УДК 004.9 : 711.168

Є. М. Крижановський[1]
В. О. Будяк[1]
І. М. Штельмах[1]

ФОРМУВАННЯ ДАТАСЕТУ ТА АВТОМАТИЗОВАНЕ СТВОРЕННЯ ПЛАНУ БУДІВЕЛЬ ЗА АЕРОФОТОЗЙОМКОЮ БУДІВЕЛЬ

[1] Вінницький національний технічний університет

Анотація

Використовуючи сучасні інформаційні технології, проведено дослідження процесу автоматизованого створення планів будівель на основі аерофотозйомки дахів. У роботі запропоновано підхід до обробки зображень за допомогою нейронних мереж, таких як DETR та Vision Transformers, що дозволяє з високою точністю ідентифікувати та сегментувати дахи будівель. Основною метою є створення точних моделей міської забудови для подальшого використання у геоінформаційних системах та плануванні територій. Розроблений підхід сприяє автоматизації процесів аналізу міської інфраструктури та підвищенню точності картографічних даних.

Ключові слова: аерофотозйомка, сегментація дахів, нейронні мережі, інформаційні технології, автоматизація, міське планування.

Abstract

Using modern information technologies, a study was conducted to automate the creation of building plans based on aerial roof photography. The approach leverages image processing through neural networks, such as DETR and Vision Transformers, enabling accurate identification and segmentation of building roofs. The primary goal is to create precise urban development models for further use in geographic information systems and territorial planning. The proposed method enhances the automation of urban infrastructure analysis and improves the accuracy of cartographic data.

Keywords: aerial photography, roof segmentation, neural networks, information technologies, automation, urban planning.

Актуальність дослідження

Створення точних геопросторових даних на основі аерофотозйомки є критично важливим для ефективного планування міської забудови, проєктування інфраструктури та кадастрових робіт. Зокрема, визначення координат фундаментів будівель за знімками дахів дозволяє уникнути тривалих і дорогих наземних вимірювань.

Сучасні методи, реалізовані в таких програмних продуктах, як ArcGIS, QGIS, Digitals, Adobe Photoshop, використовують підходи сегментації зображень за кольорами [1-3], співставлення з існуючими картографічними матеріалами [4] та традиційні методи аналізу геопросторових даних і зображень [4–9]. Проте, їх ефективність знижується за умов складного рельєфу місцевості, нестандартної геометрії дахів, низької контрастності між дахами та іншими елементами ландшафту, а також великої кількості схожих об'єктів на знімках.

Сучасні методи машинного навчання, зокрема трансформери (ViTs) та підходи на основі псевдоміток, показують значний прогрес у задачах аналізу зображень [10]. Їх адаптація для вирішення задач аерофотозйомки може забезпечити суттєве підвищення точності та автоматизації процесу ідентифікації фундаментів будівель.

Таким чином, дослідження, спрямоване на розробку і впровадження сучасних інтелектуальних технологій для аналізу аерофотознімків, є актуальним як з наукової, так і з практичної точки зору.

Формування датасету.

Для ефективного вирішення задачі визначення контурів фундаменту будівель на основі аерофотозйомки доцільно використовувати датасет "Aerial Imagery for Roof Segmentation", розміщений на платформі Kaggle. Цей датасет був створений спеціально для задач, пов'язаних із сегментацією дахів

будівель на аерофотознімках, що робить його одним із найцінніших інструментів для досліджень у сфері комп'ютерного зору. Використання такого набору даних забезпечує доступ до високоякісних і добре розмічених зображень, які дозволяють навчати складні моделі глибинного навчання для точного виявлення та ідентифікації дахів будівель на зображеннях з висоти. Завдяки цьому досягається значне підвищення точності розроблених алгоритмів, що допомагає в автоматизації процесів аналізу великих масивів даних і використовується для вирішення задач у сфері міського планування та управління територіями.

Набір даних "Aerial Imagery for Roof Segmentation" є незамінним ресурсом для вирішення завдань сегментації дахів, які є критичним елементом при створенні детальних топографічних карт і планів міст. Використання цих карт і планів дозволяє містам приймати інформовані рішення щодо майбутнього розвитку інфраструктури, проектування нових житлових або промислових зон, а також ефективного використання земельних ресурсів. У зв'язку з швидким темпом урбанізації та постійним зростанням обсягу даних, отриманих з аерофотозйомок, автоматизація аналізу цих даних є життєво важливою для підвищення ефективності процесів планування та розвитку міських просторів. Впровадження передових методів глибинного навчання, таких як згорткові нейронні мережі (CNN), дозволяє суттєво підвищити швидкість і точність обробки аерофотознімків, що в свою чергу покращує якість прийняття рішень щодо управління міськими ресурсами.

