

Ярошенко Олена Іванівна

кандидат економічних наук, доцент,
доцент кафедри економіко-математичного моделювання,
Чернівецький національний університет імені Юрія Федьковича
ORCID: <https://orcid.org/0000-0002-5976-6559>

Семенович Богдан Степанович

аспірант кафедри економічної теорії, менеджменту і адміністрування,
Чернівецький національний університет імені Юрія Федьковича
ORCID: <https://orcid.org/0009-0000-8779-5966>

Григорків Василь Степанович

доктор фізико-математичних наук, професор,
завідувач кафедри економіко-математичного моделювання,
Чернівецький національний університет імені Юрія Федьковича
ORCID: <https://orcid.org/0000-0003-4866-946X>

Григорків Марія Василівна

доктор економічних, професор,
професор кафедри економіко-математичного моделювання,
Чернівецький національний університет імені Юрія Федьковича
ORCID: <https://orcid.org/0000-0003-3327-991X>

**ЗАСТОСУВАННЯ МАШИННОГО НАВЧАННЯ
ТА КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ ПРОГНОЗУВАННЯ ПОКАЗНИКІВ
ЗАБРУДНЕННЯ ПОВІТРЯ ЯК ФАКТОРУ ДОБРОБУТУ СУСПІЛЬСТВА**

Розроблено методику інтеграції технологій машинного зору з ансамблевими методами машинного навчання для прогнозування показників забруднення атмосферного повітря на основі автоматизованого аналізу транспортних потоків урбанізованих територій. Запропонована система використовує модель YOLO11s для розпізнавання та класифікації транспортних засобів різних категорій та алгоритм ExtraTrees для визначення концентрацій дрібнодисперсних частинок PM_{2.5} і PM₁₀ у повітрі. Емпірична валідація на більш ніж 50 тис. спостереженнях продемонструвала високу точність і стабільність прогнозування та його придатність для автоматизованих систем екологічного моніторингу з метою покращення екологічної ситуації та підвищення добробуту суспільства. Підхід забезпечує переваги порівняно з традиційними методами екологічного моніторингу, включаючи масштабованість та незалежність від додаткової сенсорної інфраструктури.

Ключові слова: машинне навчання, машинний зір, забруднення повітря, прогнозування, добробут суспільства, транспортний моніторинг

Olena Yaroshenko

Candidate of Economic Sciences, Associate Professor,
Associate Professor of the Department of Economic
and Mathematical Modeling,
Yuriy Fedkovych Chernivtsi National University

Bohdan Semenovych

Postgraduate Student at the Department of Economic Theory,
Management and Administration,
Yuriy Fedkovych Chernivtsi National University

Vasyl Hryhorkiv

Doctor of Physical and Mathematical Sciences, Professor,
Head of the Department of Economic and Mathematical Modeling,
Yuriy Fedkovych Chernivtsi National University

Mariia Hryhorkiv

Doctor of Economics, Professor,
Professor of the Department of Economic and Mathematical Modeling,
Yuriy Fedkovych Chernivtsi National University



APPLICATION OF MACHINE LEARNING AND COMPUTER VISION FOR AIR POLLUTION INDICATORS FORECASTING AS A FACTOR OF SOCIAL WELFARE

A methodology for integrating computer vision technologies with ensemble machine learning methods has been developed for atmospheric air pollution forecasting based on automated traffic flow analysis. The proposed approach addresses critical environmental challenges in urban areas where transportation contributes significantly to air quality degradation, directly impacting public health and social welfare. The system architecture employs the YOLO11s model for automated recognition and classification of six vehicle categories: private cars, taxis, commercial vehicles, medium trucks, heavy trucks, and buses in complex urban traffic conditions. The ExtraTrees ensemble algorithm ensures high-precision forecasting of PM_{2.5} and PM₁₀ fine particulate matter concentrations in ambient air through multi-output regression methodology. The research approach incorporates spatio-temporal traffic flow analysis, lag variable utilization for capturing temporal dependencies, and systematic hyperparameter optimization through GridSearchCV with cross-validation techniques. Data integration combines real-time traffic monitoring from computer vision systems with environmental sensor networks, creating comprehensive datasets for machine learning model training and validation. Empirical validation conducted on more than 50,000 observations demonstrates high forecasting accuracy with R^2 values of 0.972 for PM₁₀ and 0.776 for PM_{2.5}, confirming the methodology's stability and suitability for automated environmental monitoring systems in urban territories. The developed approach provides significant advantages over traditional monitoring methods, including scalability, real-time data processing capabilities, reduced operational costs, and independence from additional sensor infrastructure deployment. Research findings contribute to evidence-based environmental policy formulation and urban mobility management strategies, providing a robust foundation for sustainable urban development and environmental protection in metropolitan areas.

