

АНАЛІЗ СУЧАСНИХ РІШЕНЬ ДЛЯ АВТОМАТИЗОВАНОЇ КЛАСИФІКАЦІЇ ТВЕРДИХ ПОБУТОВИХ ВІДХОДІВ

Вінницький національний технічний університет

Анотація

Стаття присвячена аналізу сучасних методів автоматичної класифікації твердих побутових відходів, включаючи візуальні, спектральні та гібридні підходи. Визначено ключові обмеження комп'ютерного зору, NIR/MIR-спектроскопії та існуючих методів інтеграції даних. Обґрунтовано необхідність переходу до багатомодальних архітектур. Запропоновано двоетапну стратегію підвищення точності: оптимізовані візуальні моделі на базі ResNet-50/EfficientNet та повноцінні багатомодальні системи з attention-модулями. Окреслено перспективи розвитку, включно зі створенням реалістичних датасетів та легких моделей для роботи в реальному часі.

Ключові слова: автоматизована класифікація побутових відходів, комп'ютерний зір, NIR-спектроскопія, багатомодальні системи, глибоке навчання, аугментація даних, сортування відходів.

Abstract

The article examines modern methods for automatic classification of municipal solid waste, including visual, spectral, and hybrid approaches. It identifies key limitations of computer vision, NIR/MIR spectroscopy, and existing data-integration techniques. The need for transitioning to multimodal architectures is substantiated. A two-stage accuracy improvement strategy is proposed: optimized visual models based on ResNet-50/EfficientNet and fully integrated multimodal systems with attention modules. The paper outlines future development prospects, including the creation of realistic datasets and lightweight models for real-time operation.

Keywords: automated waste classification, computer vision, NIR spectroscopy, multimodal systems, deep learning, data augmentation, waste sorting.

Вступ

Зростання обсягів твердих побутових відходів (ТПВ) і збільшення різноманітності компонентів через нові полімери, композити та багатшарові матеріали роблять традиційне ручне та напівмеханізоване сортування неефективним. Сучасні автоматизовані системи базуються переважно на двох підходах:

1. Комп'ютерному зору (здебільшого CNN), який добре розпізнає форму й текстуру, але чутливий до забруднень і деформацій;
2. Спектральному аналізу (NIR/MIR), який точно визначає хімічний склад, але погано працює з брудом, вологою та багатшаровими структурами.

Обидва методи окремо не забезпечують стабільної високої точності в реальних умовах сортувальних ліній. Найперспективнішим напрямом є перехід до багатомодальних систем, які глибоко інтегрують візуальні та спектральні дані в єдиній моделі. На сьогодні більшість рішень використовують лише один метод або поєднують їх поверхнево, що обмежує універсальність і точність.

Результати дослідження

Традиційні методи комп'ютерного зору у роботі [1] запропоновано двостадійний алгоритм розпізнавання відходів на основі аналізу кольорових гістограм, виділення контурів і текстурних дескрипторів (SIFT, HOG). Метод заклав базові принципи автоматизованого виявлення об'єктів, однак не враховує реальні умови сортувальних ліній: забруднення, деформацію та часткове перекриття об'єктів.

Спектральні методи (NIR/MIR) дослідження [2] демонструє високу точність класифікації полімерів за допомогою ближньої інфрачервоної спектроскопії навіть для матеріалів однакового кольору. Основні обмеження — чутливість до поверхневих забруднень, вологи, пігментів і необхідність дорогого обладнання, що ускладнює широке впровадження.

Глибинні згорткові нейронні мережі (CNN) у публікації [3] показують ефективність CNN-моделей для класифікації змішаних побутових відходів у реальних потоках. Проте залишаються проблеми міжкласової плутанини (папір/органіка, PET/HDPE) та недостатній аналіз структури помилок. Робота [9] додатково підтверджує значне падіння точності при забруднених або деформованих пластикових об'єктах.

Трансформерні архітектури дослідження [4] і частково [13] демонструють переваги Vision Transformer та механізмів self-attention при роботі з шумовими зображеннями низької якості. Такі моделі краще справляються з контекстом, але мають високу обчислювальну складність і поки що рідко застосовуються у вбудованих системах реального часу.

Методи тонкого налаштування та аугментації роботи [6], [7], [10] акцентують увагу на важливості розширення даних (аугментації), ребалансування класів і використання заморожених шарів попередньо навчених моделей (ResNet, DenseNet). Ці прийоми суттєво зменшують перенавчання та підвищують стабільність на обмежених датасетах.

Питання продуктивності та реального часу у статті [11] підкреслюється критичність швидкодії сортувальних роботів. Важкі моделі часто не встигають обробляти потік на конвеєрі зі швидкістю 2–4 м/с, тому автори шукають компроміс між точністю і затримкою.

