

МЕТОДИ КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ РОЗПІЗНАВАННЯ ДОРОЖНІХ ЗНАКІВ У СИСТЕМАХ ІНТЕЛЕКТУАЛЬНОГО ТРАНСПОРТУ

Казаров Д.Ю.

e-mail: dmytro.kazarov@nure.ua

Науковий керівник – канд. техн. наук, доц. Любченко В.А.
Харківський національний університет радіоелектроніки, каф. ІНФ
м. Харків, Україна

Traffic sign recognition is an important task in modern intelligent transportation systems and autonomous vehicles. Accurate detection and classification of road signs allow vehicles to correctly interpret traffic rules and improve road safety. This paper considers computer vision methods used for solving the problem of traffic sign recognition. The general pipeline of such systems includes image acquisition, preprocessing, sign detection, feature extraction and classification. Traditional approaches based on color and shape analysis are widely used for detecting candidate regions of road signs. Feature extraction methods allow systems to describe visual characteristics of detected objects. Machine learning algorithms are then applied for classification of traffic signs. Recently, deep learning approaches, especially convolutional neural networks, have significantly improved recognition accuracy. These methods allow systems to work reliably even in complex real-world conditions.

Автоматичне розпізнавання дорожніх знаків є ключовою функцією інтелектуальних транспортних систем та систем активної безпеки водія (ADAS – Advanced Driver Assistance Systems). Коректна ідентифікація знаків обмеження швидкості, заборонних і попереджувальних знаків дозволяє транспортному засобу адаптувати поведінку в реальному часі, що безпосередньо впливає на зниження аварійності. Метою роботи є огляд і порівняння методів комп'ютерного зору для вирішення цієї задачі.

Узагальнена система включає чотири етапи: захоплення зображення камерою, попередня обробка, виявлення областей-кандидатів та класифікація знака. На етапі попередньої обробки застосовується фільтрація шуму та перетворення у кольоровий простір HSV (Hue-Saturation-Value – тон, насиченість, яскравість), стійкіший до змін освітлення порівняно з RGB [1]. Узагальнену структуру системи наведено на рисунку 1.



Рисунок 1 – Блок-схема системи розпізнавання дорожніх знаків

Одним із найпоширеніших класичних підходів є поєднання дескриптора HOG (Histogram of Oriented Gradients – гістограма орієнтованих градієнтів) з класифікатором на основі методу опорних векторів (SVM – Support Vector Machine). Зображення розбивається на клітинки розміром 8x8 пікселів, у кожній з яких обчислюються градієнти яскравості та будується гістограма їх орієнтацій з 9 бінами в діапазоні 0°–180°. Магнітуда та орієнтація градієнта визначається так [2]:

$$|G| = \sqrt{G_x^2 + G_y^2}, \theta = \arctan\left(\frac{G_y}{G_x}\right)$$

де: $G_x = I(x+1, y) - I(x-1, y)$, $G_y = I(x, y+1) - I(x, y-1)$

Отриманий вектор HOG-ознак передається у класифікатор SVM з RBF-ядром. Метод демонструє точність 95-97% на датасеті GTSRB (German Traffic Sign Recognition Benchmark) при помірних обчислювальних витратах, що робить його придатним для вбудованих систем з обмеженими ресурсами [1].

Сучасні системи розпізнання дорожніх знаків використовують згорткові нейронні мережі (CNN – Convolutional Neural Networks), які автоматично навчаються виділяти ієрархічні ознаки без ручного проєктування дескрипторів. Базова операція згортки визначається як [3]:

$$S(i, j) = \sum_m \sum_n I(i + m, j + n) * K(m, n)$$

Після кожного згорткового шару застосовується функція активації ReLU (Rectified Linear Unit) та шар підвибірки (max pooling), що зменшує розмірність і підвищує інваріантність до зсувів об'єкта. Для виявлення та класифікації в реальному часі застосовується архітектури типу YOLO (You Only Look Once), які обробляють зображення за одне проходження мережі зі швидкістю 30-100 кадрів на секунду. На датасеті GTSRB CNN-моделі досягають точності 98-99% перевищуючи класичні підходи, особливо в умовах поганого освітлення та часткового перекриття знаків [3]. Приклад архітектури CNN наведено на рисунку 2.

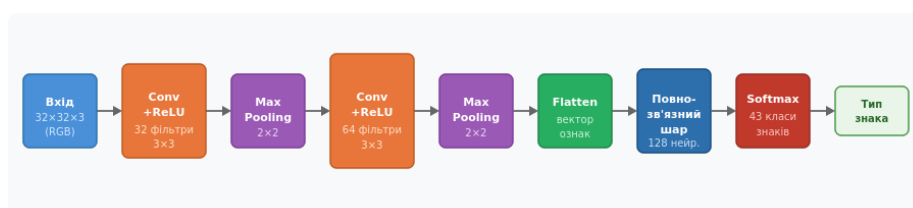


Рисунок 2 – Архітектура CNN для класифікації дорожніх знаків

Порівняння розглянутих методів показує, що підхід на основі HOG+SVM забезпечує точність 95-97% при помірних апаратних вимогах і є придатним для вбудованих пристроїв. CNN-архітектури, зокрема YOLO, досягають точності 98-99% з обробкою в реальному часі, однак потребують значно більших обчислювальних ресурсів. Найпростіший підхід на основі аналізу кольору та форми у просторі HSV поступається обом методам за точністю (~85%), проте залишається актуальним для систем із жорсткими обмеженнями за потужністю.

Таким чином, вибір методу визначається вимогами конкретної системи щодо балансу між точністю, швидкістю та апаратними обмеженнями. Метод HOG+SVM залишається актуальним для вбудованих пристроїв, тоді як CNN-архітектури є пріоритетним вибором для систем, де критична максимальна точність розпізнавання. Перспективним напрямом є оптимізація CNN-моделей для розгортання на периферійних пристроях у складі інтелектуальних транспортних систем.

Список використаних джерел:

1. Stallkamp J., Schlipsing M., Salmen J., Igel C. Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition. *Neural Networks*. 2012. Vol. 32. P. 323-332. URL: <https://doi.org/10.1016/j.neunet.2012.02.016>.
2. Dalal N., Triggs B. Histograms of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2005. P. 886-893. URL: <https://doi.org/10.1109/CVPR.2005.177>
3. Lecun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. 1998. Vol. 86, no. 11. P. 2278-2324. URL: <https://doi.org/10.1109/5.726791>.
4. Lyubchenko V., Veretelnyk K., Kots P., Lyashenko V. Digital image segmentation procedure as an example of an NP-problem. *Multidisciplinary Journal of Science and Technology*. 2024. Vol. 4, no. 4. P. 170–177. URL: <https://openarchive.nure.ua/handle/document/26162>.