Цей датасет містить високоточні аерофотознімки дахів будівель, а також відповідні бінарні маски, які чітко позначають контури дахів і відокремлюють їх від навколишнього середовища. Це дозволяє моделі в процесі навчання отримувати точні уявлення про форми дахів і їх характеристики, що надалі дає можливість правильно ідентифікувати будівлі на нових знімках. Ілюстрація основних типів дахів, що зустрічаються на аерофотознімках, наведена на рисунку 1.

Рис. 1. Ілюстрація типів дахів

Датасет розділений на три основні частини: тренувальні дані (train), дані для валідації (val) та тестові дані (test). Кожен з цих розділів відіграє важливу роль у процесі навчання і тестування моделі, допомагаючи на кожному етапі оцінювати її продуктивність і точність.

Тренувальний набір є найбільшим і найважливішим розділом датасету. Він включає значну кількість зображень разом із відповідними масками, що дозволяє моделям навчитися розпізнавати та класифікувати дахи будівель. Моделі, треновані на великому обсязі таких даних, отримують здатність ефективно визначати різні типи дахів і виводити закономірності, що забезпечує високу точність при сегментації. Валідуючий набір використовується для перевірки, наскільки добре модель узагальнює отримані знання і може застосовувати їх до нових даних. Це важливий етап у процесі навчання, який допомагає виявляти проблеми перенавчання і гарантувати, що модель здатна працювати з різними типами зображень, не адаптуючись лише до конкретних випадків. Тестовий набір, який складається з нових зображень і масок, не використаних на етапах тренування і валідації, дає можливість об'єктивно оцінити остаточну ефективність моделі на абсолютно нових даних.

Технічні характеристики датасету мають велике значення для його застосування в задачах сегментації дахів. Завдяки високій роздільній здатності зображень можна з великою точністю ідентифікувати об'єкти навіть на найдетальніших аерофотознімках, що є особливо важливим для завдань, які вимагають високої точності. Крім того, всі зображення мають однакові розміри, що спрощує їх обробку перед використанням у моделях глибинного навчання. Уніфіковані розміри дозволяють уникнути додаткових етапів попередньої обробки, таких як нормалізація чи зміна масштабу, що значно полегшує процес підготовки даних для моделювання.

Особливо важливою перевагою цього датасету є висока якість розмітки. Маски сегментації представлені у бінарному вигляді, що чітко розділяє дахи будівель від фону і дає можливість моделі точно вчитися розпізнавати ці об'єкти на нових зображеннях. Якість розмітки є критичною, оскільки навіть найменші похибки у цьому процесі можуть значно вплинути на результат роботи моделі. Точна розмітка гарантує, що моделі зможуть навчитися правильно виділяти будівлі на основі точних даних, що забезпечує високу точність сегментації. Крім того, датасет є масштабованим і може бути застосований до задач, які потребують обробки великих обсягів аерофотознімків. Це робить його ідеальним вибором для вирішення реальних практичних завдань, таких як моніторинг міської інфраструктури або контроль за забудовою територій.

Незважаючи на те, що цей датасет був створений для вирішення конкретної задачі сегментації дахів, його потенційне застосування є значно ширшим. Він може бути використаний для інших завдань комп'ютерного зору, таких як класифікація об'єктів на аерофотознімках, моніторинг міських забудов, оцінка стану будівель тощо. Це робить датасет універсальним ресурсом для виконання широкого спектра задач, пов'язаних із аналізом міської інфраструктури та управління міськими просторами. У сучасних умовах, коли автоматизація аналізу аерофотознімків стає все більш актуальною, використання таких інструментів дозволяє значно скоротити час і зусилля, необхідні для обробки великих обсягів даних.

Цей датасет також є чудовою базою для навчання глибинних нейронних мереж, таких як згорткові нейронні мережі (CNN), які мають широке застосування у задачах обробки зображень. Високоякісні дані, надані у датасеті, дозволяють моделі вчитися на прикладах реальних аерофотознімків і досягати високих результатів при вирішенні практичних завдань. Завдяки цьому такі моделі можуть бути застосовані для аналізу міської інфраструктури, автоматизованого планування територій, оцінки стану будівель, а також для створення детальних карт забудованих територій.

Загалом, датасет "Aerial Imagery for Roof Segmentation" є одним із найбільш корисних і потужних інструментів для вирішення задач, пов'язаних із сегментацією дахів на аерофотознімках. Він відкриває великі можливості для розробки високоточних моделей глибинного навчання, які можуть бути застосовані для різних галузей, включаючи міське планування, аналіз інфраструктури, моніторинг міських забудов і багато інших сфер, де необхідна автоматизація процесів обробки супутникових і аерофотознімків.

