



А. Є. Батюк, І. Ю. Опалінський

Національний університет "Львівська політехніка", м. Львів, Україна

## ПОРІВНЯННЯ ЕФЕКТИВНОСТІ ЗАСТОСУВАННЯ МОДЕЛЕЙ КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ КОНТРОЛЮ ПАРКУВАННЯ

Проаналізовано сучасні методи комп'ютерного зору, які використовують для автоматизованого контролю паркування. Досліджено особливості функціонування глибоких нейронних мереж, зокрема моделей YOLO, Faster R-CNN, SSD, EfficientDet, Mask R-CNN і DETR. Встановлено, що продуктивність цих моделей залежить від таких параметрів, як швидкість оброблення даних, точність виявлення об'єктів та ефективність використання ресурсів. Проведено порівняння моделей на підставі експериментальних даних, що враховують реальні умови функціонування паркувальних систем. Визначено, що одностайні моделі (YOLO, SSD) забезпечують значну швидкість оброблення відеопотоку, що робить їх придатними для застосування в режимі реального часу. Водночас, двоетапні моделі (Faster R-CNN, Mask R-CNN) мають вищу точність виявлення об'єктів, особливо у випадках, коли необхідне розпізнавання дрібних або частково перекритих транспортних засобів. Запропоновано використання трансформаторних архітектур (моделей DETR і RT-DETR) для підвищення точності детекції об'єктів без потреби використання немаксимального подавлення NMS (англ. *Non-Maximum Suppression*). Досліджено можливість удосконалення автоматизованих систем паркування шляхом застосування гібридних підходів, що комбінують одну та двоетапні моделі. Виявлено, що такі підходи дають змогу знизити затримки процедури оброблення даних і підвищити надійність виявлення транспортних засобів у складних умовах. Розроблено методику вибору оптимальної моделі комп'ютерного зору для автоматизованих паркувальних систем залежно від умов експлуатації. Наведено рекомендації щодо вибору моделей для великих відкритих паркінгів, підземних стоянок, мобільних застосунків і систем з обмеженими ресурсами. Визначено перспективні напрями проведення подальших досліджень, зокрема інтеграцію адаптивних трансформаторних архітектур у процес виявлення транспортних засобів для підвищення точності виявлення об'єктів та швидкості оброблення відео в реальному часі. Оцінено практичне застосування досліджених моделей у реальних паркувальних умовах. З'ясовано, що вибір оптимальної моделі залежить від конкретного сценарію використання та вимог до точності і швидкодії. Встановлено, що подальший розвиток автоматизованих систем контролю паркування має орієнтуватися на гібридні та адаптивні підходи, що поєднують переваги різних архітектур.

**Ключові слова:** нейронні мережі; автоматизований контроль; методи розпізнавання зображень; глибоке навчання; алгоритми виявлення об'єктів; швидкість оброблення відеоданих; удосконалення моделей.

### Вступ / Introduction

Сучасний розвиток технологій штучного інтелекту, зокрема глибокого навчання та машинного зору, значно покращив можливості автоматизованих систем контролю паркування. Проте, вибір оптимальної моделі комп'ютерного зору залишається відкритою проблемою, оскільки наявні рекомендації та аргументовані рішення мають компроміси між швидкістю оброблення відеоданих, точністю прогнозування та ресурсовитратністю.

Методи машинного зору активно досліджують провідні науковці, такі як Йозеф Редмон (англ. *Joseph Redmon*), Росс Гіршик (англ. *Ross Girshick*), Каймінг Хе (англ. *Kaiming He*) та Шаоцін Жень (англ. *Shaoqing Ren*), які розробили низку популярних архітектур виявлення об'єктів, зокрема моделі YOLO (англ. *You Only Look Once*), Faster R-CNN (англ. *Region-based Convolutional Neural Networks*) та Mask R-CNN. Останні дослідження демонструють ефективність трансформатор-

них архітектур, таких як моделі DETR (англ. *DEtection TRansformer*) та RT-DETR (англ. *Real-Time DETR*), які можуть покращити продуктивність розпізнавання об'єктів у складних умовах.

Попри значний прогрес у цій сфері, актуальною є проблема підбору моделі машинного зору для конкретних умов експлуатації. Одностайні моделі, такі як YOLO та SSD, забезпечують значну швидкість оброблення даних, але мають нижчу точність виявлення об'єктів у складних сценах [21]. Двоетапні моделі, такі як Faster R-CNN, забезпечують кращу точність виявлення об'єктів, але є більш ресурсовитратними [23, 8]. Недостатньо досліджено ефективність гібридних моделей та їхню адаптацію до реальних умов паркування, що потребує подальших досліджень.

**Об'єкт дослідження** – застосування моделей комп'ютерного зору для автоматизованих систем контролю паркування.

### Інформація про авторів:

**Батюк Анатолій Євгенович**, канд. техн. наук, доцент, кафедра автоматизованих систем управління.

Email: [anatolii.y.batiuk@lpnu.ua](mailto:anatolii.y.batiuk@lpnu.ua); <https://orcid.org/0000-0001-7650-7383>

**Опалінський Іван Юрійович**, магістр, кафедра автоматизованих систем управління. Email: [ivanopaliskij@gmail.com](mailto:ivanopaliskij@gmail.com)

**Цитування за ДСТУ:** Батюк А. Є., Опалінський І. Ю. Порівняння ефективності застосування моделей комп'ютерного зору для автоматизованої системи контролю паркування. Науковий вісник НЛТУ України. 2025, т. 35, № 1. С. 119–126.

**Citation APA:** Batiuk, A. Ye., & Opalinskiy, I. Yu. (2025). Comparison of the efficiency of different computer vision models for the automated parking control system. *Scientific Bulletin of UNFU*, 35(1), 119–126. <https://doi.org/10.36930/40350115>

*Предмет дослідження* – методи і засоби встановлення ефективності використання різних моделей комп'ютерного зору для виявлення транспортних засобів, що дасть змогу порівняти їхню точність виявлення об'єктів, швидкість оброблення даних і економію ресурсів у реальних умовах.

