УДК 004.8, 004.67, 656.1/.5

КОМП'ЮТЕРНИЙ ЗІР ЯК СКЛАДОВА ЧАСТИНА СИСТЕМИ КЕРУВАННЯ ІНТЕЛЕКТУАЛЬНИМИ ТРАНСПОРТНИМИ ЗАСОБАМИ

Горобій П.С.[1]; Олійник А.О., д.т.н, проф[2]
[1,2] Національний Університет «Запорізька політехніка»
[1,2] Україна, Запоріжжя
[1] ORCID: 0009-0007-5609-6784; [2] ORCID: 0000-0002-6740-6078

Анотація. *Розглядається проблематика використання комп'ютерного зору, як складової частини комплексної системи керування інтелектуальними транспортними засобами.*

Ключові слова: *комп'ютерний зір; інтелектуальний транспортний засіб; транспортна інфраструктура.*

Вступна частина. Впровадження безпечного і екологічного наземного транспорту є однією з умов розв'язання задачі поліпшення якості життя мешканців густонаселених міст. Збільшення кількості наземного транспорту в містах закономірно породжує проблеми пов'язані із ускладненням використанням спільного міського простору: затори, нестача паркувальних місць, зниження доступності міської інфраструктури (заклади охорони здоров'я, заклади освіти, рекреаційні зони, тощо), збільшення кількості дорожньо-транспортних пригод (ДТП) та погіршення екологічного стану. Частина цих проблем може бути повністю вирішена лише на етапі планування транспортної інфраструктури міста. Але зниження кількості ДТП та раціональне використання ресурсів (пальне, електричної енергії, час в дорозі) можна досягти адаптувавши наземний транспорт до концепції заміни людини-водія цільовим апаратно-програмний комплексом керування інтелектуальним транспортним засобом (ІТЗ). Під здатністю до керування слід розуміти скоординоване виконання встановленого набору операцій. Саме на цьому етапі швидкість реакції і точність виконання відповідних дій інтелектуальним апаратно-програмним комплексом керування ІТЗ ліквідує вплив людського фактору (фізичні вади, втома або керування у нетверезому стані, особистісні риси характеру, навички водіння та інші фактори) та демонструє більш прийнятні результати, ніж людина-водій [1]. Також такі системи здатні значно покращити обрання оптимального маршруту та режиму руху (прискорення, гальмування, випередження і обгони на вільних ділянках маршруту), що позитивно вплине на зменшення вуглецевого сліду ІТЗ. Щоб забезпечити функціонування ІТЗ у режимі детекції і подолання характерних у міському середовищі динамічних перешкод необхідне отримання візуальної інформації (зображення та відео) за допомогою візуальних датчиків та детекторів, як елемент апаратного забезпечення. І системи комп'ютерного зору (КЗ), як елемент програмного забезпечення, які дозволяють машинам отримувати значущу інформацію з цифрових зображень, відео та інших візуальних даних, а також діяти на основі цієї інформації [2].

Мета дослідження. Огляд основних досягнень та проблем у впровадженні КЗ у системах управління ІТЗ в умовах густонаселеного міста.

Основна частина. Основною метою використання КЗ у ІТЗ є миттєве надання програмному комплексу ситуаційної обізнаності однозначної інформації щодо фізичних об'єктів довкола транспортного засобу, а також щодо характеру їх руху. Отримання такої кодифікованої інформації відбувається у процесі відокремлення та класифікації значущих груп пікселів на зображенні. Основні досягнення останніх років у цій царені зумовлені застосуванням передових методологій, таких як згорткові нейронні мережі (CNN), моделі візуальних трансформаторів (ViTs), генеративно-суперницькі мережі (GAN) та візуально-лінгвістичні моделі (VLM). Одним з найвідоміших досліджень є метод YOLO (You Only Look Once) який визначається своєю ефективністю і точністю детекції об'єктів у реальному часі [3].

Також можна відзначити, R-CNN (регіональна згорткова нейронна мережа) [4], що об'єднує регіональні мережі пропозицій із глибоким навчанням, значно покращуючи ефективність виявлення. Семантична сегментація передбачає класифікацію кожного пікселя зображення, щоб зрозуміти весь контекст сцени.

Приклади роботи включають DeepLab, розроблену в Google Research, яка використовує глибокі згорткові мережі та складне згортання для дрібнозернистих просторових ієрархій [5], і SegNet, призначену для попіксельної семантичної сегментації при економії пам'яті та обчислювальних ресурсів, що робить її практичною для програм реального часу [6].

Помітні досягнення включають EfficientNet від команди Google Brain, який використовує наскрізний підхід до масштабування для досягнення найсучасніших результатів з точки зору ефективності та точності [7], і Vision Transformers (ViTs), які демонструють перспективу масштабування та фіксації довгострокових залежностей [8].

Незалежно від використаного методу обробки візуальних сигналів швидкість, якість та повнота класифікованої інформації має критично важливе значення для прийняття комплексом керування якісних рішень щодо порядку сигналів які будуть відправлені на елементи керування ІТЗ.

