

ВИЗНАЧЕННЯ ТИПУ ЛІСУ З ВИКОРИСТАННЯМ КЛАСИЧНИХ ТА СУЧАСНИХ МЕТОДІВ МАШИННОГО ТА ГЛИБИННОГО НАВЧАННЯ НА ОСНОВІ ЧАСОВИХ РЯДІВ СУПУТНИКОВИХ ДАНИХ

Є. В. Салій^{1,2}, А. М. Лавренюк¹

¹ Навчально-науковий Фізико-технічний інститут

² Інститут космічних досліджень НАН України та ДКА України

Анотація

У даній роботі розглянуто використання супутникових даних, вегетаційних індексів на їх основі, ембедінгів отриманих за допомогою фундаментальної моделі Presto та їх комбінацій як вхід для моделі випадкового лісу для вирішення задачі семантичної сегментації типу лісу на основі часових рядів супутникових даних Sentinel-1 та Sentinel-2.

Було виявлене погіршення результатів сегментації внаслідок використання моделі Presto на 0.09 в рамках міри F1. Найкращі результати, зі значенням F1 на рівні 0.83, продемонструвало як пряме використання випадкового лісу над супутникових даних, так і їх комбінація з вегетаційними індексами.

Ключові слова: фундаментальні моделі, випадковий ліс, семантична сегментація, супутникові дані

Вступ

Світ сьогодні знаходиться в процесі швидких кліматичних та соціально-економічних змін, які потребують просунутого моніторингу та прогнозування динаміки змін земного покриву та землекористування.

У багатьох випадках, для побудови мап земного покриву використовують модель випадкового лісу [1, 2]. Однак в останній час набувають популярність фундаментальні моделі бачення (vision foundation models) для задач дистанційного зондування, які потенційно може бути альтернативою. Фундаментальна модель зазвичай визначається як великомасштабна модель глибинного навчання, яка натренована на широкому наборі даних, яка може бути адаптована до широкого спектра прикладних задач [3]. З червня 2021 було випущено щонайменше 58 [4] таких моделей, серед яких є результати співпраці NASA та IBM у вигляді моделі Prithvi [5, 6], модель NASA Harvest - Presto [7], та багато інших [8].

У рамках Україно-Швейцарської співпраці, в межах проекту DT4LC «Розробка масштабованих моделей цифрових двійників для виявлення змін земного покриву з використанням машинного навчання» пропонується використовувати двохетапний підхід до цифрового двійника. Один із цих етапів включає в собі використання фундаментальних моделей для полегшення та прискорення побудови мап землекористування.

Дане дослідження буде зосереджене на аналізі релевантності моделі Presto [7] у порівнянні з моделлю випадкового лісу в контексті класифікації типу лісу.

1. Постановка задачі та опис даних

Для цього дослідження було взято мультимодальний датасет для виявлення пошкоджень лісу [9]. Даний датасет містить часові ряди Sentinel-1 та Sentinel-2 з просторовим розрізненням починаючи з 10 метрів. Окрім супутникових даних він також містить інформацію про погоду з ERA-5. Датасет покриває 18 локацій на території України та включає як маски пошкодженого лісу (разом із інформацією про тип пошкодження), так і маски типу лісу з Forest Type 2018.

З переліченого ми будемо використовувати лише супутникові дані та маски типу лісу. Таким чином, перед нами стоїть задача семантичної сегментації типу лісу на основі часових рядів супутникових даних. Оскільки маски типу лісу були взяті з Forest Type 2018, то сегментація відбуватиметься в межах чотирьох можливих класів, а саме: не ліс, листяний ліс, хвойний ліс та мішаний ліс.

Приклад даних з використаного датасету можна побачити на Рис. 1. На Рис. 1(а) продемонстровано дані Sentinel-1 з двома наявними в ньому поляризаціями – VV (вертикальна-вертикальна) та VH (вертикально-горизонтальна). На Рис. 1(б) показано кольорову частину Sentinel-2, а саме червону (B4), зелену (B3) та синю (B2) частину спектру. Маска ж типу лісу на Рис. 1(в) демонструє класи в наступному кодуванні кольорів: чорний – не ліс, світло зелений – листяний ліс, темно зелений – хвойний ліс, жовтий – мішаний ліс.