Отже, для розв'язання поставленої задачі пропонується удосконалена інтелектуальна технологія, архітектура якої складається із трьох невід'ємних компонентів, призначених для підвищення точності та ефективності всього процесу.

По-перше, запропонована технологія включає в себе механізм локалізації, призначений для ідентифікації груп дахів на великомасштабних зображеннях. Даний початковий крок має ключове значення для виділення регіонів інтересу шляхом зосередження виключно на відповідних областях. Використовуючи моделі детекції об'єктів, такі як DETR, система дозволяє визначити скупчення дахів серед складних міських ландшафтів або різних умов зображення.

По-друге, після фази локалізації запропонована архітектура включає модель сегментації для створення детальних масок сегментації даху. Завдяки використанню ViTs, спеціально навчених для сегментації даху, система може ефективно фіксувати складні деталі та окреслювати окремі дахи. Алгоритм сегментації дахів, що описує цей процес, зображено на рисунку 2.

Рис. 2. Блок-схема алгоритму сегментації дахів

На рисунку 3 зображено загальну взаємодію між компонентами інтелектуальної технології, що включають механізм локалізації на основі DETR, модель сегментації на базі ViTs, а також методи післяобробки, що сприяють точному визначенню контурів дахів та їх подальшому аналізу.

Рис. 3. Взаємодія моделей та інструментів

Охарактеризуємо основні моменти програмної реалізації удосконаленої інтелектуальної технології автоматизованого створення плану будівель за аерофотозйомкою будівель. Виявлення ребер даху зображено на рисунку 6 реалізовано через алгоритм ліній Хафа, який знаходить ребра (лінії) на даху й формує зображення з їх позначенням. Для забезпечення відтворюваності результатів використовується функція seed_everything, яка фіксує генератори випадкових чисел у Python, NumPy та PyTorch.

Завантаження даних зображенньь здійснюється через клас DataRetriever, який завантажує зображення та маски сегментації, нормалізує їх і перетворює в тензори для роботи з моделлю. Шляхи до наборів даних і списки файлів для тренування та валідації налаштовуються окремо. Аугментація даних проводиться за допомогою бібліотеки Albumentations: для тренувальних даних виконується випадкове обрізання до 512×512 пікселів і нормалізація, а для валідаційних даних застосовується лише нормалізація.

Вхідні зображення перетворюються на тензори X, що містять числові представлення, та мітки y, які визначають класи. Аугментація, як-от обертання чи масштабування, покращує здатність моделі узагальнювати результати.

Рисунок 4 демонструє аналіз аерофотознімків: оригінальні зображення порівнюються з масками, які автоматично виділяють дахи будівель. Це допомагає створювати 3D-моделі, аналізувати щільність забудови й автоматизувати картографування.

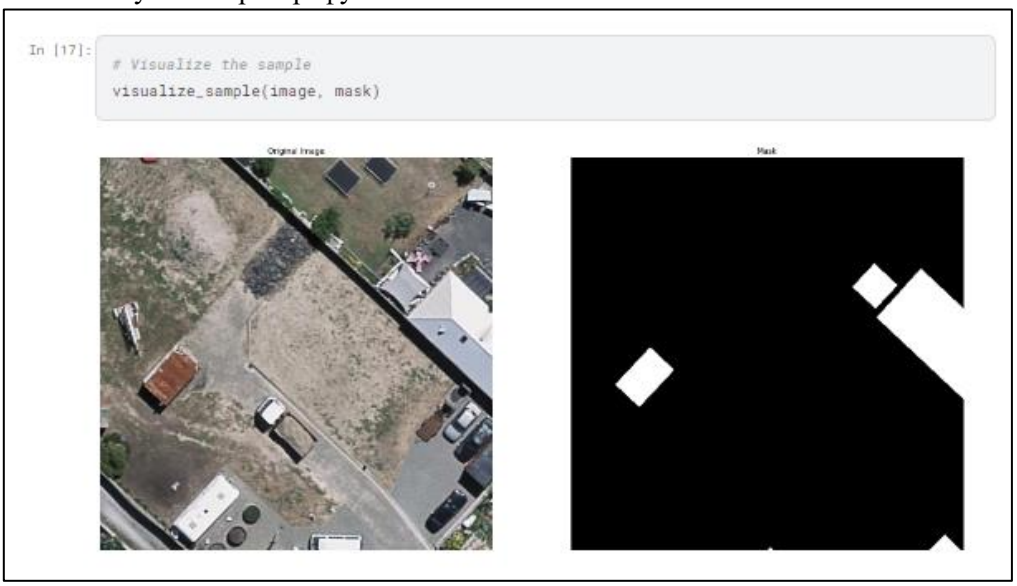

Рис. 4. Результати аналізу аерофотозйомки для визначення дахів будівель

Під час тренування моделі машинного навчання було проведено 10 епох. Основні результати включають значення втрат і точності на тренувальному та валідаційному наборах. Валідаційні показники використовувалися для оцінки здатності моделі узагальнювати на нові дані.