Keywords: machine learning, computer vision, air pollution, forecasting, social welfare, traffic monitoring, ensemble methods, environmental monitoring

Постановка проблеми. Забруднення повітря становить одну з найбільш критичних екологічних проблем сучасності, безпосередньо впливаючи на здоров'я населення та якість життя людей усього світу. За даними Інституту впливу на здоров'я (Health Effects Institute, 2024), забруднення атмосферного повітря є другим за значущістю фактором ризику смертності, спричиняючи понад 8 мільйонів смертей у глобальному масштабі [1]. Всесвітня організація охорони здоров'я констатує, що майже все населення планети (99%) вдихає повітря, яке містить високі концентрації забруднювачів [2]. Ця глобальна криза особливо загострюється в межах міських агломерацій, де висока щільність населення та інтенсивна промислова та у цілому економічна діяльність.

Ключову роль у формуванні екологічної ситуації урбанізованих територій відіграє транспортний сектор. Міські транспортні потоки генерують широкий спектр забруднювачів, включаючи дрібнодисперсні частинки (PM_{2.5}, PM₁₀) та діоксид азоту (NO₂), що мають безпосередній вплив на респіраторні та серцево-судинні захворювання у населення [1]. Частка транспорту у викидах парникових газів ЄС-27 очікувано зростатиме, оскільки декарбонізація європейської економіки відбувається швидше в інших секторах, зокрема у виробництві енергії [3]. Це підкреслює необхідність розробки ефективних систем моніторингу та прогнозування транспортного забруднення.

Традиційні методи моніторингу якості повітря, що базуються на стаціонарних сенсорних станціях, характеризуються певними обмеженнями: високою вартістю розгортання та обслуговування, обмеженим просторовим покриттям та відсутністю можли-

вості динамічного реагування на зміни транспортних потоків. Проте сучасні досягнення в галузі комп'ютерного зору та глибокого навчання відкривають нові перспективи для створення економічно ефективних, масштабованих систем екологічного моніторингу.

Аналіз останніх досліджень і публікацій. Класичні підходи до моніторингу атмосферного забруднення базуються на використанні стаціонарних вимірювальних станцій, обладнаних високоточними аналітичними приладами. Kumar et al. (2015) продемонструвала, що такі системи забезпечують високу точність вимірювань, проте характеризуються обмеженим просторовим розподілом та високими експлуатаційними витратами [4]. Альтернативних підходом є супутникові методи моніторингу, розроблені van Donkelaar et al. (2016). Вони забезпечують глобальне покриття, але характеризуються низькою часовою роздільною здатністю та обмеженою точністю для міських територій через вплив хмарності та аерозолів [5].

Паралельно з розвитком традиційних методів моніторингу, зміни в галузі комп'ютерного зору відкрили нові можливості для аналізу транспортних потоків. Redmon et al. (2016) розробили архітектуру YOLO (You Only Look Once), яка забезпечила ідентифікацію об'єктів у реальному часі [6]. Подальші вдосконалення цієї архітектури, включаючи YOLOv11, досягли точності виявлення транспортних засобів понад 95% при збереженні високої швидкості обробки.

Ансамблеві методи машинного навчання демонструють значні переваги у прогнозуванні екологічних параметрів. Geurts et al. (2006) розробили

алгоритм Extremely Randomized Trees (ExtraTrees), який показав високу ефективність у задачах регресії з множинними виходами [7].

Інтеграція комп'ютерного зору з екологічним моніторингом стала предметом досліджень останніх років. Zheng et al. (2020) використовували дані відеоспостереження для оцінювання викидів від транспорту [8]. Економетричні дослідження Karagulian et al. (2015) підтвердили статистично значущий зв'язок між інтенсивністю транспортного руху та концентраціями забруднювачів у міському повітрі [9].

Виділення невирішених раніше частин загальної проблеми. Незважаючи на значні досягнення у галузі комп'ютерного зору та моніторингу довкілля, існуючі дослідження здебільшого зосереджені на окремих аспектах проблеми. Відсутні комплексні моделі, які ефективно інтегрують візуальний аналіз транспортних потоків з прогнозуванням показників забруднення повітря, враховуючи просторово-часові характеристики урбанізованого та економічно активного середовища.