Через те, що в даному датасеті дані для яких наявні дані різняться від локації до локації і загалом

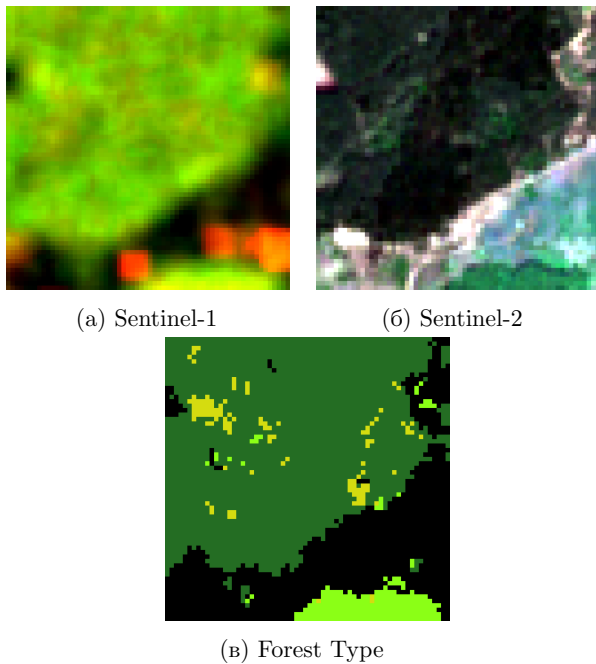


Рис. 1. Приклад даних

їхні розміри є різними, було проведено попередню обробку даних. Дана обробка включає нарізання даних на фрагменти 64x64 пікселів та медіанну фільтрацію темпоральної компоненти кожного окремого каналу. Таким чином, вихідні дані містять 3 часові компоненти, які були отримані як медіана за лютий–квітень, травень–липень, серпень–жовтень для каналів Sentinel-1 та березень–травень, травень–липень, серпень–жовтень для Sentinel-2.

Після цього отримані 13447 фрагменти були розподілені між навчальними, валідаційними та тестовими підмножинами у співвідношенні 70:15:15. У рамках даного дослідження з навчальної підмножини було виділено 80264 ділянки 3x3 пікселі, що використовуватимуться для навчання, серед яких 5264 належать класу мішаного лісу, а решта класів містить по 25 тисяч ділянок кожна. Для тестової підмножини було виконано аналогічні кроки, що дало 868293 ділянки з 118293 мішаного лісу та по 250 тисяч екземплярів інших класів.

2. Методологія

Для вирішення задачі сегментації типу лісу було розглянуто 2 підходи, кожен з яких включає використання моделі випадкового лісу для класифікації класу для кожного пікселя.

Перший підхід полягає в прямому використанні випадкового лісу для наявних супутникових даних. Другий – в використанні енкодер-декодер архітектури, де в ролі енкодера виступає фундаментальна модель Presto [7], а вже отримані за допомогою неї ембедінги подаватимуться на декодер (випадковий ліс) для отримання результатів класифікації.

Моделю Presto (Pretrained Remote Sensing Transformer) [7] була представлена в квітні 2023. Це модель з відкритими вагами з 402 тисячами

параметрами, що була переднавчена на попіксельних часових рядах. Вона була спроектована для роботи з широким спектром вхідних даних, включаючи дані з супутників Sentinel-1 та Sentinel-2, кліматичні дані ERA-5 (а саме температури та вологість на висоті 2 метри), карти висоти (DEM), мапи Dynamic World Land Cover (класи земного покриття), а також просторову (які координати) та часову (за яку дату знімок) інформацію про використані дані. Проте з цього переліку лише довільна підмножина каналів Sentinel-1 і Sentinel-2 разом із мета інформацією про дату та місце знаходження знімку є обов'язковими для роботи моделі. Внаслідок попіксельної природи попереднього навчання, Presto обробляє кожен кожен піксель окремо і відповідно продукує незалежні ембедінги.

Оскільки вегетаційні індекси [10] та інженерія ознак в цілому [11] довели свої спроможності покращувати результати сегментації, дане дослідження також включатиме використання вегетаційних індексів знайдених за допомогою методу з [10].

3. Експеримент

3.1. Пошук інформативного набору ознак

Оскільки метод відбору інформативних ознак з [10] розрахований на роботу з бінарною класифікацією та лише з одним часовим інтервалом, для пошуку вегетативних індексів було використано стратегію one-vs-rest (один проти всіх) за часовий проміжок травень–липень. Таким чином, ми незалежно знаходили оптимальний набір ознак для відокремлення одного класу від всіх інших.

