

УДК 004.93:004.032.6:[004.932.2]

**МУЛЬТИМЕДІЙНА СИСТЕМА ВАЛІДАЦІЇ ОБ'ЄКТІВ
МІСЬКОЇ ІНФРАСТРУКТУРИ: ПОРІВНЯЛЬНИЙ АНАЛІЗ
НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ ТА ЕКСПЕРТНИХ ОЦІНОК**

Зіма В.О.

e-mail: volodymyr.zima@nure.ua

Науковий керівник – канд. техн. наук, ст. викл. Кобилін І.О.
Харківський національний університет радіоелектроніки, каф. ІНФ
м. Харків, Україна

This paper presents a hybrid multimedia validation system for crowdsourced urban accessibility reports that combines convolutional neural networks with a Human-in-the-Loop verification interface. Uploaded photos are first processed by CNN-based classifiers and perceptual-hashing to detect and de-duplicate candidate obstacles (e.g., ramps, high curbs, missing tactile tiles). Each prediction is accompanied by a confidence score; low-confidence or ambiguous cases are flagged as disputed and routed to users for manual review. User confirmations and corrections are fed back into an active-learning pipeline to incrementally improve model performance. The proposed two-stage architecture balances automated scalability with human judgment, minimizing erroneous automatic decisions in challenging visual conditions. We evaluate the approach on a representative crowdsourced dataset and discuss metrics for accuracy, annotation effort, and user validation time. Practical considerations for privacy, edge vs. cloud deployment, and UX design are addressed. The system aims to produce reliable, georeferenced accessibility maps to support safer route planning for mobility-restricted populations.

При зборі інформації з краудсорсингових платформ, а зокрема і для міської інфраструктури, генерується великий потік візуальних даних (фото перешкод, пандусів тощо). Валідація та обробка цих даних є однією з основних проблем, що перешкоджають ефективній роботі програми. При цьому ручна модерація кожного окремого зображення може бути неможливою через обсяги інформації, а повністю автоматична оцінка з використанням комп'ютерного зору може часто давати помилку, як приклад, тінь від дерева нейромережа може сприймати як глибоку яму, або наприклад, неправильно оцінювати кут нахилу пандуса.

Спеціально для вирішення цієї проблеми пропонується розробити гібридну мультимедійну систему валідації, що базуватиметься на парадигмі Human-in-the-Loop. Основною ідеєю цього підходу, буде часткове зняття відповідальності з автоматичної системи, на користь рішень, що приймає сам користувач, тобто людина сама остаточно обирає маршрут з врахуванням всіх можливих перешкод на шляху. Паралельно цьому, автоматична система має поступово покращувати свої передбачення, навчаючись з урахуванням рішень людини. Запропонована архітектура має розділяти процес

верифікації на два етапи: автоматизований аналіз та експертну оцінку користувачем.

На першому (автоматизованому) етапі кожне завантажене зображення проходить попередню обробку за допомогою згорткової нейронної мережі (CNN). Архітектура моделі налаштована на вирішення задачі багато-класової класифікації елементів інфраструктури (наприклад, «пандус», «високий бордюр», «відсутність тактильної плитки»). Результатом роботи моделі є не лише визначення ймовірного класу об'єкта, але й обчислення метрики впевненості мережі (confidence score). Також, для швидкого семантичного порівняння та відсіювання повних дублікатів фотографій можуть бути використані алгоритми перцептивного хешування (pHash).

На другому етапі, результати машинного аналізу формують та додають в мультимедійний інтерфейс системи у вигляді інформаційних «карток об'єктів». Така картка містить оригінальне зображення, просторові метадані та попередній висновок нейромережі з рівнем її впевненості. Якщо показник достовірності моделі падає нижче заданого порогу (наприклад, за складних умов освітлення або під нестандартними кутами), система не прийматиме рішення самостійно. Натомість картка буде позначена як «спірна» та представлена користувачеві застосунку для ручного підтвердження або спростування гіпотези III.