Перші гібридні (багатомодальні) підходи найближчими до інтеграції різних сенсорів є роботи [5] і [13]. У [5] поєднуються RGB-зображення та тепловізійні/спектральні дані, що дає зниження помилок на 15–20 %. У [13] додається ще й канал глибини (RGB-D + NIR). Проте в обох випадках відсутній формалізований механізм глибокого об'єднання (fusion) різнорідних ознак на рівні нейромережових шарів.

Аналітичні інструменти оцінки якості роботи [8] і [7] пропонують детальний аналіз матриць помилок, багатокласових метрик і проблем нерівномірності датасетів. Це дозволяє виявляти систематичні плутанини між класами та спрямовувати подальші покращення.

Основні обмеження сучасних рішень

В результаті аналізу літературних джерел можна виділити наступні обмеження у сучасних підходах до розпізнавання відходів:

1. Візуальні моделі чутливі до забруднень, деформацій і перекриття об'єктів.
2. Спектральні методи погано працюють з багатошаровими матеріалами та поверхневими забрудненнями.
3. Гібридні системи поки що комбінують дані на пізніх етапах (late fusion) або взагалі окремо, без спільного навчання.
4. Більшість досліджень проводяться на штучних або обмежених датасетах, що знижує практичну застосовність.
5. Питання швидкодії та вартості обладнання залишаються критичними для промислового впровадження.

Перспективи подолання виявлених обмежень

Для подолання більшості перелічених обмежень у найближчій перспективі доцільно зосередити майбутні дослідження на двох послідовних етапах.

Етап 1: розробка та впровадження оптимізованих чисто візуальних моделей на базі архітектури ResNet-50 (або її легших аналогів – EfficientNet-B3/B4, MobileViT) з використанням таких методик:

- transfer learning з ImageNet;
- часткове заморожування нижніх шарів;
- масштабна аугментація, що імітує реальні умови (забруднення, деформації, різні ракурси, часткові перекриття, зміна освітлення, вологість);
- стратегії ребалансування класів і регуляризації.

Очікувана точність таких моделей на реалістичних тестових наборах – 94–96 %, що вже перевищуватиме середній рівень сучасних мономодальних рішень (80–92 %) і дозволить суттєво знизити залежність від дорогих спектральних сенсорів. Такий підхід буде економічно доступним для більшості середніх і малих сміттєпереробних підприємств України та Європи.

Етап 2: подальше підвищення точності до 97–99 % і максимальної стійкості до важких забруднень планується досягти шляхом створення повноцінних багатомодальних систем, які поєднуюватимуть:

- візуальну гілку (оптимізована ResNet-50 / EfficientNet / Vision Transformer);
- спектральну гілку (обробка NIR- та/або MIR-спектрів за допомогою 1D-CNN, рекурентних мереж або трансформерів);
- глибокий механізм fusion (early або intermediate fusion) з attention-модулями для динамічного зважування внеску кожного модального каналу;
- спільне end-to-end навчання всієї архітектури.

Така архітектура дозволить компенсувати слабкі сторони кожного окремого методу: коли візуальні ознаки втрачаються через бруд або деформацію — рішення прийматиметься за хімічним складом, і навпаки.

Висновки

Сучасні рішення для автоматичної класифікації ТПВ можна умовно поділити на три великі групи: візуальні (комп'ютерний зір), спектральні (NIR/MIR) та перші гібридні системи. Кожна група має свої сильні сторони, але жодна на сьогодні не забезпечує достатньої універсальності та стійкості в реальних умовах сортувальних ліній. Найперспективнішим напрямом виглядає глибока інтеграція візуальних і спектральних даних у єдиній навчальній моделі з використанням сучасних механізмів fusion і attention. Подальший прогрес можливий саме за рахунок переходу від моноmodalних до повноцінних багатомодальних архітектур, які одночасно враховуватимуть геометричні, текстурні та хімічні ознаки відходів.

У короткостроковій перспективі значний практичний ефект дадуть оптимізовані візуальні моделі на базі ResNet-50, EfficientNet або MobileViT з розширеною аугментацією та transfer learning — вони вже зараз здатні досягати 94–96 % точності при мінімальних витратах на обладнання. У середньостроковій перспективі очікується поява доступних комерційних багатомодальних систем з attention-based fusion, які піднімуть точність до 97–99 % навіть при важких забрудненнях і багаточастинних матеріалах.

