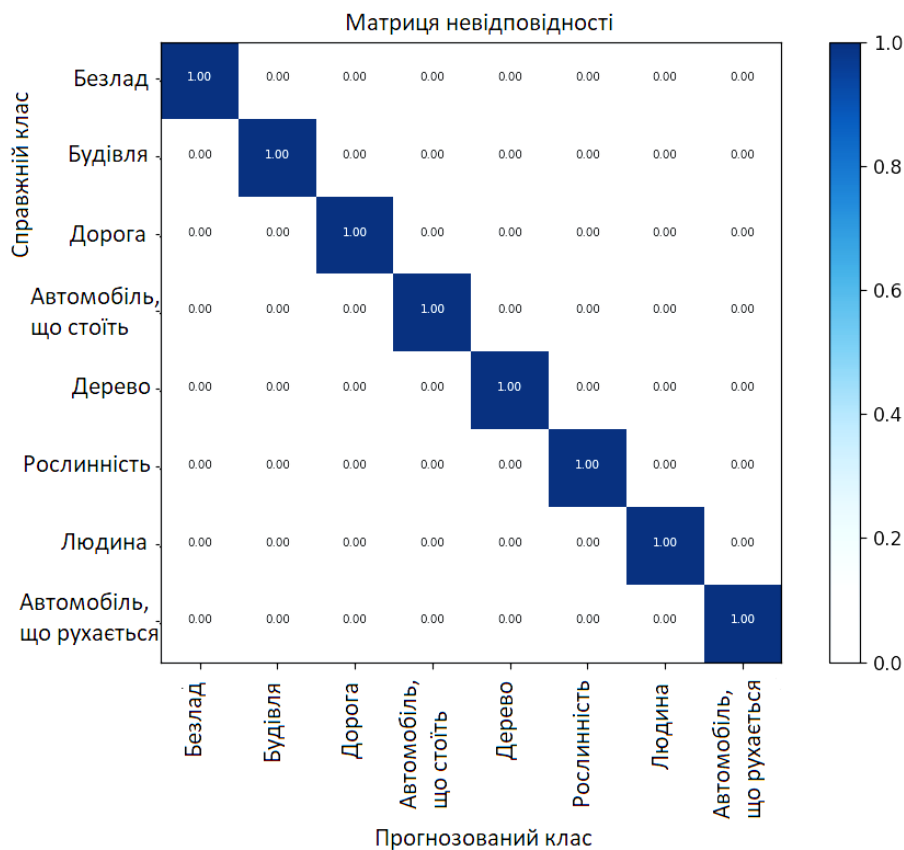


Рис. 19. Матриця невідповідності

Розрахована для отриманих результатів матриця невідповідності зображена на рисунку 19.

Порівнявши отриманий результат з вхідними зображеннями, можна зробити висновок про правильно виконану обробку даних. Іншими словами, кожен з класів визначено вірно та шаблони є ідентичними.

Висновок

Безпілотні літальні апарати зараз є найбільш досліджуваною темою. Вони стають невід'ємною частиною багатьох сфер діяльності. Цьому сприяло зменшення обчислюваної техніки, камер та поява систем глобального позиціонування. Основною сферою використання безпілотних літаючих апаратів є збір фото та відео інформації. Ці дані потребують обробки, тому з розвитком технологій виникли методи, що вирішують проблеми аналізу складних зображень. У кожній з галузей, де необхідна обробка та аналіз фото та відео інформації, використовуються різні алгоритми. Їх вибір залежить від конкретних умов та потреб у цій галузі, бо наразі не існує єдиного універсального рішення.

В ході роботи було проведено аналіз предметної області, розглянуто вже існуючі моделі сегментації зображень, їх недоліки та обмеження. На основі виявлених недоліків та нюансів сформовані вимоги

до системи. Під час проведення аналізу предметної галузі виявлено та описано основні проблеми та визначено задачу, що підлягає автоматизації. В результаті основні процеси були декомпозовані на підзадачі.

Після того як програмне забезпечення було розроблене, воно було протестоване на якість виконання сегментації фото та відеоінформації. Для оцінки якості проведеної обробки даних були використані метрики – IoU, піксельна точність та матриця невідповідності. У ході тестування було визначено помилки, що виникають під час виконання сегментації даних та пояснені можливі причини їх виникнення.

Виходячи з отриманих результатів можна зробити висновок про коректність проведеної сегментації даних, класи розпізнані правильно.

Вдосконалити програмне забезпечення можна шляхом автоматизації процесу створення еталонних шаблонів, провести навчання моделі на більшій кількості даних та інтегрувати створене програмне забезпечення у більшу систему, яка могла б працювати з дроном напряму.

Внесок авторів: формулювання проблеми – Д. М. Крицький; огляд та аналіз інформаційних джерел – О. В. Попов, О. О. Кравцова; реалізація програмного забезпечення – Н. І. Шкуренко; аналіз

результатів – Д. М. Крицький, О. В. Попов;
верифікація результатів – Д. М. Крицький,
О. О. Кравцова.

Усі автори прочитали та погодилися з
опублікованою версією рукопису.