Human-in-the-Loop Validation System for Crowdsourced Urban Accessibility Reports

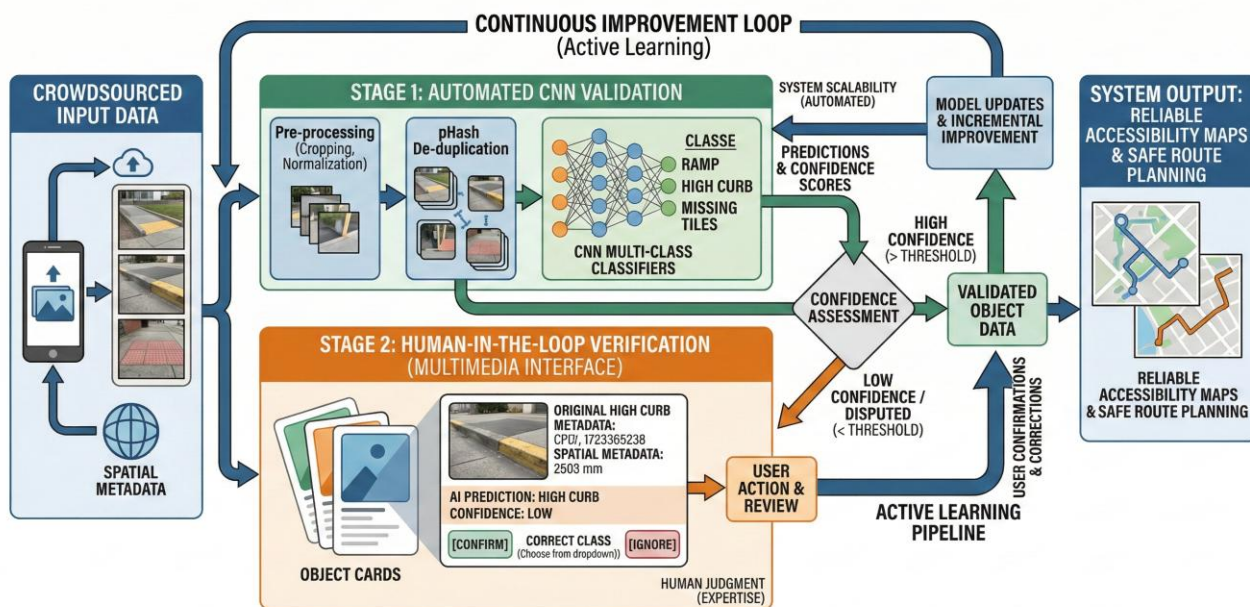


Рисунок 1 – Схема реалізації підходу Human-in-the-Loop у системі валідації

Експертні оцінки, отримані з використанням мультимедійного інтерфейсу карток, формують безперервний цикл зворотного зв'язку за

принципом активного навчання (Active Learning). Коли користувач підтверджує або виправляє автоматичну класифікацію, цей прецедент зберігається та розширює валідований набір даних. Ефективність такого підходу підтверджується сучасними дослідженнями у сфері комп'ютерного зору: впровадження парадигми Human-in-the-Loop дозволяє суттєво зменшити кількість помилок моделі. Ефективність такого підходу підтверджується сучасними дослідженнями у сфері комп'ютерного зору. Зокрема, роботи, присвячені активному навчанню та Human-in-the-Loop підходам, демонструють покращення якості класифікації та зменшення потреби у суцільній ручній розмітці даних за рахунок селективної перевірки лише невпевнених або складних випадків [3, 4].

Отже, запропонований гібридний підхід є оптимальним концептуальним рішенням для систем краудсорсингового мапінгу доступності міської інфраструктури. Він вирішує ключову проблему таких платформ, а саме пошук балансу між автоматизованою обробкою великих масивів візуальних даних та збереженням високої достовірності інформації. Делегуючи початкову класифікацію нейронній мережі та залишаючи вирішення спірних випадків людині, в такому випадку система не лише гарантує безпеку кінцевих маршрутів для маломобільних груп населення, але й генерує унікальний, стійкий до аномалій набір просторових даних для подальшого аналізу міського середовища.

Список використаних джерел:

1. Xingjiao Wu, L. Xiao, Y. Sun, J. Zhang, T. Ma, and L. He, "A Survey of Human-in-the-loop for Machine Learning," *arXiv preprint arXiv:2108.00941*, Aug. 2021. Available: <https://arxiv.org/abs/2108.00941>
2. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). <https://doi.org/10.1109/CVPR.2009.5206848>
3. Abdelwahab, A., Albahli, M. A., & Alshehri, S. A. (2023). Active deep learning for image classification. *Electronics*, 13(1). <https://doi.org/10.3390/electronics13010169>
4. Sarah E. Huebner, Meredith S. Palmer, Craig Packer (2024). Evaluating machine learning for wildlife image classification in Snapshot Safari. *Citizen Science: Theory and Practice*, 9(1). <https://doi.org/10.5334/cstp.752>