*Мета роботи* – проаналізувати наявні моделі комп'ютерного зору для автоматизованих систем контролю паркування, що дасть можливість порівняти ефективність їх застосування, визначити основні переваги та недоліки, а також окреслити перспективні напрями проведення подальших досліджень.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- проаналізувати найпоширеніші моделі комп'ютерного зору, які використовують в автоматизованих системах контролю паркування, зокрема моделей YOLO, Faster R-CNN, SSD, EfficientDet та Mask R-CNN, що дасть змогу виявити їхні переваги і недоліки у контексті використання для задач автоматизованого паркування.
- порівняти ефективність моделей у режимі реального часу, зокрема щодо швидкості оброблення даних, точності виявлення об'єктів та ресурсовитратності, що дасть можливість визначити найпридатніші для різних умов експлуатації.

У наступних розділах буде проаналізовано ключові моделі комп'ютерного зору, таких як YOLO, Faster R-CNN, SSD, EfficientDet та Mask R-CNN, та проаналізовано їхні можливості та обмеження в контексті автоматизованих систем контролю паркування.

**Аналіз останніх досліджень та публікацій.** Розвиток автоматизованих систем контролю паркування, заснованих на комп'ютерному зорі, залишається актуальним напрямом досліджень. Одним з ключових особливостей цього напрямку є вибір моделі глибокого навчання, яка дає можливість ефективно виявляти транспортні засоби на підставі відеоаналізу.

Перші дослідження в галузі використання комп'ютерного зору для автоматизованих систем контролю паркування фокусувалися на розробленні методів виявлення транспортних засобів на підставі аналізу зображень та відео. Однією з найбільш ранніх та популярних моделей для цієї мети стала модель YOLO (англ. *You Only Look Once*), яку було наведено у роботі [22]. Модель YOLO показала значний прорив у швидкості виявлення об'єктів завдяки одноетапному підходу, де зображення обробляють тільки один раз. Незважаючи на значну швидкість оброблення даних, початкові версії моделі YOLO страждали від меншої точності виявлення об'єктів у складних сценаріях їхнього перебування з дрібними або перекритими об'єктами.

Подальші дослідження спрямовувалися на підвищення точності виявлення об'єктів за допомогою двоетапних моделей, таких як Faster R-CNN, яку розробили Shaoqing Ren, Kaiming He, Ross Girshick та Jian Sun [23]. У роботі показано, що двоетапний підхід дає змогу досягти вищої точності виявлення об'єктів у роботі зі складними сценами, де важливо точно виявляти транспортні засоби навіть в умовах часткового перекриття. Проте це збільшує обчислювальну складність і знижує швидкість оброблення даних, що робить двоетапні моделі менш придатними для роботи в реальному часі [8]. Дослідження [26] також порівнює Faster R-CNN та Mask R-CNN у завданні виявлення тріщин на дорожньому покритті, що може бути аналогічним випадком

для виявлення транспортних засобів у складних умовах. Автори відзначають, що модель Mask R-CNN має переваги у точності розпізнавання, однак використовує більше обчислювальних ресурсів.

Окрім Faster R-CNN, альтернативна модель – SSD (англ. *Single Shot MultiBox Detector*), запропонована в роботі [18], намагається поєднати переваги одноетапного та двоетапного підходів, забезпечуючи оптимальний баланс між швидкістю та точністю. Вона стала популярною для завдань, де важливо зберегти продуктивність у реальному часі, при цьому забезпечуючи достатню точність для виявлення транспортних засобів на парковках. Дослідження [29] аналізує ефективність Faster R-CNN та YOLOv4 у задачах розпізнавання дорожніх знаків, що демонструє високу продуктивність YOLOv4 у реальному часі.

Окрему увагу варто приділити моделі YOLO та її вдосконаленням. У роботі [4] наведено огляд архітектурних наступників YOLO, враховуючи покращення моделей YOLOv4 та YOLOv5, а також їх використання у реальних застосуваннях. Основні висновки: YOLOv4 демонструє значну швидкість і точність для загальних завдань виявлення об'єктів, що робить її ідеальним кандидатом для реального часу; YOLOv5 продовжує вдосконалювати модель, знижуючи затримки та підвищуючи ефективність розпізнавання дрібних об'єктів.

Робота [7] розглядає використання YOLO-v2 для виявлення малих транспортних засобів в умовах інтелектуального паркування. Автори пропонують удосконалення моделі шляхом додавання резидуальних модулів та покращення глибини мережі для збільшення точності.

Отже, аналіз наявних досліджень показує, що розвиток комп'ютерного зору для автоматизованих систем контролю паркування залишається перспективною галуззю, яка вимагає подальших досліджень та удосконалень для підвищення ефективності використання ресурсів та точності виявлення об'єктів.

## Результати дослідження та їх обговорення / Research results and their discussion

Дослідження комп'ютерного зору для систем контролю паркування розпочнемо з можливості використання одно- і двоетапних моделей, з підвищення їхньої продуктивності, з аналізу проблеми належної точності виявлення об'єктів у складних сценаріях їхнього перебування та з дрібними перекритими об'єктами, а також проаналізуємо можливості подальшого вдосконалення таких моделей та комбіновані підходи до їхнього використання.

**1. Використання одноетапних моделей для контролю паркування.** Одноетапні моделі, такі як YOLO (англ. *You Only Look Once*) та SSD (англ. *Single Shot MultiBox Detector*), набули великої популярності через свою високу продуктивність і здатність працювати в режимі реального часу. Архітектура моделі YOLO є простішою порівняно з іншими моделями, оскільки зображення розбивається на сітку, і кожна клітинка сітки відповідає за передбачення об'єкта та його положення.

Основна перевага моделі YOLO полягає у значній швидкості оброблення даних: вона здатна обробляти до 45 кадрів за секунду (FPS), що робить її придатною для використання у реальних системах контролю паркування, де важлива швидка реакція на зміни [22]. Проте, модель YOLO також має деякі недоліки. Основним з них є її менша точність виявлення об'єктів порівняно з дво-

етапними моделями, особливо в разі виявлення дрібних об'єктів або об'єктів, що частково перекриваються.