Для простоти вегетаційні індекси обиралися зі схеми нормалізованих різниць (normalized difference, $\frac{A-B}{A+B}$).

Під час кожного з пошуків використовувалися наступні параметри оптимізації:

- Розмір набору: 3
- Кількість епох: 250
- Розмір популяції: 250
- Ймовірність мутації: 10%
- Кількість батьків: 3
- Поріг інформативності та незалежності: 0.05
- Функція незалежності: геометричне середнє

Оскільки набір даних містить 4 різних класи, то фінальна кількість обраних вегетаційних індексів рівна 12.

3.2. Конфігурації навчання

На Рис. 2 можна побачити всі 7 схем навчання. Всі вони використовують модель випадкового лісу для фінальної класифікації, але приймають на вхід різні дані.

Так схема S-RF (Рис. 2(a)) є найбільш стандартною, де супутникові дані напряму передаються в модель. Оскільки вони містять 12 каналів (2 канали Sentinel-1 та 10 каналів Sentinel-2) для кожного з трьох часових інтервалів, то загальна кількість ознак – 36. У S-VI-RF (Рис. 2(б)), супутникові дані

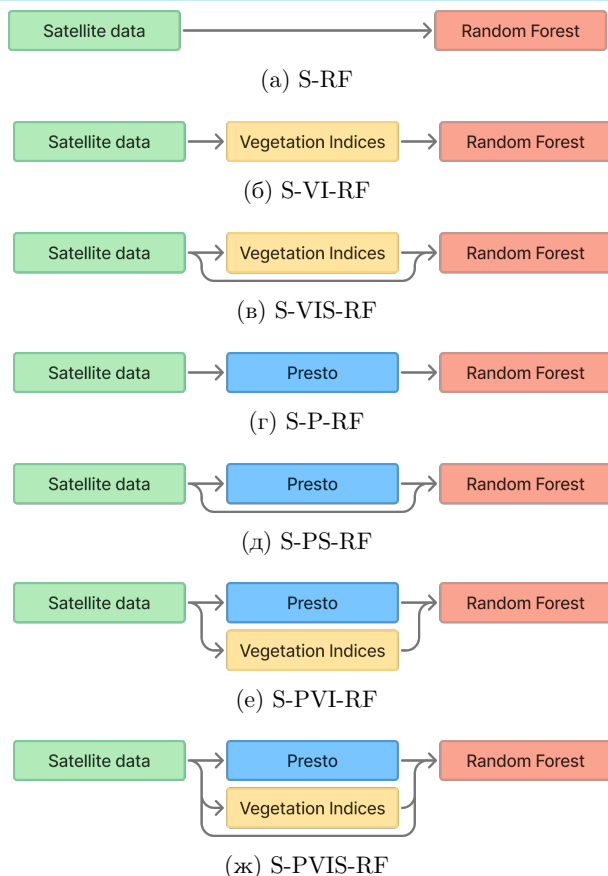


Рис. 2. Розглянуті конфігурації моделі сегментації.

використовуються для обчислення вегетаційних індексів, оскільки їх так само як і каналів – 12, то і кількість ознак також 36. У S-VIS-RF (Рис. 2(в)), на вхід подається об'єднання каналів та вегетаційних індексів, що відповідає 72 вхідним ознакам.

У схемах на Рис. 2(г) – 2(ж) використовуються ембедінги, що отримуються при використанні декодера моделі Presto. У своїй архітектурі дана модель передбачає вихід в 128 елементному ембедінгу. Таким чином S-P-RF (Рис. 2(г)) містить 128 ознак, S-PS-RF (Рис. 2(д)) та S-PVI-RF (Рис. 2(е)) – 164, S-PVIS-RF (Рис. 2(ж)) – 200.

Імплементація моделі випадкового лісу була взята з бібліотеки scikit-learn з наступними параметрами:

- Кількість дерев: 256;
- Мінімальна кількість елементів для спліту: 128;
- Мінімальна кількість елементів в листках: 32;
- Критерій: Джині.

4. Результати

4.1. Обрані вегетаційні індекси

У Таблиці 1 перелічені обрані вегетаційні індекси для найкращого відокремлення клас від всіх інших у порядку від найбільш інформативного (Індекс 1) до найменш інформативного (Індекс 2).