Найкращою виявилася шоста епоха з мінімальною валідаційною втратою 0.2496 і скором NaN. Загальний час тренування склав 3:42:39. Протягом 10 епох тренувальні втрати поступово знижувалися (точність варіювалася від 0.1598 до 0.2108), що свідчить про покращення моделі.

Висновки

Було проведено аналіз тренування моделі машинного навчання з метою оцінки її ефективності та здатності узагальнювати на нові дані. Найкращі результати були досягнуті на шостій епосі, що відзначалася найнижчими валідаційними втратами. Протягом тренування спостерігалося поступове зниження втрат на тренувальному наборі, що свідчить про покращення моделі. Аналіз було виконано із застосуванням сучасних інформаційних технологій та мовою програмування Python, що дозволило провести точні розрахунки моделі.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Є. М. Крижановський, В.Б. Мокін, А.Р. Ящолт, та Л.М. Скорина, Системний аналіз та проектування ГІС. Вінниця, Україна: ВНТУ, 2015.

2. А. І. Зубик, ГІС в урбаністиці та просторовому плануванні: навчально методичний посібник для аудиторної та самостійної роботи студентів з курсу “Використання ГІС в урбаністиці та просторовому плануванні”. Львів, Україна: ЛНУ, 2021.

3. В.Б. Мокін, І.В. Варчук, та Є.М. Крижановський, Інформаційна технологія аналізу та оптимізації топологічної спостережуваності багатозв'язних геоінформаційних систем: монографія. Вінниця, Україна: ВНТУ, 2019.

4. Qi Chen, Lei Wang, Yifan Wu, Guangming Wu, Zhiling Guo, Steven L. Waslander, TEMPORARY REMOVAL: Aerial imagery for roof segmentation: A large-scale dataset towards automatic mapping of buildings, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 147, 2019, Pages 42-55, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2018.11.011.

5. Q. Li et al., "Instance Segmentation of Buildings Using Keypoints," IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 2020, pp. 1452-1455, doi: 10.1109/IGARSS39084.2020.9324457.

6. Guo M, Liu H, Xu Y, Huang Y. Building Extraction Based on U-Net with an Attention Block and Multiple Losses. Remote sensing (Basel, Switzerland). 2020;12(9):1400-. doi:10.3390/rs12091400

7. Sariturk, Batuhan, et al. "FEATURE EXTRACTION FROM SATELLITE IMAGES USING SEGNET AND FULLY CONVOLUTIONAL NETWORKS (FCN)." International Journal of Engineering and Geosciences, vol. 5, no. 3, Oct. 2020, pp.

8. Gaston Lenczner, Adrien Chan-Hon-Tong, Bertrand Le Saux, Nicola Luminari, Guy Le Besnerais. DIAL: Deep Interactive and Active Learning for Semantic Segmentation in Remote Sensing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15, pp.3376 - 3389. ff10.1109/jstars.2022.3166551f, https://ieeexplore.ieee.org/document/9324457.

9. Agustsson, E., et al.: Interactive full image segmentation by considering all regions jointly. In: CVPR. pp. 11622–11631. IEEE (2019).

10. Ю.О. Карпінський, А.А. Лященко, Н.Ю. Лазоренко-Гевель, Основи ГІС. Стандартизація географічної інформації: навч. посіб. Київ, Україна: КНУБА, 2021.

Євгеній Миколайович Крижановський – канд. техн. наук, доцент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: kruzhan@gmail.com;

Будяк Владислав Олександрович – студент групи 2ІСТ-23м, Факультет інтелектуальних інформаційних технологій та автоматизації Вінницького національного технічного університету, Вінниця, budyakqwe@gmail.com.

Ігор Миколайович Штельмах – канд. техн. наук, асистент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: igor.shtelmakh@vntu.edu.ua.

Kryzhanovsky, Evgeniy M. – Cand. Sc. (Eng), Department of Systems Analysis, Vinnytsia National Technical University, Vinnytsia, Vinnytsia National Technical University, Vinnytsia, e-mail: kruzhan@gmail.com;

Budiak Vladyslav O. – student of 2IST-23m group, Faculty of Intelligent Information Technologies and Automation of Vinnytsia National Technical University, Vinnytsia, budyakqwe@gmail.com.

Shtelmah, Igor M. – Cand. Sc. (Eng), Department of Systems Analysis, Vinnytsia National Technical University, Vinnytsia, Vinnytsia National Technical University, Vinnytsia, e-mail: igor.shtelmakh@vntu.edu.ua.