Відомо, що модель YOLO може мати проблеми з виявленням малих транспортних засобів, таких як мотоцикли або велосипеди, на паркувальних майданчиках з великою кількістю автомобілів. За статистикою, точність моделі YOLO (міра AP, Average Precision) може змінюватись від 57 до 65 %, залежно від складності сцени [22]. Проте її продуктивність робить її привабливою для задач, де ключовою є швидкість оброблення даних, наприклад, для великих паркувальних майданчиків із великим потоком транспорту.

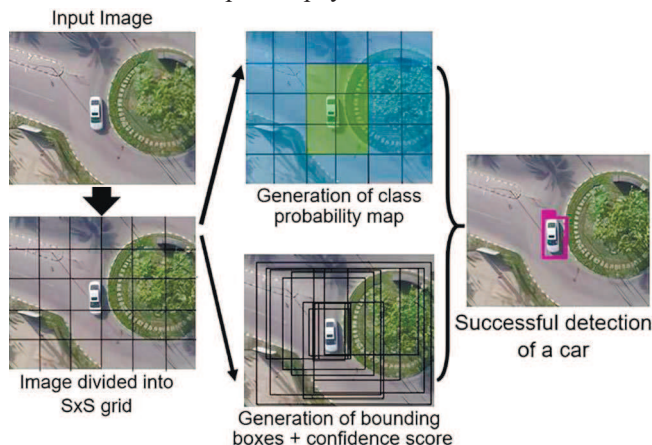


Рис. 1. Архітектура моделі YOLO / YOLO architecture [22]

Останні дослідження, зокрема робота моделі DETRs Beat YOLOs on Real-time Object Detection [32], показують, що DETR-моделі можуть конкурувати з моделями YOLO в реальному часі завдяки усуненню етапу NMS (англ. *Non-Maximum Suppression*). Це відкриває нові можливості для використання трансформаторних архітектур у системах автоматизованого контролю парку-

вання. Оскільки такі системи працюють у динамічному середовищі, швидкість оброблення відео та якість виявлення об'єктів є критично важливими показниками їхньої роботи. Використання моделі DETR дає можливість усунути можливі затримки, зумовлені етапом NMS, а значить потенційно підвищити ефективність розпізнавання транспортних засобів.

Модель SSD має збалансованішу архітектуру, яка використовує багаторівневі особливості нейронної мережі для виявлення об'єктів на різних масштабах зображення. SSD-модель налаштована під швидке оброблення зображення та детекцію об'єктів навіть у реальному часі, де вона забезпечує високу як точність, так і швидкість оброблення зображень. Це надає можливість моделі SSD ефективніше справлятися з виявленням об'єктів різних розмірів, забезпечуючи високу якість їхнього розпізнавання навіть у складних умовах середовища. Завдяки використанню спеціалізованих шарів, що обробляють інформацію на різних рівнях, модель SSD здатна точніше розпізнавати об'єкти різного масштабу, зберігаючи свою високу продуктивність. Особливо це корисно у випадках, коли потрібно розпізнавати об'єкти з різними контрастами розмірів або деталізації, що сприяє її універсальності в реальних умовах.

Як і модель YOLO, модель SSD є одноетапною моделлю, що дає змогу їй працювати в реальному часі, досягаючи продуктивності до 22 FPS. Така швидкість оброблення даних є важливою перевагою для багатьох застосувань, де тривалість реакції критично важлива. Модель SSD також має вищу точність виявлення об'єктів у випадках, де важливо працювати з об'єктами різного масштабу, що дає їй змогу стати хорошим вибором для задач, де необхідно враховувати широкий спектр варіацій у розмірах об'єктів [18].

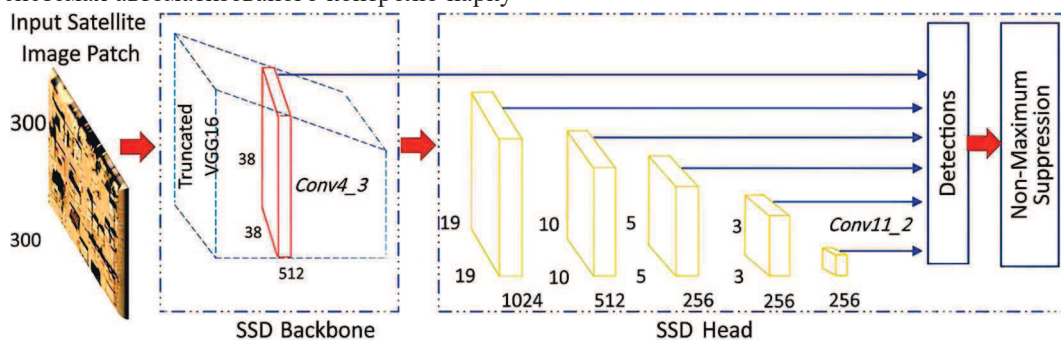


Рис. 2. Архітектура моделі SSD / SSD architecture [18]

Архітектурно модель SSD використовує метод багаторівневої детекції, аналізуючи зображення на декількох рівнях. Наприклад, великі об'єкти виявляються на нижчих рівнях, а дрібні – на вищих. Точність моделі SSD (AP) досягає 72 % на стандартних наборах даних для виявлення об'єктів, таких як COCO [18].

**2. Використання двоетапних моделей для підвищення точності виявлення об'єктів.** Двоетапні моделі, такі як Faster R-CNN і Mask R-CNN, є значно точнішими завдяки використанню регіонів інтересу (RoI) та складної архітектури, що містить два етапи оброблення зображення. Модель Faster R-CNN, наприклад, спочатку виділяє потенційні регіони інтересу на зображенні за допомогою регіон-пропонуєчої мережі RPN (англ. *Region Proposal Network*), а потім виконує точну класифікацію цих регіонів і визначення меж об'єктів [23]. Це надає можливість моделі забезпечувати належ-

ну точність виявлення об'єктів у складних сценах, де об'єкти можуть перекриватися або бути частково видимими. Архітектура моделі Faster R-CNN є складнішою і ресурсомісткою, що призводить до зниження швидкості оброблення даних (до 7 кадрів за секунду на високопродуктивних системах) [23].