Література

1. *The Cityscapes Dataset for Semantic Urban Scene Understanding [Text]* / M. Cordts, M. Omran, S. Ramos, T. Rehfeld, R. Benenson, U. Franke, S. Roth, B. Shiele // *arXiv*. – 2016. – Article No. 1604.01685. – P. 1-29. DOI: 10.48550/arXiv.1604.01685.

2. *Segmentation and Recognition Using Structure from Motion Point Clouds [Text]* / J. G. Brostow, J. Shotton, J. Fauqueur, R. Cipolla // *Computer Vision – ECCV 2008. ECCV 2008. Lecture Notes in Computer Science*. – Springer, Berlin, Heidelberg, 2008. – Vol. 5302. – P. 44-57. DOI: 10.1007/978-3-540-88682-2_5.

3. Junmo, K. *Highway driving dataset for semantic video segmentation [Text]* / K. Byungju, Y. Junho, K. Junmo // *arXiv*. – 2011. – Article No. 2011.00674. – P. 1-12. DOI: 10.48550/arXiv.2011.00674.

4. *Method for Identifying and Counting Objects [Text]* / O. Pohudina, D. Kritskiy, A. Bykov, T. Szalay // *Integrated Computer Technologies in Mechanical Engineering. Advances in Intelligent Systems and Computing*. – Springer, Cham, 2020. – Vol. 1113. – P. 161-172. DOI: 10.1007/978-3-030-37618-5_15.

5. *Топчієв, Б. С. Алгоритмічно-програмний метод колоризації зображень [Електронний ресурс] : магістерська дис. : 121 Інженерія програмного забезпечення / Топчієв Борис Сергійович. – Київ, 2020. – 148 с. – Режим доступу: <https://ela.kpi.ua/handle/123456789/33832>. – 29.03.2023.*

6. *Чиж, І. Г. Трикомпонентні лінзові зум-афокальні системи трансфокаторів [Текст] / І. Г. Чиж, І. Г. Лісняк // Наукові вісті КІП. – 2019. – №3. – С. 73-79.*

7. *Піпко, А. С. Інтелектуальна система фотореалістичного перенесення стилів між зображеннями [Електронний ресурс] : магістерська дис. : 124 Системний аналіз / Піпко Анна Сергіївна. – Київ, 2018. – 72 с. – Режим доступу: <https://ela.kpi.ua/handle/123456789/23756>. – 29.03.2023.*

8. Patel, M. J. *A novel approach for semantic segmentation of automatic road network extractions from remote sensing images by modified UNet [Text]* / M. J. Patel, A. M. Kothari, H. P. Koringa // *Radioelectronic and computer systems*. – 2022. – No. 3. – P. 161-173. DOI: 10.32620/reks.2022.3.12.

9. *Васильєва, І. К. Метод автоматической кластеризации данных дистанционного зондирования [Текст] / И. К. Васильева, А. В. Попов // Авіаційно-космічна техніка і технологія. – 2019. – №3. – С. 64-75. DOI: 10.32620/akt.2019.3.08.*

10. *Microsoft COCO: Common Objects in Context [Text]* / Ts. Lin, M. Marie, S. Belongie, L. Bourdev,

R. Girshick // *arXiv*. – 2015. – Article No. 1405.0312. – P. 1-15. DOI: 10.48550/arXiv.1405.0312.

References

1. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Benenson, R., Franke, U., Roth, S. & Shiele, B. The cityscapes dataset for semantic urban scene understanding. *arXiv*, 2016, article no. 1604.01685, pp. 1-29. DOI: 10.48550/arXiv.1604.01685.

2. Brostow, G. J., Shotton, J., Fauqueur, J. & Cipolla, R. Segmentation and recognition using structure from motion point clouds. *Computer Vision – ECCV 2008. ECCV 2008. Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg, 2008, vol. 5302, pp. 44-57. DOI: 10.1007/978-3-540-88682-2_5.

3. Byungju, K., Junho, Y. & Junmo, K. Highway driving dataset for semantic video segmentation. *arXiv*, 2011, article no. 2011.00674, pp. 1-12. DOI: 10.48550/arXiv.2011.00674.

4. Pohudina, O., Kritskiy, D., Bykov, A. N. & Szalay, T. Method for Identifying and Counting Objects. *Integrated Computer Technologies in Mechanical Engineering. Advances in Intelligent Systems and Computing*, Springer, Cham, 2020, vol. 1113, pp. 161-172. DOI: 10.1007/978-3-030-37618-5_15.

5. *Топчієв, Б. С. Алгоритмічно-програмний метод колоризації зображень. Магістерська дис. [Algorithmic-program method of image colorization. Master's thesis].* Kiyv, 2020. 148 p. Available at: <https://ela.kpi.ua/handle/123456789/33832> (accessed March 29, 2023).

6. Chyzh, I. H. & Lisnyak, I. H. Trykomponentni lizovi zum-afokal'ni systemy transfokatoriv [Three-component lens zoom-afocal zoom systems]. *Naukovi visti KPI – Scientific news of KPI*, 2019, no. 3, pp. 73-79.