Метою дослідження є розробка та емпірична валідація комплексного підходу застосування ансамблевих методів машинного навчання для прогнозування рівнів забруднення атмосферного повітря на основі автоматизованого моніторингу транспортних потоків із використанням технологій комп'ютерного зору та врахуванням просторово-часових характеристик урбанізованого середовища із достатньо інтенсивною економічною підсистемою.

Виклад основного матеріалу дослідження. Розроблена система базується на п'ятиетапній архітектурі обробки даних, що включає збір та попередню обробку візуальної інформації, автоматичне розпізнавання транспортних засобів, інтеграцію з екологічними даними з платформи OpenAQ та просторово-часове зіставлення, створення консолідованого датасету для машинного навчання та прогнозне моделювання.

Основним джерелом даних про транспортні потоки слугувала платформа NYC Open Data (<https://opendata.cityofnewyork.us>), що містить систематизовані дані про міську мобільність, включаючи відеоматеріали з камер дорожнього спостереження та статистику транспортних потоків за 2016-2024 роки [10].

Відповідно до стандартів NYC Open Data, транспортні засоби класифіковані за наступними шістьма категоріями:

- Auto – приватні легкові автомобілі, включаючи седани, хетчбеки, універсали та позашляховики, призначені для перевезення пасажирів;
- Taxi – ліцензовані таксі та транспортні засоби ride-sharing сервісів (наприклад, Uber, Lyft);
- Commercial – легкі комерційні транспортні засоби, включаючи фургони доставки та службові автомобілі;

– Medium Truck – вантажівки середнього класу, що використовуються для міських перевезень;

– Heavy Truck – автомобілі з вантажопідйомністю близько 15 тонн;

– Bus – автобуси громадського транспорту, шкільні автобуси та туристичні автобуси.

Для отримання даних про якість повітря використовувалася глобальна платформа OpenAQ (<https://openaq.org>), яка агрегує інформацію з понад 12,000 моніторингових станцій у 108 країнах світу [11]. Платформа забезпечує доступ до погодинних вимірювань концентрацій основних атмосферних забруднювачів, включаючи PM2.5, PM10, NO2, SO2, O3 та CO. Для даного дослідження було зосереджено увагу на двох ключових показниках:

– PM2.5 (дрібнодисперсні частинки діаметром менше 2.5 мікрометрів);

– PM10 (частинки діаметром менше 10 мікрометрів).

Вибір саме цих параметрів зумовлений їх критичним впливом на здоров'я населення та наявністю достатньої кількості якісних даних для проведення статистично значущого аналізу у досліджуваному регіоні.

Для реалізації компоненту машинного зору було обрано модель YOLO11s (You Only Look Once), останню версію архітектури You Only Look Once, що забезпечує оптимальний баланс між точністю розпізнавання та швидкістю обробки. Її архітектура складається з трьох компонентів: backbone (екстракція ознак), neck (агрегація ознак) та head (прогнозування).

Базова модель YOLOv11s, попередньо навчена на датасеті COCO, була адаптована до умов міського транспортного середовища за допомогою transfer learning із фіксацією ваг основної частини мережі та подальшою адаптацією вихідних шарів детектора до задачі розпізнавання шести категорій транспортних засобів. Спеціалізований датасет містив 1000 анотованих кадрів міського трафіку з розміткою через платформу Roboflow. Це дозволило досягти значного приросту точності розпізнавання для класів Taxi, Bus та Commercial у складних умовах міського трафіку.

Для забезпечення точного підрахунку транспортних засобів та запобігання дублювання використувався алгоритм SORT (Simple Online and Realtime Tracking). Цей підхід присвоює кожному виявленому об'єкту унікальний ідентифікатор (vehicle_id) та відстежує його траєкторію між кадрами відео.

Результати роботи системи візуального розпізнавання демонструють рис. 1 та 2, де показано автоматичне виявлення та класифікацію транспортних засобів при щільному потоці транспорту у вечірній час, що критично важливо для надійного моніторингу у реальних умовах експлуатації (номерні знаки приховані з міркувань приватності).