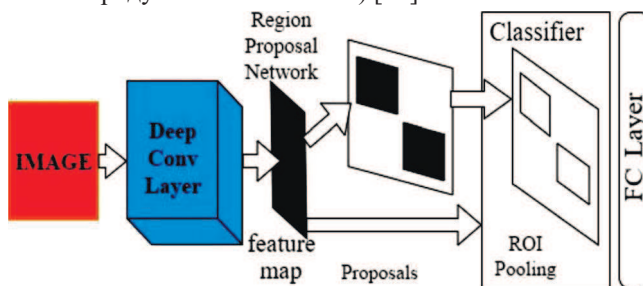


Рис. 3. Архітектура моделі Faster R-CNN / Faster R-CNN architecture [10]

Проте точність цієї моделі є однією з найвищих серед всіх моделей для виявлення об'єктів: на стандартних наборах даних COCO або PASCAL VOC, модель Faster R-CNN показує належну точність виявлення об'єктів до 84 % (AP), що робить її ідеальною для застосувань, де важлива належна точність виявлення об'єктів і де не критичні вимоги до оброблення в реальному часі.

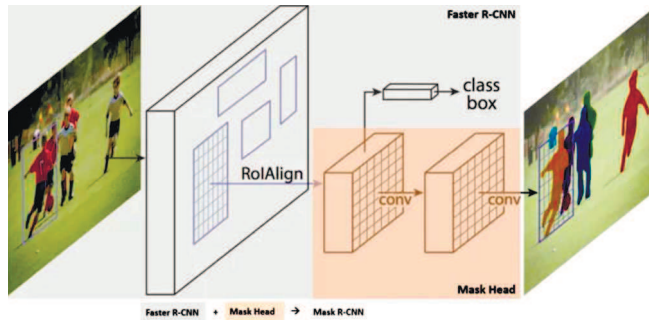


Рис. 4. Архітектура моделі Mask R-CNN / Mask R-CNN architecture [20]

Модель Mask R-CNN, яка є розширенням моделі Faster R-CNN, додає можливість сегментації об'єктів, що дає змогу моделі не тільки визначати контури об'єктів, але й виділяти кожен піксель, що належить до об'єкта. Це особливо корисно в завданнях, де важлива детальна сегментація, наприклад, для точного оцінювання розмірів транспортного засобу або візуального контролю на паркувальних майданчиках [8]. Точність моделі

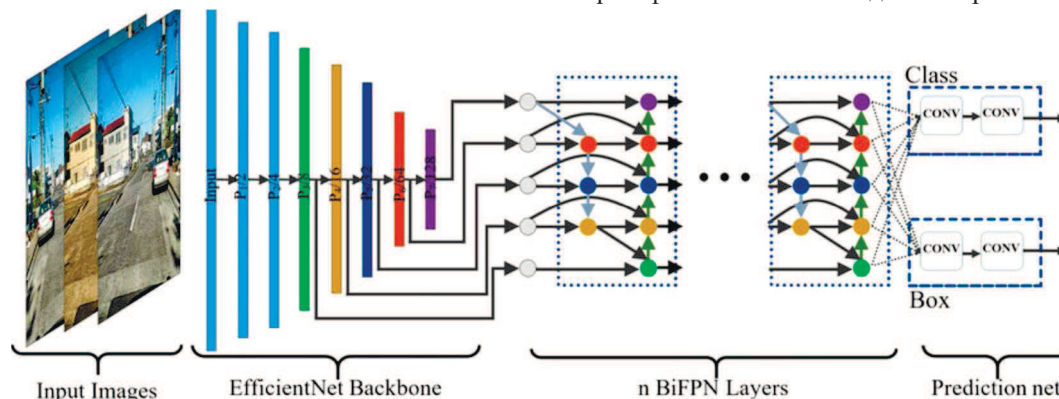


Рис. 5. Архітектура моделі EfficientDet / EfficientDet architecture [19]

Архітектура моделі EfficientDet базується на концепції комбінації різних масштабів зображення в єдину багатомасштабну модель, що надає можливість обробляти об'єкти різного розміру одночасно [24]. Це робить її дуже ефективною в завданнях виявлення об'єктів на паркувальних майданчиках з різними транспортними засобами (великими і малими). При цьому вона демонструє точність виявлення об'єктів на рівні 80 % (AP), що дещо нижче, ніж у двоетапних моделей, але значно перевищує продуктивність одноетапних підходів.

**4. Проблеми точності виявлення об'єктів у складних сценаріях їхнього перебування та з дрібними перекритими об'єктами.** Одноетапні моделі, такі як YOLO і SSD, мають загальні обмеження у точності виявлення об'єктів при виявленні дрібних або перекритих об'єктів. Проблема полягає в тому, що ці моделі намагаються обробляти всю сцену одночасно, що може призводити до втрати деталей на дрібних об'єктах [22]. Відомо, що точність моделі YOLO у роботі з дрібними перекритими об'єктами може знижуватися до 40-50 %, що є критичним для задач, де важливо виявляти навіть

Mask R-CNN може досягати 87 % у стандартних завданнях сегментації.

**3. Підвищення продуктивності виявлення об'єктів за допомогою моделі EfficientDet.** Вибір оптимальної архітектури нейронної мережі може бути складним завданням. Необхідно знайти баланс між високою точністю класифікації об'єктів та ефективним використанням ресурсів. Це містить налаштування глибини мережі, розміру фільтрів та інших параметрів. У 2019 році команда дослідників із Google AI представила вирішення цієї проблеми. Вони розробили серію архітектур моделей під назвою EfficientNet. Ці моделі відрізняються високим ступенем ефективності роботи та легко налаштовуються під різні завдання. Вони дають змогу класифікувати зображення з високою точністю, споживаючи мінімальну кількість ресурсів. Модель EfficientNet стало значним кроком уперед у розвитку нейронних мереж для класифікації зображень і продовжує бути актуальною ще й досі.