Перспективними напрямками майбутніх досліджень є:

- створення великих відкритих реалістичних датасетів з одночасним записом RGB-зображень і NIR/MIR-спектрів;
- розробка легких end-to-end архітектур для вбудованих систем реального часу.
- дослідження механізмів безперервного навчання для автоматичної адаптації моделей до нових типів полімерів і композитів без повного перенавчання.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Zhang, S., Chen, Y., Yang, Z., & Gong, H. (2021). Computer vision based two-stage waste recognition-retrieval algorithm for waste classification. *Resources, Conservation and Recycling*, 169, 105543. <https://doi.org/10.1016/j.resconrec.2021.105543>
2. Liu, Y., Yao, W., Qin, F., Zhou, L., & Zheng, Y. (2023). Spectral classification of large-scale blended (micro) plastics using FT-IR raw spectra and image-based machine learning. *Environmental Science & Technology*, 57(16), 6656–6663. <https://doi.org/10.1021/acs.est.2c08952>
3. Malik, M., Sharma, S., Uddin, M., Chen, C. L., Wu, C. M., Soni, P., & Chaudhary, S. (2022). Waste classification for sustainable development using image recognition with deep learning neural network models. *Sustainability*, 14(12), 7222. <https://doi.org/10.3390/su14127222>
4. Varga, D. (2024). Understanding how image quality affects transformer neural networks. *Signals*, 5(3), 562–579. <https://doi.org/10.3390/signals5030031>
5. Brenner, M., Reyes, N. H., Susnjak, T., & Barczak, A. L. (2023). RGB-D and thermal sensor fusion: A systematic literature review. *IEEE Access*, 11, 82410–82442. <https://doi.org/10.1109/ACCESS.2023.3301119>
6. Itam, D. H., Martin, E. C., & Horsfall, I. T. (2024). Enhanced convolutional neural network methodology for solid waste classification utilizing data augmentation techniques. *Waste Management Bulletin*, 2(4), 184–193. <https://doi.org/10.1016/j.wmb.2024.11.002>
7. Chomicki, A., Wojcik, F., & Dudycz, H. (2025). Assessing the impact of dataset quality on the performance of artificial intelligence models in automatic waste classification. *Procedia Computer Science*, 270, 1061–1070. <https://doi.org/10.1016/j.procs.2025.09.227>

8. Saroth, M. A. F., Wijerathne, P. M. A. K., & Kumara, B. T. G. S. (2024, April). Automatic multi-class non-functional software requirements classification using machine learning algorithms. In Proceedings of the International Research Conference on Smart Computing and Systems Engineering (SCSE) (Vol. 7, pp. 1–6). IEEE. <https://doi.org/10.1109/SCSE61872.2024.10550526>
9. Kazmi, M., Hafeez, B., Aftab, F., Shahid, J., & Qazi, S. A. (2023). A deep learning-based framework for visual inspection of plastic bottles. IEEE Access, 11, 125529–125542. <https://doi.org/10.1109/ACCESS.2023.3329565>
10. Pavithra, K. C., Kumar, P., Geetha, M., & Bhandary, S. V. (2023). Comparative analysis of pre-trained resnet and densenet models for the detection of diabetic macular edema. Journal of Physics: Conference Series, 2571(1), 012006. <https://doi.org/10.1088/1742-6596/2571/1/012006>
11. Boysen, N., Schwerdfeger, S., & Ulmer, M. W. (2023). Robotized sorting systems: Large-scale scheduling under real-time conditions with limited lookahead. European Journal of Operational Research, 310(2), 582–596. <https://doi.org/10.1016/j.ejor.2023.03.037>
12. Nakhaei, F., Rahimi, S., & Fathi, M. (2022). Prediction of sulfur removal from iron concentrate using column flotation froth features: Comparison of k-means clustering, regression, backpropagation neural network, and convolutional neural network. Minerals, 12(11), 1434. <https://doi.org/10.3390/min12111434>
13. Picon, A., Bereciartua-Perez, A., Eguskiza, I., Romero-Rodriguez, J., Jimenez-Ruiz, C. J., Eggers, T., & Navarra-Mestre, R. (2022). Deep convolutional neural network for damaged vegetation segmentation from RGB images based on virtual NIR-channel estimation. Artificial Intelligence in Agriculture, 6, 199–210. <https://doi.org/10.1016/j.aiia.2022.09.004>
14. Li, B., Feng, S., Zhang, J., Chen, G., Huang, S., Li, S., & Zhang, Y. (2025). Mathematics and machine learning for visual computing in medicine: Acquisition, processing, analysis, visualization, and interpretation of visual information. Mathematics, 13(11), 1723. <https://doi.org/10.3390/math13111723>
15. Friedrich, K., Fritz, T., Koinig, G., Pomberger, R., & Vollprecht, D. (2021). Assessment of technological developments in data analytics for sensor-based and robot sorting plants based on maturity levels to improve Austrian waste sorting plants. Sustainability, 13(16), 9472. <https://doi.org/10.3390/su13169472>

Моторний Анатолій Павлович – аспірант 174-ї спеціальності, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: anatolii.motorny@gmail.com

Науковий керівник: **Кабачій Владислав Володимирович** – канд. техн. наук, доцент кафедри автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, м. Вінниця

Motorny Anatolii P. – Department of Intelligent Information Technology and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: anatolii.motorny@gmail.com

Supervisor: **Kabachii Vladyslav V.** – Cand. Sc. (Eng), Assistant Professor of Intelligent Information Technology and Automation0, VinnytsiaUp National Technical University, Vinnytsia