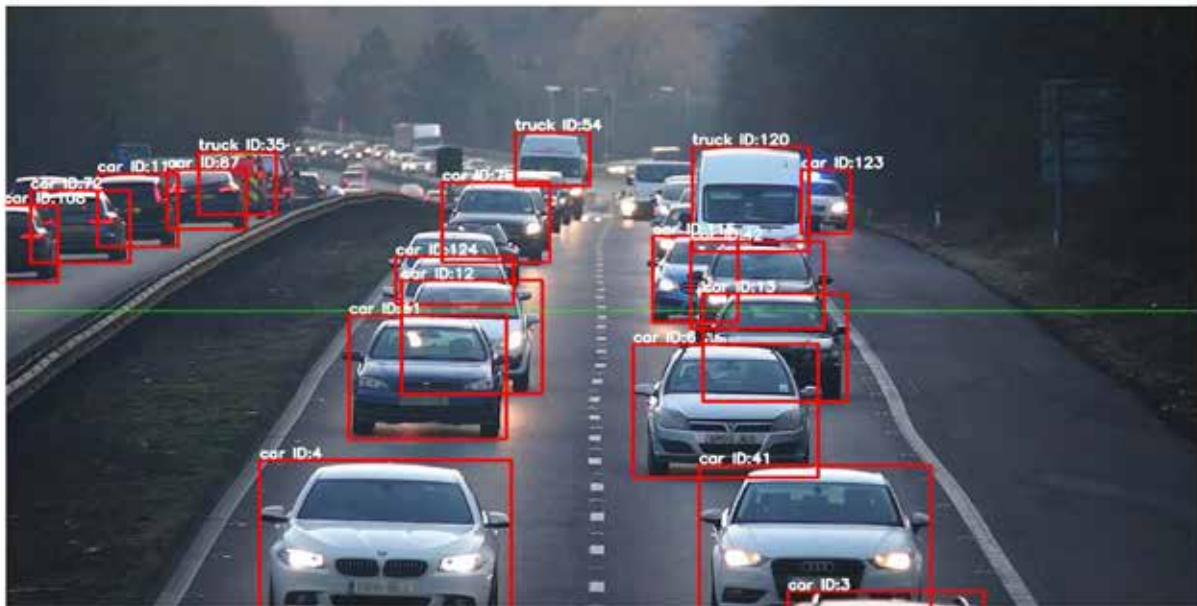


Рис. 1. Результати автоматичної класифікації транспортних засобів системою машинного зору в умовах міського трафіку