Модель EfficientDet надає оптимальний баланс між точністю виявлення об'єктів та ефективністю використання ресурсів завдяки використанню архітектури EfficientNet для зменшення кількості параметрів [24]. Ця модель відзначається відмінною продуктивністю на мобільних пристроях та в умовах обмежених обчислювальних ресурсів, що робить її привабливим вибором для систем контролю паркування, де критично важливими є енергоефективність та швидкість оброблення даних.

невеликі транспортні засоби, такі як мотоцикли або велосипеди.

Попри значну швидкість оброблення даних, моделі YOLO мають низку обмежень, які можуть впливати на ефективність у реальних умовах паркування [22]:

- потреба використання ETAПу NMS для видалення зайвих детекцій знижує продуктивність у високонавантажених сценах;
- висока чутливість до налаштувань порогових значень (IoU, Confidence) потребує модифікації для кожного конкретного середовища;
- складнощі з розпізнаванням об'єктів у випадках часткового перекриття та низького освітлення.

Враховуючи ці недоліки, актуальним є пошук альтернативних підходів, зокрема використання DETR-моделей, які здатні усувати залежність від ETAПу NMS та більш ефективно обробляти складні сцени.

Двоетапні моделі, такі як Faster R-CNN і Mask R-CNN, справляються з цими проблемами завдяки використанню регіонів інтересу та багаторівневому обробленню зображень [8, 23]. Їхня точність, у разі виявлення дрібних об'єктів, може досягати 80-85 %, що робить

їх незамінними для завдань, де важлива максимальна точність, наприклад, у багаторівневих паркінгах або в міських умовах, де транспортні засоби часто перекриваються один одним. На підставі проведених досліджень,

можна зробити висновок, що вибір моделі комп'ютерного зору для систем контролю паркування значно залежить від конкретних вимог до швидкості оброблення даних і точності виявлення об'єктів.

Табл. 1. Порівняння моделей / Comparison of models

| Модель       | Тип        | Переваги                                                                                                                        | Недоліки                                                                                                                          |
|--------------|------------|---------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------|
| YOLO         | Одноетапна | Значна швидкість оброблення даних.<br>Проста архітектура.<br>Підходить для реального часу.                                      | Менша точність виявлення об'єктів порівняно з двоетапними моделями.<br>Може мати проблеми з дрібними перекритими об'єктами.       |
| EfficientDet | Одноетапна | Баланс між точністю та швидкістю.<br>Ефективна архітектура для ресурсозалежних середовищ.<br>Підходить для мобільних пристроїв. | Трохи відстає від двоетапних моделей за точністю виявлення об'єктів.<br>Може мати проблеми з дуже дрібними перекритими об'єктами. |
| SSD          | Одноетапна | Хороший баланс між швидкістю та точністю.<br>Підтримка реального часу.                                                          | Менша точність виявлення об'єктів порівняно з двоетапними моделями.<br>Вразлива до перекриття об'єктів.                           |
| Faster R-CNN | Двоетапна  | Належна точність виявлення об'єктів.<br>Краще виявляє дрібні об'єкти.                                                           | Повільніша порівняно з одноетапними моделями.<br>Складніший процес навчання.                                                      |
| Mask R-CNN   | Двоетапна  | Належна точність виявлення об'єктів для сегментації.<br>Піксельна точність виявлення об'єктів для масок.                        | Найповільніша з наведених моделей.<br>Більш ресурсовитратна.                                                                      |

У роботах [18, 23, 24] наголошено на тому, що хоча двоетапні моделі забезпечують вищу точність виявлення об'єктів, одноетапні моделі є більш придатними для реального часу завдяки своїй швидкості оброблення даних. Моделі EfficientDet та SSD є оптимальними виборами для умов з обмеженими ресурсами, тоді як моделі Faster R-CNN та Mask R-CNN доцільно використовувати у випадках, коли критично важлива максимальна точність виявлення об'єктів [8, 23, 24].

5. **Можливості подальшого вдосконалення моделей та комбіновані підходи.** Для подальшого розвитку систем автоматизованого контролю паркування перспективними є гібридні підходи, що поєднують переваги одно- і двоетапних моделей [23, 24]. Наприклад, можна використовувати моделі YOLO або SSD для пер-

винного виявлення транспортних засобів в умовах реального часу, а потім застосовувати моделі Faster R-CNN або Mask R-CNN для детальнішої перевірки або виявлення у складних сценах, таких як перекриття транспортних засобів або виявлення дрібних об'єктів [8, 23, 24].

Гібридні системи можуть забезпечити не тільки значну швидкість оброблення даних, але й належну точність виявлення об'єктів у складних умовах, що робить їх перспективними для широкого впровадження в системи управління паркувальними майданчиками.

На підставі аналізу літератури та сучасних досліджень було розроблено рекомендації щодо вибору моделей залежно від специфічних умов їхньої експлуатації (табл. 2).

Табл. 2. Рекомендації щодо вибору моделі / Recommendations for choosing a model

| Умови використання                                                             | Рекомендовані моделі              | Переваги                                                                           |
|--------------------------------------------------------------------------------|-----------------------------------|------------------------------------------------------------------------------------|
| Великий відкритий паркінг із високим транспортним потоком                      | YOLOv8, SSD                       | Значна швидкість оброблення даних, оброблення великих обсягів трафіку              |
| Підземний або критий паркінг з низьким освітленням                             | RT-DETR, Faster R-CNN, Mask R-CNN | Стабільна робота в умовах слабкого освітлення                                      |
| Паркувальна зона з частковим перекриттям авто                                  | DETR, RT-DETR, Mask R-CNN         | Краща локалізація при складних ракурсах                                            |
| Мобільні та вбудовані системи з обмеженими ресурсами                           | EfficientDet, YOLO-NAS, Tiny YOLO | Низьке споживання ресурсів, ефективність на Edge AI                                |
| Паркувальна система з високими вимогами до точності виявлення об'єктів         | DINO, Deformable DETR, RT-DETR    | Максимальна точність виявлення об'єктів                                            |
| Реальний час, потрібна значна швидкість детекції                               | YOLOv8, SSD, YOLO-NAS             | Миттєве розпізнавання, мінімальна затримка                                         |
| Гібридний підхід для точності виявлення об'єктів та швидкості оброблення даних | YOLOv8 + DETR                     | Оптимальне співвідношення швидкості оброблення даних і точності виявлення об'єктів |

Грунтуючись на проведеному аналізі, можна виділити два основні напрями для подальшого розвитку [4]:

- подальша модифікація архітектури та алгоритмів моделей комп'ютерного зору для підвищення точності виявлення об'єктів та продуктивності одночасно, особливо в умовах обмежених ресурсів;
- використання гібридних підходів – комбінування одно- і двоетапних моделей може забезпечити оптимальне поєднання швидкості оброблення даних і точності виявлення об'єктів для контролю паркування в реальних умовах.