7. *Піпко, А. С. Інтелектуальна система фотореалістичного перенесення стилів між зображеннями. Магістерська дис. [Intelligent system of photorealistic transfer of styles between images. Master's thesis].* Kiyv, 2018. 72 p. Available at: <https://ela.kpi.ua/handle/123456789/23756> (accessed March 29, 2023).

8. Patel, M. J., Kothari, A. M. & Koringa, H. P. A novel approach for semantic segmentation of automatic road network extractions from remote sensing images by modified UNet. *Radioelektronni i komp'uterni sistemi – Radioelectronic and computer systems*, 2022, no. 3, pp. 161-173. DOI: 10.32620/reks.2022.3.12.

9. *Васильєва, І. К. & Попов, А. В. Метод автоматической кластеризации данных дистанционного зондирования [Remote sensing data automatic clustering method].* *Авіаційно-космічна техніка і технологія – Aerospace technic and technology*, 2019, no. 3 (155), pp. 64-75. DOI: 10.32620/akt.2019.3.08.

10. Lin, Ts., Marie, M., Belongie, S., Bourdev, L. & Girshick, R. *Microsoft COCO: Common Objects in Context. arXiv*, 2015, article No. 1405.0312, pp. 1-15. DOI: 10.48550/arXiv.1405.0312.

Надійшла до редакції 19.12.2022, розглянута на редколегії 12.06.2023

DEVELOPMENT OF SOFTWARE FOR DATA SEGMENTATION BY PHOTO AND VIDEO INFORMATION

*Dmytro Kritskiy, Nadiya Shkurenko, Oleksii Popov,
Olexandra Kravtsova*

The object of this research is the process of perception of individual objects in photos and videos for further analysis of the situation in the city. The subject of the research is image processing for selection, classification, and further use of the obtained information about the objects in the photo. The goal is to create information technology for the segmentation of individual segments in photos and videos obtained with the help of unmanned aerial vehicles (in particular, drones) in urban scenes. Main tasks: analysis of existing methods of data segmentation; implementation of an algorithm that would perform segmentation of data by various terrain objects; software testing; and formation of research results. The following results were obtained: existing models of image segmentation were examined for limitations and shortcomings. On the basis of identified shortcomings, the requirements for the developed system were formed. During the analysis of the subject area, the main problems were identified and described and the task to be automated was determined. During the design of algorithms for solving problems, the sequence of the execution of each process was determined. The design of the software included a description of the information space of the system and the user interface. The user interface was chosen after analyzing the main types of user interfaces. To assess the quality of the data processing, existing metrics were used – IoU, pixel accuracy, and the discrepancy matrix. During testing, errors that occur during data segmentation and their causes were determined, after their elimination, the correct result of image segmentation was shown. Conclusions. The scientific novelty of the obtained results is that although there is still no segmentation method that fully satisfies all needs for the analysis of the category of images obtained by drones, the developed software allows uploading images and templates to them as benchmarks for artificial intelligence. perform segmentation and verify the correctness of this segmentation visually and using the IoU metric, the mismatch matrix. The software can be improved by automating the process of creating a reference pattern, training the model on more data, and integrating the created software into a larger system that would work directly with the drone.

Keywords: semantic segmentation; algorithm; metric; UAV; template; mismatch matrix; software.

Крицький Дмитро Миколайович – канд. техн. наук, доцент каф. інформаційних технологій проектування, декан факультету літакобудування, Національний аерокосмічний університет ім. М. С. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Шкурєнко Надія Іванівна – бакалавр каф. «Інформаційні технології проектування», Національний аерокосмічний університет ім. М. С. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Попов Олександр Вікторович – аспірант каф. інформаційних технологій проектування, Національний аерокосмічний університет ім. М. С. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Кравцова Олександра Олександрівна – студ. каф. «Інформаційні технології проектування», Національний аерокосмічний університет ім. М. С. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Dmytro Kritskiy – PhD, associate professor of the Department “Information Technology of Design”, Dean of Aircraft Design Faculty, National Aerospace University «Kharkiv Aviation Institute», Kharkiv, Ukraine, e-mail: d.krickiy@khai.edu, ORCID: 0000-0003-4919-0194, Scopus Author ID: 57195913632, Researcher ID: W-8092-2019, <https://scholar.google.com.ua/citations?user=a4aQ-UQAAAAJ&hl=ru>.

Nadiya Shkurenko – bachelor of the Department "Information design technologies", National Aerospace University «Kharkiv Aviation Institute», Kharkiv, Ukraine, e-mail: n.i.chupryna@student.khai.edu, ORCID: 0000-0001-6824-0148.

Oleksii Popov – PhD student of the Department “Information Technology of Design”, National Aerospace University «Kharkiv Aviation Institute», Kharkiv, Ukraine, e-mail: o.v.popov@khai.edu, ORCID: 0000-0002-2526-9140.

Olexandra Kravtsova – student of the Department "Information design technologies", National Aerospace University «Kharkiv Aviation Institute», Kharkiv, Ukraine, e-mail: o.o.kravtsova@student.khai.edu, ORCID: 0000-0003-1729-9582.