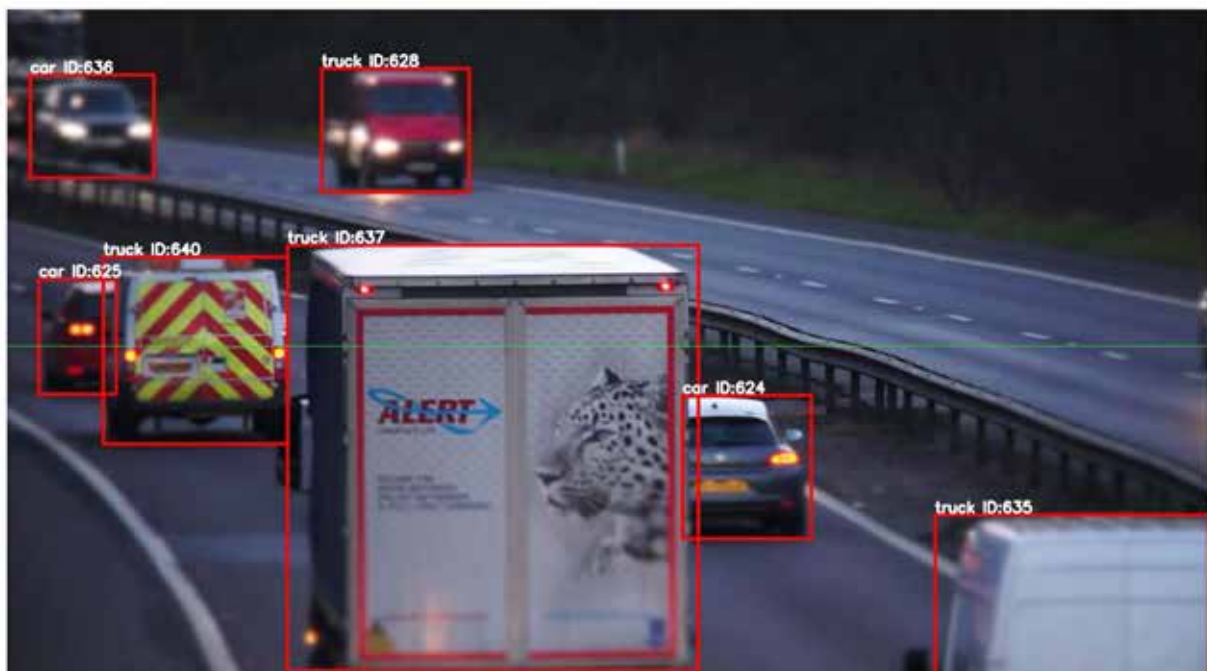


Рис. 2. Виявлення та ідентифікація різних категорій транспортних засобів у змішаному потоці у вечірній час

Просторово-часова інтеграція даних здійснювалася через зіставлення координат розташування відеокамер з найближчими сенсорами якості повітря. Для кожної спостережної точки визначалися релевантні пари “відеопотік → сенсор” на основі мінімальної географічної відстані.

Агрегація транспортних даних здійснювалася у погодинному розрізі з підрахунком кількості тран-

спортних засобів кожного типу. Фінальний аналітичний масив об'єднував транспортну статистику з погодинними показниками якості повітря, створюючи основу для подальшого статистичного аналізу та моделювання.

Емпірична валідація розробленої методології проводилася на основі 52500 даних мегаполісу Нью-Йорк за період 2016–2024 років. Дослідження охо-

пило аналіз взаємозв'язку між інтенсивністю руху різних типів транспортних засобів та концентраціями забруднювачів атмосферного повітря.

Для побудови прогнозової моделі застосовувався алгоритм ExtraTrees (Extremely Randomized Trees) – ансамблевий метод машинного навчання, що базується на множині дерев рішень.

Для одночасного прогнозування концентрацій PM_{2.5} та PM₁₀ застосовувався підхід MultiOutputRegressor, що дозволяє навчити єдину модель для множинних цільових змінних. Цей підхід має суттєві переваги над навчанням окремих моделей: модель враховує кореляції між цільовими змінними, здійснює одноразове навчання замість множинних моделей та забезпечує узгодженість між прогнозами різних забруднювачів.

Особливу увагу приділено кодуванню часових характеристик через тригонометричні функції. Традиційне лінійне кодування години доби (0-23) створює штучний розрив між 23-ю та 0-ю годинами, що не відповідає циклічній природі часових процесів. Тригонометричне кодування вирішує цю проблему через проекцію лінійного часу на одиничне коло. Це забезпечує неперервність часової репрезентації та дозволяє моделі правильно інтерпретувати близькість 23:00 та 01:00. Аналогічно кодуються дні тижня та місяці року для захоплення їх циклічної природи.

Модель використовує наступні групи предикторів:

1. Транспортні змінні: кількість автомобілів кожного з шести типів: Auto (V_1), Taxi (V_2),

Commercial (V_3), Medium Truck (V_4), Heavy Truck (V_5), Bus (V_6);

– lag-змінні (дані за попередню годину) для врахування інерційності транспортних потоків.

2. Часові характеристики:

– година доби (0-23): $H(t)$;

– день тижня (1-7): $D(t)$;

– місяць року (1-12): $M(t)$;

– бінарна змінна "вихідний день": $W(t)$.

3. Історичні екологічні показники:

– концентрація PM_{2.5} за попередню годину: $PM_{2.5}(t-1)$;

– концентрація PM₁₀ за попередню годину: $PM_{10}(t-1)$.

Lag-змінні відіграють критичну роль у моделі, оскільки концентрації забруднювачів демонструють значну автокореляцію через фізичні процеси дисперсії та накопичення у атмосфері. Включення транспортних lag-змінних дозволяє врахувати затримку між емісією та її вимірюванням сенсорами.

На основі введених змінних прогнозна модель формалізована як:

$$PM_i(t) = f(V_1(t-1), V_2(t-1), V_3(t-1), V_4(t-1), V_5(t-1), V_6(t-1), H(t), D(t), M(t), W(t), PM_i(t-1)) + I(t-1) + \varepsilon(t),$$

де $I(t)$ – промислова активність у момент часу t ;

$\varepsilon(t)$ – випадкова похибка у момент часу t .