Незважаючи на значний прогрес у розробленні та застосуванні моделей комп'ютерного зору для виявлення об'єктів, проблема знаходження універсального рішення як щодо їхнього вибору, так і можливостей використання залишається відкритою.

Дослідження показало, що традиційні YOLO-моделі мають значні переваги у швидкості оброблення даних, однак їх ефективність обмежується необхідністю використання етапу NMS. Альтернативні підходи, такі як RT-DETR, демонструють кращі результати в умовах реального часу, надаючи можливість покращити точність виявлення об'єктів без додаткових витрат обчислювальних ресурсів.

З огляду на отримані результати, можна зробити висновок, що трансформаторні архітектури комп'ютерного зору можуть стати перспективним напрямом у розробленні систем автоматизованого паркування.

**Обговорення результатів дослідження.** У дослідженні [27], опублікованому на платформі Keylabs, було

проведено порівняльний аналіз моделей YOLOv8 та Faster R-CNN. Результати показали, що модель YOLOv8 перевершує модель Faster R-CNN як за точністю виявлення об'єктів, так і за швидкістю оброблення відеоданих. Зокрема, YOLOv8 досягла при метриці mAP@500.62 з затримкою на GPU 1.3 мс, тоді як модель Faster R-CNN показала при метриці mAP@500.41 з затримкою 54 мс. Ці дані узгоджуються з нашими результатами, де моделі YOLO також демонструє кращу продуктивність у реальному часі.

У роботі [25], опублікованій у журналі Sensors, було проведено порівняння моделей YOLO V2, Faster R-CNN та SSD для класифікації кутів третіх молярів на панорамних рентгенівських знімках. Дослідження показало, що моделі YOLO V2 та SSD демонструють вищу точність виявлення об'єктів та швидкість оброблення даних порівняно з моделлю Faster R-CNN, що підтверджує їхню придатність для застосувань, де важлива швидкість оброблення даних і точність виявлення об'єктів. Ці висновки збігаються з нашими спостереженнями щодо ефективності моделі YOLO та SSD у завданнях контролю паркування.

Дослідження [14] демонструє систему розпізнавання зайнятості паркувальних місць за допомогою комп'ютерного зору. Система використовує відеопотік з камер для відстеження зайнятості кожного паркувального місця, що дає можливість економити час водіїв та підвищувати комфорт і безпеку дорожнього руху. Цей підхід підтверджує актуальність використання моделей комп'ютерного зору в автоматизованих системах контролю паркування, що узгоджується з нашими результатами.

Стаття [28], яка опублікована на платформі Labelvisior, показує детальний аналіз та порівняння моделей YOLOv8 та Mask R-CNN. Згідно з дослідженнями, YOLOv8 відзначається високою швидкістю та точністю, що робить її придатною для застосувань у реальному часі. З іншого боку, модель Mask R-CNN забезпечує належну точність виявлення об'єктів, але з меншою швидкістю оброблення відеоданих, що може бути обмеженням для деяких застосувань. Ці висновки підтверджують наші результати щодо компромісу між швидкістю та точністю у виборі моделі для автоматизованих систем контролю паркування.

У дослідженні [30] було проведено навчання та тестування трьох моделей: YOLO, Faster R-CNN та SSD для задачі виявлення хмар. Результати показали, що всі три моделі досягли високих показників точності виявлення об'єктів, проте YOLO та SSD продемонстрували кращу швидкість оброблення даних, що є критичним для застосувань у реальному часі. Ці результати узгоджуються з нашими висновками щодо ефективності моделей YOLO та SSD у завданнях контролю паркування.

У дослідженні [27] проведено порівняння моделей YOLOv8 та Faster R-CNN для задачі виявлення об'єктів на супутникових знімках. Результати показали, що модель YOLOv8 перевершує модель Faster R-CNN як за точністю виявлення об'єктів, так і за швидкістю оброблення відеоданих, що робить її більш придатною для застосувань, де важлива оперативність.

У дослідженні [15] проведено порівняння моделей YOLO, Faster R-CNN та SSD для задачі виявлення дефектів на поверхнях. Результати показали, що модель, запропонована авторами, перевершує базову модель YOLOv5s на 6,3 % за точністю виявлення об'єктів, при

цьому зменшуючи кількість параметрів та обчислювальну складність. Це свідчить про потенціал використання покращених моделей глибокого навчання для завдань, що вимагають високої точності виявлення об'єктів за обмежених обчислювальних ресурсів.

Результати цього дослідження підтверджують наші висновки щодо компромісу між точністю та швидкістю різних моделей. Зокрема, одноетапні моделі, такі як YOLO та SSD, показують кращу швидкодію, що робить їх придатними для застосувань у реальному часі. Водночас, модель Faster R-CNN забезпечує вищу точність виявлення об'єктів, що важливо для завдань, де критично необхідне розпізнавання дрібних деталей. Аналогічні результати спостерігаються у нашому дослідженні, що підтверджує ефективність вибору конкретної архітектури залежно від вимог до продуктивності системи контролю паркування.

Отже, внаслідок виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

*Наукова новизна отриманих результатів дослідження* – удосконалено методику вибору моделей комп'ютерного зору для автоматизованих систем контролю паркування, яка, на відміну від наявних, поєднує адаптивні алгоритми глибокого навчання, трансформаторні архітектури та гібридні підходи до виявлення транспортних засобів, а також дає можливість підвищити точність їх виявлення та швидкість оброблення відеоданих у різних експлуатаційних умовах.