Аналізуючи отриманий набір, варто відзначити, що 7 з 12 ознак використовуються канали Sentinel-1. Це свідчить про важливу роль радарних даних в розпізнаванні типів земного покриття, навіть попри те,

Таблиця 1. Обрані вегетаційні індекси

Клас	Індекс 1	Індекс 2	Індекс 3
Не ліс	$\frac{B5-VH}{B5+VH}$	$\frac{VV-VH}{VV+VH}$	$\frac{B6-B4}{B6+B4}$
Листяний ліс	$\frac{B7-B6}{B7+B6}$	$\frac{B2-B11}{B2+B11}$	$\frac{B8A-VH}{B8A+VH}$
Хвойний ліс	$\frac{B7-VH}{B7+VH}$	$\frac{B5-B11}{B5+B11}$	$\frac{B3-B6}{B3+B6}$
Мішаний ліс	$\frac{VV-B5}{VV+B5}$	$\frac{VV-B11}{VV+B11}$	$\frac{VV-B2}{VV+B2}$

що оптичні дані Sentinel-2 традиційно вважаються основними для аналізу рослинності. Окрім радарних даних активно використовується перший канал короткохвильового інфрачервоного діапазону (B11). Серед обраних вегетаційних індексів жодний не є широковідомим.

4.2. Чисельні результати моделей

У Таблиці 2 перелічені зважені значення точності (precision), повноти (recall) та F1 метрики. З неї можна побачити, що найкращий чисельний результат був досягнутий моделлю з конфігурації S-VIS-RF, що використовує комбінацію супутникових даних та вегетаційних індексів для класифікації типу лісу. Дана конфігурація на 0.01 (0.84 в порівнянні з 0.83) випередила базову модель S-RF в повноті та повторила значення точності та F1.

Варто відзначити, що модель з використанням лише вегетаційних індексів (S-VI-RF) показала досить близькі (на 0.02 гірші по всіх параметрах) результати у порівнянні з S-VIS-RF.

Конфігурації ж з використанням ембедінгів з фундаментальної моделі Presto показали значно гірші результати. В базовій версії S-P-RF всі метрики на 0.09 гірші в порівнянні з S-RF. Додавання до ембедінгу вегетаційних індексів ситуацію суттєво не змінили, піднявши метрика на 0.02 (0.76 у порівнянні з 0.074). Лише при додаванні оригінальних супутникових даних у конфігураціях S-PS-RF та S-PVIS-RF призвели до аналогічних S-RF значень точностей. При цьому S-PS-RF відпрацювала краще ніж S-PVIS-RF в precision та F1 на 0.01. Гірші результати моделі з більшою кількістю вхідних ознак може свідчити про наявність надлишкових та нерелевантних ознак, що ускладнюють сегментацію класів.

Висновки

У ході роботи було проведено дослідження того, наскільки ефективним може бути використання фундаментальної моделі Presto для вирішення задачі семантичної сегментації типу лісу та яким чином можливо покращити результати піксельної класифікації в межах цієї задачі.

Як результат, фундаментальна модель Presto показала себе посередньо при використанні її як енкодер для отримання ознак на основі якої проводилася б класифікація. У всіх розглянутих конфігураціях з її використанням, тобто включаючи комбінування з

Таблиця 2. Результати моделей

Конфігурація	Precision	Recall	F1-score
S-RF	0.83	0.83	0.83
S-VI-RF	0.81	0.82	0.81
S-VIS-RF	0.83	0.84	0.83
S-P-RF	0.74	0.74	0.74
S-PS-RF	0.83	0.83	0.83
S-PVI-RF	0.76	0.76	0.76
S-PVIS-RF	0.82	0.83	0.82

супутниковими даними та вегетаційними індексами, результати залишилися не краще ніж у базового підходу з прямим використанням супутникових даних як вхід для моделі випадкового лісу.

Знайдені за допомогою інженерії ознак вегетаційні індекси продемонстрували важливість радарних даних Sentinel-1 для дискримінації класів навіть в задачах, що традиційно переважно вирішується на основі Sentinel-2. Отримані ознаки, що компонують як радарні так і оптичні дані, демонструють широкий простір для подальшого дослідження. Використання лише вегетаційних індексів для вирішення поставленої задачі сегментації хоч і показали гірші результати, ніж базова модель, але її відставання не є надто значним, тримаючись в межах 0.02 від еталонного значення. Їхнє ж використання у поєднанні з оригінальними супутниковими даними допомогло підняти на 0.01 повноту (з 0.83 до 0.84) отриманої моделі.