Графічний аналіз результатів (рис. 3 та 4) демонструє високу кореляцію між прогнозованими та фактичними значеннями концентрацій забруднювачів. Порівняння фактичних та прогнозованих значень

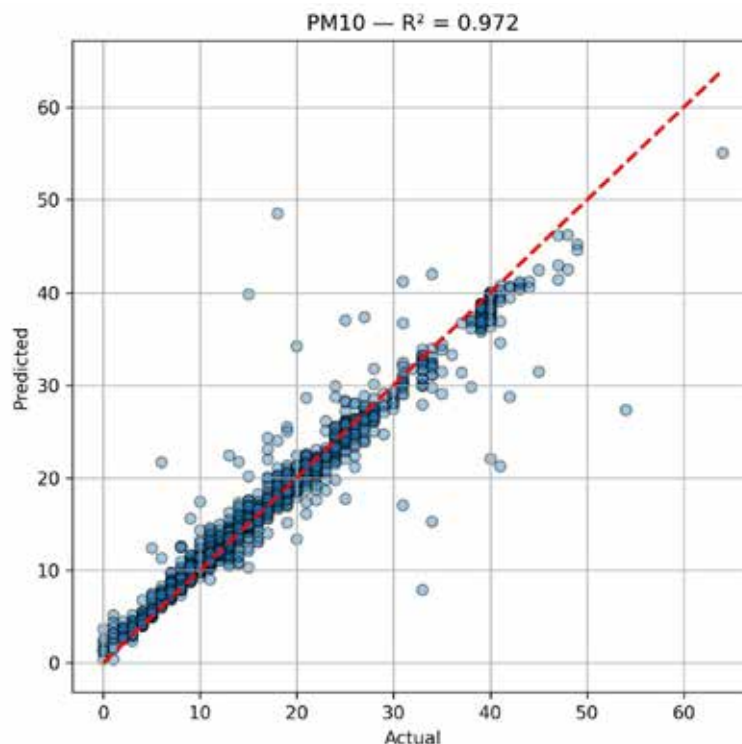


Рис. 3. Порівняння фактичних та прогнозованих значень концентрації PM₁₀ ($R^2=0.972$)

концентрацій PM10 (рис. 3) демонструє високу точність моделі з $R^2 = 0.972$, при цьому розподіл точок навколо лінії ідеального прогнозу свідчить про відсутність систематичних похибок.

Для PM2.5 (рис. 4) спостерігається дещо більша варіабельність з $R^2 = 0.776$, що пояснюється більшою складністю динаміки дрібнодисперсних частинок, які піддаються впливу додаткових факторів, включаючи вторинне утворення в атмосфері та далекий транспорт. Проте загальна тенденція залишається чітко вираженою з мінімальними відхиленнями у високих концентраціях.

Аналіз важливості змінних у моделі ExtraTrees виявив наступні закономірності:

- Таксі та Commercial мають найвищий вплив на концентрації PM10 та PM2.5, що, ймовірно, пов'язано з їх інтенсивною експлуатацією в бізнес-години, постійним рухом у міському циклі та тривалим перебуванням у заторах;

- Heavy Truck та Medium Truck, незважаючи на високі питомі викиди, демонструють менший сумарний вплив, ймовірно, через їх переважну експлуатацію в періоди низького навантаження транспортної системи;

- часові характеристики (година доби, день тижня) суттєво впливають на модельні оцінки.

Перехресна валідація методом k-fold ($k=5$) підтвердила стабільність результатів моделювання. Середні значення R^2 для різних підвибірок складають:

- PM10: $R^2 = 0.981 \pm 0.008$;

- PM2.5: $R^2 = 0.768 \pm 0.015$.

Висновки. Розроблений підхід застосування машинного зору для моніторингу транспортного навантаження та прогнозування забруднення повітря демонструє високу ефективність та практичну значущість для вирішення екологічних проблем сучасних великих міст.

Ключові результати дослідження можна поділити на такі категорії:

- *методологічні*: успішна інтеграція технологій комп'ютерного зору з ансамблевими методами машинного навчання для створення комплексної системи екологічного моніторингу, використання тригонометричного кодування часових характеристик для врахування циркадних ритмів транспортних потоків та MultiOutputRegressor для одночасного прогнозування множинних забруднювачів;

- *технічні*: адаптація моделі YOLO11s для точного розпізнавання та класифікації міських транспортних засобів з високим рівнем точності для різних категорій транспорту; досягнення високої прогнозовної точності з $R^2 = 0.972$ для PM10 та $R^2 = 0.776$ для PM2.5 на основі валідації 52,500 спостережень з підтвердженням стабільності результатів через перехресну валідацію;