*Практична значущість результатів дослідження* – результати дослідження можна застосувати для розроблення автоматизованих систем контролю паркування, що використовують вдосконалені моделі комп'ютерного зору та глибокого навчання для виявлення транспортних засобів у режимі реального часу, а також можна впровадити у відеоаналітичні модулі інтелектуальних транспортних систем, що дасть можливість підвищити ефективність управління паркувальним простором, зменшити затримки в обробленні відеоданих і оптимізувати використання обчислювальних ресурсів у міських умовах та системах з обмеженими обчислювальними можливостями.

## Висновки / Conclusions

Проаналізовано наявні моделі комп'ютерного зору для автоматизованих систем контролю паркування, що дало можливість порівняти ефективність їх застосування, визначити основні переваги та недоліки, а також окреслити перспективні напрями подальших досліджень. За результатами проведеного дослідження можна зробити такі основні висновки:

- 1) Проаналізовано продуктивність найпоширеніших моделей детекції об'єктів, таких як YOLO, Faster R-CNN, SSD, EfficientDet та Mask R-CNN, з урахуванням їх точності виявлення об'єктів, швидкості оброблення даних і ефективності використання обчислювальних ресурсів.
- 2) З'ясовано, що одноетапні моделі, такі як YOLO та SSD, забезпечують значну швидкість оброблення відеоданих, що робить їх придатними для застосування в режимі реального часу, тоді як двоетапні моделі Faster R-CNN та Mask R-CNN демонструють вищу точність виявлення об'єктів, але значно поступаються за швидкістю.
- 3) Встановлено, що трансформаторні архітектури, зокрема моделі DETR та RT-DETR, можуть усунути потребу використання етапу NMS (англ. *Non-Maximum Suppres-*

sion), що потенційно підвищує якість розпізнавання транспортних засобів у складних паркувальних умовах із частковим перекриттям об'єктів.

- 4) Досліджено можливості використання гібридних моделей, що поєднують одно- та двоетапні архітектури, для покращення ефективності роботи автоматизованих паркувальних систем, що дає можливість досягти оптимального балансу між швидкістю та точністю.
- 5) Запропоновано методологію вибору моделі комп'ютерного зору залежно від специфічних умов експлуатації, що дає змогу адаптувати системи контролю паркування для роботи у великих відкритих зонах, підземних паркінгах та мобільних застосунках із обмеженими обчислювальними ресурсами.
- 6) Оцінено ефективність застосування моделей у реальних умовах їхньої експлуатації, яке було обґрунтовано теоретично, що підтвердило доцільність використання адаптивних і гібридних підходів для підвищення точності виявлення об'єктів та швидкості оброблення відеоданих.

## References

1. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv: 2004.10934. <https://doi.org/10.48550/arXiv.2004.10934>
2. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-End Object Detection with Transformers. In: European Conference on Computer Vision, (pp. 213–229). Springer, Cham. [https://doi.org/10.1007/978-3-030-58452-8\\_13](https://doi.org/10.1007/978-3-030-58452-8_13)
3. Dai, J., Li, Y., He, K., & Sun, J. (2016). R-FCN: Object Detection via Region-based Fully Convolutional Networks. In: Advances in Neural Information Processing Systems, (pp. 379–387). <https://doi.org/10.48550/arXiv.1605.06409>
4. Diwan, Tausif, G. Anirudh, and Jitendra V. Tembhurne. "Object detection using YOLO: Challenges, architectural successors, datasets and applications." multimedia Tools and Applications 82.6 (2023): 9243-9275. <https://doi.org/10.1007/s11042-022-13644-y>
5. Fu, C. Y., Liu, W., Ranga, A., Tyagi, A., & Berg, A. C. (2017). DSSD: Deconvolutional Single Shot Detector. arXiv preprint arXiv: 1701.06659. <https://doi.org/10.48550/arXiv.1701.06659>
6. Girshick, R. (2015). Fast R-CNN. In: *Proceedings of the IEEE International Conference on Computer Vision*, (pp. 1440–1448). <https://doi.org/10.1109/ICCV.2015.169>
7. Han, Xiaohong, Jun Chang, and Kaiyuan Wang. "Real-time object detection based on YOLO-v2 for tiny vehicle object." *Procedia Computer Science* 183 (2021): 61-72. <https://doi.org/10.1016/j.procs.2021.02.031>
8. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2020). Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 386–397. <https://doi.org/10.1109/TPAMI.2018.2844175>
9. Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A.,... & Murphy, K. (2017). Speed/Accuracy Trade-offs for Modern Convolutional Object Detectors. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 7310–7311). <https://doi.org/10.1109/CVPR.2017.351>
10. Jha, S., Dey, A., Kumar, R., & Kumar, V. (2019). A novel approach on visual question answering by parameter prediction using faster region based convolutional neural network. *International Journal of Interactive Multimedia and Artificial Intelligence (IJIMAI)*, 5(5), 30–37. URL: <https://www.ijimai.org/journal/bibcite/reference/2688>
11. Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., Feng, Z., & Qu, R. (2019). A Survey of Deep Learning-based Object Detection. *IEEE Access*, 7, 128837–128868. <https://doi.org/10.1109/ACCESS.2019.2939201>
12. Kirillov, A., Girshick, R., He, K., & Dollár, P. (2019). Panoptic Feature Pyramid Networks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 6399–6408). <https://doi.org/10.1109/CVPR.2019.00655>
13. Kong, T., Yao, A., Chen, Y., & Sun, F. (2017). HyperNet: Towards Accurate Region Proposal Generation and Joint Object Detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 845–853). <https://doi.org/10.1109/CVPR.2016.95>
14. Kreshchenko, Taras., & Yushchenko, Yury. (2022). Parking Spot Occupancy Classification Using Deep Learning. *NaUKMA Research Papers. Computer Science*, 5, 72–78. <https://doi.org/10.18523/2617-3808.2022.5.72-78>
15. Liang, Han, Lee, Seong-Cheol., & Seo, Suyoung. (2022). Automatic recognition of road damage based on lightweight attentional convolutional neural network. *Sensors*, 22(24), article ID 9599. <https://doi.org/10.3390/s22495999>
16. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2020). Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 318–327. <https://doi.org/10.1109/TPAMI.2018.2858826>
17. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D.,... & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In: *European Conference on Computer Vision*, (pp. 740–755). Springer, Cham. [https://doi.org/10.1007/978-3-319-10602-1\\_48](https://doi.org/10.1007/978-3-319-10602-1_48)
18. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. In: *European Conference on Computer Vision*, (pp. 21–37). Springer, Cham. [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2)
19. Naddaf-Sh, S., Naddaf-Sh, M. M., Kashani, A. R., & Zargarzadeh, H. (2020, December). An efficient and scalable deep learning approach for road damage detection. In: *2020 IEEE International Conference on Big Data*, (pp. 5602–5608). IEEE. URL: <https://ieeexplore.ieee.org/abstract/document/9377751>
20. Object segmentation with Mask R-CNN | by Abhijeet Pujara | Medium. (n.d.). Retrieved January 1, 2024. URL: <https://abhijeetpujara.medium.com/object-segmentation-with-mask-r-cnn-6e9c06cef81c>
21. Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767. <https://doi.org/10.48550/arXiv.1804.02767>
22. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 779–788). <https://doi.org/10.1109/CVPR.2016.91>
23. Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
24. Tan, M., Pang, R., & Le, Q. V. (2020). EfficientDet: Scalable and Efficient Object Detection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 10781–10790). <https://doi.org/10.1109/CVPR42600.2020.01080>
25. Vilcapoma, Piero, et al. (2024). Comparison of Faster R-CNN, YOLO, and SSD for Third Molar Angle Detection in Dental Panoramic X-rays." *Sensors* 24.18, article ID 6053. <https://doi.org/10.3390/s24186053>
26. Xu, Xiangyang, et al. (2022). Crack detection and comparison study based on faster R-CNN and mask R-CNN. *Sensors* 22.3, article ID 1215. <https://doi.org/10.3390/s22031215>
27. YOLOv8 vs Faster R-CNN: A Comparative Analysis – Keylabs. (2024, January). URL: [https://keylabs.ai/blog/yolov8-vs-faster-r-cnn-a-comparative-analysis/?utm\\_source=chatgpt.com](https://keylabs.ai/blog/yolov8-vs-faster-r-cnn-a-comparative-analysis/?utm_source=chatgpt.com)
28. YOLOv8 vs Mask R-CNN: In-depth Analysis and Comparison – Labelvisor. (n.d.). URL: <https://www.labelvisor.com/yolov8-vs-mask-r-cnn-in-depth-analysis-and-comparison>
29. Youssouf, Njayer. "Traffic sign classification using CNN and detection using faster-RCNN and YOLOV4." *Heliyon* 8.12 (2022). <https://doi.org/10.1016/j.heliyon.2022.e11792>