Перелік використаних джерел

1. Semi-Supervised Forest Type Mapping in Europe on Satellite Data / N. Kussul, A. Shelestov, B. Yailymov, H. Yailymova // 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). Т. 1. — 2023. — С. 454—458. — DOI: [10.1109/IDAACS58523.2023.10348948](https://doi.org/10.1109/IDAACS58523.2023.10348948).
2. Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data—A Case Study of Czechia / J. Svoboda, P. Štych, J. Laštovička, D. Paluba, N. Kobliuk // Remote Sensing. — 2022. — Т. 14, № 5. — ISSN 2072-4292. — DOI: [10.3390/rs14051189](https://doi.org/10.3390/rs14051189). — URL: <https://www.mdpi.com/2072-4292/14/5/1189>.
3. On the Opportunities and Risks of Foundation Models / R. Bommasani [та ін.]. — 2022. — arXiv: [2108.07258](https://arxiv.org/abs/2108.07258) [cs.LG]. — URL: <https://arxiv.org/abs/2108.07258>.
4. Vision Foundation Models in Remote Sensing: A Survey / S. Lu, J. Guo, J. R. Zimmer-Dauphinee, J. M. Nieusma, X. Wang, P. VanValkenburgh, S. A. Wernke, Y. Huo. — 2024. — DOI: [10.48550/ARXIV.2408.03464](https://doi.org/10.48550/ARXIV.2408.03464). — URL: <https://arxiv.org/abs/2408.03464>.
5. Foundation Models for Generalist Geospatial Artificial Intelligence / J. Jakubik [та ін.]. — 2023. — DOI: [10.48550/ARXIV.2310.18660](https://doi.org/10.48550/ARXIV.2310.18660). — URL: <https://arxiv.org/abs/2310.18660>.
6. Prithvi-EO-2.0: A Versatile Multi-Temporal Foundation Model for Earth Observation Applications / D. Szwarcman [та ін.]. — 2024. — DOI: [10.48550/ARXIV.2412.02732](https://doi.org/10.48550/ARXIV.2412.02732). — URL: <https://arxiv.org/abs/2412.02732>.
7. Lightweight, Pre-trained Transformers for Remote Sensing Timeseries / G. Tseng, R. Cartuyvels, I. Zvonkov, M. Purohit, D. Rolnick, H. Kerner. — 2023. — arXiv: [2304.14065](https://arxiv.org/abs/2304.14065) [cs.CV].
8. Foundation Models for Remote Sensing and Earth Observation: A Survey / A. Xiao, W. Xuan, J. Wang, J. Huang, D. Tao, S. Lu, N. Yokoya. — 2024. — eprint: [arXiv:2410.16602](https://arxiv.org/abs/2410.16602).
9. A Multimodal Dataset for Forest Damage Detection and Machine Learning / H. Yailymova, B. Yailymov, Y. Salii, V. Kuzin, N. Kussul, A. Shelestov // IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium. — 2024. — С. 2949—2953. — DOI: [10.1109/IGARSS53475.2024.10641873](https://doi.org/10.1109/IGARSS53475.2024.10641873).
10. Salii Y., Lavreniuk A., Kussul N. Statistical methods of feature engineering for the problem of forest state classification using satellite data // System research and information technologies. — 2024. — Бєр. — № 1. — С. 86—98. — ISSN 1681-6048. — DOI: [10.20535/srit.2308-8893.2024.1.07](https://doi.org/10.20535/srit.2308-8893.2024.1.07). — URL: <http://dx.doi.org/10.20535/SRIT.2308-8893.2024.1.07>.
11. Transfer learning and single-polarized SAR image preprocessing for oil spill detection / N. Kussul, Y. Salii, V. Kuzin, B. Yailymov, A. Shelestov // ISPRS Open Journal of Photogrammetry and Remote Sensing. — 2025. — Січ. — Т. 15. — С. 100081. — ISSN 2667-3932. — DOI: [10.1016/j.ophoto.2024.100081](https://doi.org/10.1016/j.ophoto.2024.100081). — URL: <http://dx.doi.org/10.1016/j.ophoto.2024.100081>.