- *соціально-економічні*: розробка масштабованої архітектури, що дозволяє адаптацію для різних міських контекстів без потреби у додаткових сен-

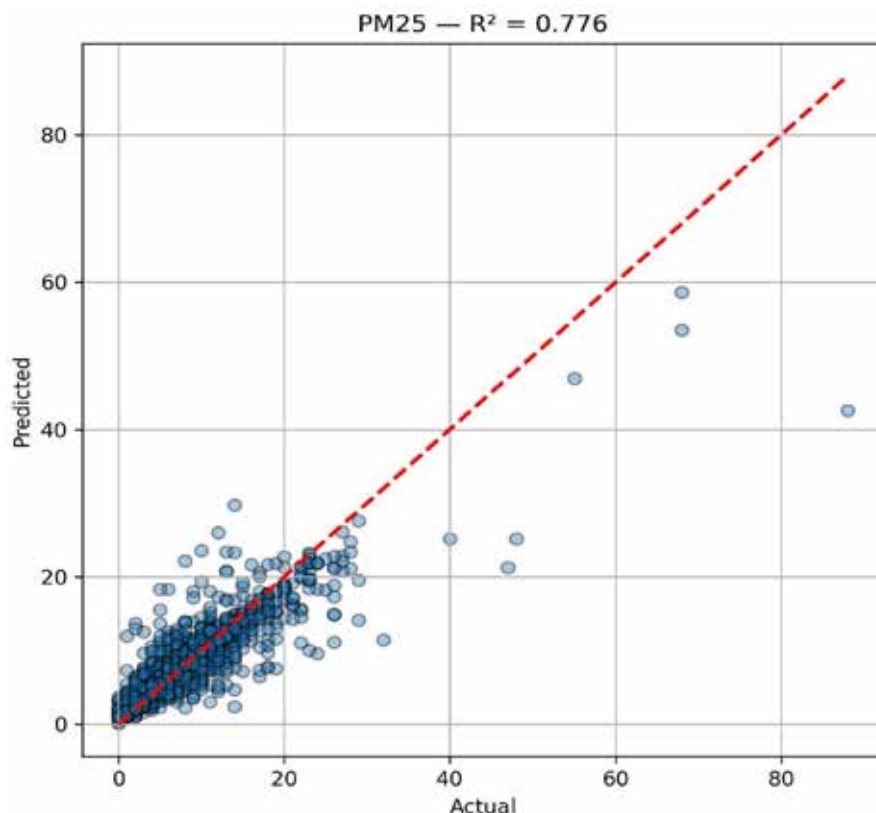


Рис. 4. Порівняння фактичних та прогнозованих значень концентрації PM2.5 ($R^2=0.776$)

сорних мережах; створення економічно ефективного рішення для автоматизованого моніторингу екологічного стану урбанізованих територій; безпосередній вплив на добробут суспільства через зниження витрат на охорону здоров'я, підвищення продуктивності праці та оптимізацію транспортних потоків для сталого розвитку міських систем.

Таким чином, результати дослідження мають важливе значення для формування науково обґрунтованої політики у сфері екологічного моніторингу

та управління міською мобільністю. Впровадження запропонованої методології сприятиме підвищенню якості життя міського населення через покращення екологічної ситуації та оптимізацію транспортних потоків, що впливає на добробут суспільства. Перспективи подальших досліджень включають розширення спектру моніторингу забруднювачів, інтеграцію з системами штучного інтелекту нового покоління та розробку платформ для прийняття управлінських рішень у реальному часі.