30. Yu, Fenglin. (2024). YOLO, Faster R-CNN, and SSD for cloud detection. *Applied and Computational Engineering*, 37, 239–247. <https://doi.org/10.54254/2755-2721/37/20230514>
31. Zhang, S., Chi, C., Yao, Y., Lei, Z., & Li, S. Z. (2020). Bridging the Gap Between Anchor-based and Anchor-free Detection via Adaptive Training Sample Selection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 9759–9768). <https://doi.org/10.1109/CVPR42600.2020.00976>
32. ZHAO, Yian, et al. (2024). Detsr beat yolos on real-time object detection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 16965–16974). URL: <https://arxiv.org/pdf/2304.08069>
33. Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object Detection with Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212–3232. <https://doi.org/10.1109/TNNLS.2018.2876865>
34. Zhu, X., Pang, J., Yang, C., Shi, J., & Lin, D. (2019). Adapting Object Detectors via Selective Cross-Domain Alignment. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (pp. 687–696). <https://doi.org/10.1109/CVPR.2019.00078>

**A. Ye. Batiuk, I. Yu. Opalinskyi**

*Lviv Polytechnic National University, Lviv, Ukraine*

## COMPARISON OF THE EFFICIENCY OF DIFFERENT COMPUTER VISION MODELS FOR THE AUTOMATED PARKING CONTROL SYSTEM

Breakthroughs in computer vision technology have opened new possibilities for developing advanced automated parking control systems that ensure accurate detection, monitoring, and management of vehicles. These systems greatly reduce the need for human involvement while notably boosting operational efficiency and overall productivity within parking facilities. By processing video data from surveillance cameras in real time, computer vision models allow for quick precise vehicle tracking and occupancy assessment. Leading models in this field include YOLO, Faster R-CNN, SSD, EfficientDet, and Mask R-CNN, each with distinct strengths and trade-offs in terms of detection speed, accuracy, and computational needs. YOLO, for instance, is known for its high-speed image processing, making it suitable for real-time applications where speed is essential, though it may not achieve the same precision as two-stage models like Faster R-CNN and Mask R-CNN, which are better suited for complex scenarios due to their greater accuracy, albeit at slower speeds. SSD provides a middle-ground approach, balancing speed and accuracy for use in mid-sized parking zones. Meanwhile, EfficientDet is optimized for resource-limited environments, such as mobile or embedded systems, maintaining a good speed-accuracy ratio while saving computational power, which makes it ideal for parking control solutions with tight resource constraints. In the course of research, we conduct a thorough comparative analysis, highlighting the advantages and limitations of each model. This analysis evaluates such factors as detection accuracy, speed, and scalability in real-world scenarios, offering practical insights on model applicability based on specific operational needs. Moreover, methods for optimizing these models are discussed to achieve a better balance of speed and accuracy, especially in demanding contexts like large parking facilities with variable vehicle volumes. Our findings suggest that hybrid models combining elements of single-stage and two-stage architectures hold promise, as they blend real-time responsiveness with high precision, marking them as a key direction for further research. In conclusion, this study underscores the need for continuous advancement in computer vision models for parking control, encouraging the development of adaptive solutions that can adjust dynamically to changing requirements in complex parking environments. Further innovations in this field could unlock even greater levels of accuracy, efficiency, and reliability in automated parking management.

**Keywords:** neural networks; automated control; image recognition methods; deep learning; object detection algorithms; video processing speed; model improvement.