На заваді вичерпної ситуаційної обізнаності систем керування ІТЗ зазвичай стають фактори, які можна розділити на наступні категорії [9]:

1. **Проблеми даних.**

1.1 ***об'єм даних*** - візуальні дані характеризуються великим об'ємом. Разом з обмеженою прохідною здатністю інформаційних каналів та затримкою у обробці значних об'ємів інформації цей фактор може стати вузьким місцем систем КЗ як на етапі навчання CNN так і на етапі практичного застосування;

1.2 ***якість (релевантність) навчальних даних*** - недостатня кількість якісних маркованих вхідних даних для навчання CNN. Спостерігається брак навчальних даних з певними групами специфічних транспортних засобів та дорожніх об'єктів. Також проблематично отримати візуальні сигнали у достатній кількості для предаварійних та аварійних сценаріїв;

1.3 ***конфіденційність*** - важливий етично-правовий аспект систем збору і обробки візуальних даних. Збереження та розповсюдження даних з обличчями людей та номерних знаків несе потенційні ризики втручання в особистий простір містян.

2. **Проблеми моделей.**

2.1 ***складність*** - програмні комплекси, засновані на моделях CNN, мають високу складність та потребують тривалих проміжків часу для навчання (від тижнів до місяців). А також значної обчислювальної потужності. Для практичного використання вже навченої моделі ця проблема також дуже актуальна. Існуючі мобільні апаратні рішення для обчислювальної складової систем КЗ дають задовільні результати, які поки що не відповідають критеріям стабільної безвідмовної роботи;

2.2 ***брак прозорості*** - робота програмних комплексів, заснованих на моделях CNN, базуються на співвідношеннях мільярдів параметрів, тому її можна дослідити лише за допомогою аналізу вихідних даних. Іншими словами такі системи мають слабо-передбачуваний характер та низьку пояснювальність вихідних сигналів;

2.3 ***універсальність і портованість*** - більшість алгоритмів навчання сильно спираються на незалежне та ідентично розподілене навчання на основі припущень під час тестування даних, що часто порушується на практиці через зсув домену. Узагальнення домену спрямоване на узагальнення моделей на нові домени без інформації про цільовий розподіл під час навчання;

2.4 ***тестування у реальних умовах*** - хоча показники релевантності вихідних сигналів покращуються зі збільшенням складності моделі та навчальних даних, тестування з реальними спотвореннями призводить до низької та дуже варіабельної продуктивності, яка сильно залежить від випадкових початкових значень, що використовуються для ініціалізації навчання.

3. **Комплексні фактори навколишнього середовища.**

3.1 ***освітлення*** - зміна часу доби, зміна пори року та погодні умови перетворюють поле зору візуальних датчиків у надзвичайно динамічну сцену з високою варіативністю, та можуть призводити до таких проблем як недостатнє експонування, розмиття та часткове перекриття. Також на об’єктивах камер може накопичуватися бруд, пил та волога, що спотворює вхідний візуальний сигнал. Ці фактори здійснюють вплив як на етапі навчання так і на етапі практичного використання;

3.2 ***оклюзія*** - цільові об'єкти можуть бути частково перекриті статичними об'єктами, такими як будівлі та ліхтарні стовпи. Динамічні об'єкти, такі як транспортні засоби, що рухаються, або інші учасники дорожнього руху, можуть перекривати один одного, наприклад, у натовпах;

3.3 ***хаотичність середовища*** - надзвичайна кількість різноманітних об’єктів, що наповнюють міський простір та впливають або не впливають на режим руху на конкретній ділянці, розмаїття кольорів, типів і форм, а також непередбачуваність унікальних сценаріїв розвитку подій, включно з випадковим або навмисним порушенням правил руху, ускладнюють відокремлення типових об’єктів на зображеннях та відео.

Висновки. Комп’ютерний зір відіграє значну роль системі керування інтелектуальними транспортними засобами. Зважене і комплексне вирішення всіх проблем детекції та класифікації об’єктів, які є невід’ємною частиною міських транспортних систем майбутнього, надасть у результаті підвищення безпеки та позитивного економічного ефекту в густонаселених містах.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Baballe, M. A. (2024). Self-Driving Car Pros. https://doi.org/10.5281/zenodo.10535786

[2] Sharma, V., Gupta, M., Kumar, A. (2021). Video Processing Using Deep Learning Techniques: A Systematic Literature Review. https://doi.org/10.1109/ACCESS.2021.3118541

[3] Mahaur, B., Mishra, K. (2023). Small-object detection based on YOLOv5 in autonomous driving systems. https://doi.org/10.1016/j.patrec.2023.03.009

[4] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

[5] Feldsar, B., Mayer, R., Rauber, A. (2023). Detecting Adversarial Examples Using Surrogate Models. https://doi.org/10.3390/make5040087

[6] Badrinarayanan, V., Kendall, A., Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. https://doi.org/10.1109/TPAMI.2016.2644615

[7] Hu, S., Liu, J., Kang, Z. (2021). DeepLabV3+/Efficientnet Hybrid Network-Based Scene Area Judgment for the Mars Unmanned Vehicle System. https://doi.org/10.3390/s21238136

[8] Zhai, X., Kolesnikov, A., Houlsby, N., & Beyer, L. (2022). Scaling vision transformers. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12104-12113).

[9] Azfar, T., Li, J., Yu, H., Cheu, R.L. (2022). Deep Learning based Computer Vision Methods for Complex Traffic Environments Perception: A Review. https://doi.org/10.48550/arXiv.2211.05120

Horobii P.S., Oliinyk A.O.

COMPUTER VISION AS A COMPONENT OF THE INTELLIGENT VEHICLE CONTROL SYSTEM

Abstract. *The research examines the issue of using computer vision as a component of a comprehensive intelligent vehicle control system.*

Keywords: *computer vision; intelligent vehicle; transport infrastructure.*