СПИСОК ЛІТЕРАТУРИ

1. Health Effects Institute. State of Global Air 2024 : Special Report. Boston, MA : Health Effects Institute, 2024. URL: https://www.stateofglobalair.org/sites/default/files/documents/2024-06/soga-2024-report_0.pdf (дата звернення: 15.03.2025).
2. WHO. Billions of people still breathe unhealthy air: new WHO data. 2022. URL: <https://www.who.int/news/item/04-04-2022-billions-of-people-still-breathe-unhealthy-air-new-who-data> (дата звернення: 15.03.2025).
3. European Environment Agency. Sustainability of Europe's mobility systems : Web report no. 01/2024. 2024. URL: <https://www.eea.europa.eu/en/analysis/publications/sustainability-of-europes-mobility-systems/climate> (дата звернення: 15.03.2025).
4. Kumar P., Morawska L., Martani C., Biskos G., Neophytou M., Di Sabatino S., Bell M., Norford L., Britter R. The rise of low-cost sensing for managing air pollution in cities. *Environmental International*. 2015. Vol. 75. P. 199-205. DOI: <https://doi.org/10.1016/j.envint.2014.11.019>
5. van Donkelaar A., Martin R. V., Brauer M., Hsu N. C., Kahn R. A., Levy R. C., Lyapustin A., Sayer A. M., Winker D. M. Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. *Environmental Science & Technology*. 2016. Vol. 50(7). P. 3762-3772. DOI: <https://doi.org/10.1021/acs.est.5b05833>
6. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA, 2016. P. 779-788. DOI: <https://doi.org/10.1109/CVPR.2016.91>
7. Geurts P., Ernst D., Wehenkel L. Extremely randomized trees. *Machine Learning*. 2006. Vol. 63(1). P. 3-42. DOI: <https://doi.org/10.1007/s10994-006-6226-1>
8. Zheng Y., Liu F., Hsieh H. P. U-Air: When urban air quality inference meets big data. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2013. P. 1436-1444. DOI: <https://doi.org/10.1145/2487575.2488188>
9. Karagulian F., Belis C. A., Dora C., Prüss-Ustün A. M., Bonjour S., Adair-Rohani H., Amann M. Contributions to cities' ambient particulate matter (PM): A systematic review of local source contributions at global level. *Atmospheric Environment*. 2015. Vol. 120. P. 475-483. DOI: <https://doi.org/10.1016/j.atmosenv.2015.08.087>
10. NYC Open Data. 2024. URL: <https://opendata.cityofnewyork.us/data/> (дата звернення: 15.03.2025).
11. OpenAQ. 2024. URL: <https://openaq.org/> (дата звернення: 15.03.2025).

REFERENCES

1. Health Effects Institute. (2024) State of Global Air 2024. *Special Report*. Boston, MA: Health Effects Institute. Available at: https://www.stateofglobalair.org/sites/default/files/documents/2024-06/soga-2024-report_0.pdf (accessed March 15, 2025)
2. WHO. (2022) Billions of people still breathe unhealthy air: new WHO data. Available at: <https://www.who.int/news/item/04-04-2022-billions-of-people-still-breathe-unhealthy-air-new-who-data> (accessed March 15, 2025)
3. European Environment Agency. (2024) Sustainability of Europe's mobility systems. *Web report no. 01/2024*. Available at: <https://www.eea.europa.eu/en/analysis/publications/sustainability-of-europes-mobility-systems/climate> (accessed March 15, 2025)
4. Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., & Britter, R. (2015) The rise of low-cost sensing for managing air pollution in cities. *Environmental International*, vol. 75, pp. 199-205. Available at: <https://doi.org/10.1016/j.envint.2014.11.019> (accessed March 15, 2025)
5. van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., Lyapustin, A., Sayer, A. M., & Winker, D. M. (2016) Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. *Environmental Science & Technology*, vol. 50(7), pp. 3762-3772. Available at: <https://doi.org/10.1021/acs.est.5b05833> (accessed March 15, 2025)
6. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 779-788. Available at: <https://doi.org/10.1109/CVPR.2016.91>
7. Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3-42. Available at: <https://doi.org/10.1007/s10994-006-6226-1> (accessed March 15, 2025)
8. Zheng, Y., Liu, F., & Hsieh, H. P. (2013) U-Air: When urban air quality inference meets big data. *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1436-1444. Available at: <https://doi.org/10.1145/2487575.2488188> (accessed March 15, 2025)
9. Karagulian, F., Belis, C. A., Dora, C., Prüss-Ustün, A. M., Bonjour, S., Adair-Rohani, H., & Amann, M. (2015) Con-

tributions to cities' ambient particulate matter (PM): A systematic review of local source contributions at global level. *Atmospheric Environment*, vol. 120, pp. 475-483. Available at: <https://doi.org/10.1016/j.atmosenv.2015.08.087> (accessed March 15, 2025)

10. NYC Open Data. (2024) Available at: <https://opendata.cityofnewyork.us/data/> (accessed March 15, 2025)

11. OpenAQ. (2024) Available at: <https://openaq.org/> (accessed March 15, 2025)

Стаття надійшла: 20.03.2025

Стаття прийнята: 23.04.2025

Стаття опублікована: 30.05.2025