

РОЗДІЛ II. ІНЖЕНЕРІЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

УДК 004.8

DOI <https://doi.org/10.26661/2786-6254-2024-1-03>

РОЗРОБКА МОДЕЛІ НАВЧАННЯ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ ІЗ ВІДЕОПОТОКУ ДЛЯ СИСТЕМ УПРАВЛІННЯ ПАРКУВАННЯМ

Марчук Д. К.

*старший викладач кафедри комп'ютерних наук
Державний університет «Житомирська політехніка»
вул. Чуднівська, 103, Житомир, Україна
orcid.org/0000-0001-8675-8047
kipz_mdk@ztu.edu.ua*

Ключові слова: *нейронна мережа, Mask R-CNN, адаптивна нейро-нечітка мережа, навчання, трансферне навчання.*

Сучасна необхідність для майже всіх великих міст – це наявність системи управління паркуванням, яка може знаходити доступні місця для паркування. Така система дозволяє користувачам легко й оперативного визначати, чи є вільне місце на конкретній парковці, й отримувати інформацію про його місцезнаходження. Mask R-CNN – це модель глибокого навчання, яка може використовуватися для розпізнавання та сегментації об'єктів на зображеннях. Її можна використовувати для створення системи розпізнавання вільних паркувальних місць, яка може допомогти водіям знайти місце для паркування. Процес навчання може розглядатися як визначення архітектури мережі, налаштування вагових коефіцієнтів або застосування різних методів донавчання для ефективного виконання поставленої задачі. У роботі досліджено трансферне навчання для моделі Mask R-CNN. Було запропоновано модель трансферного навчання для розпізнавання вільних місць для паркування. Модель складається із трьох етапів. Перший етап – це навчання на великому наборі даних COCO. Навчання на великому наборі даних дозволяє моделі узагальнювати візуальні особливості транспортних засобів і паркувальних місць. Другий етап – це донавчання на маленькому наборі даних, створеному з кадрів відеопотоку. Доновчання на малому наборі даних дозволяє моделі адаптуватися до специфіки паркувальних місць у конкретному місті. Третій етап – це донавчання з використанням адаптивної нейро-нечіткої мережі. Цей етап використовує адаптивну нейро-нечітку мережу для доопрацювання базових характеристик. Загальна модель навчання дає змогу ефективно розпізнавати вільні місця на парковці. Використання трансферного навчання робить цю модель більш ефективною та продуктивною. Трансферне навчання може бути корисним у багатьох випадках: коли доступні обмежені дані для навчання моделі для конкретного завдання (у цьому випадку трансферне навчання може допомогти моделі досягти кращої точності, ніж якщо би її навчали з нуля); коли потрібно швидко навчити модель для нового завдання (трансферне навчання може допомогти скоротити час навчання моделі); коли потрібно навчити модель для завдання, для якого немає доступних даних.

DEVELOPMENT OF A LEARNING MODEL FOR OBJECT RECOGNITION FROM VIDEO STREAM FOR PARKING MANAGEMENT SYSTEMS

Marchuk D. K.

Senior Lecturer at the Department of Computer Sciences

Zhytomyr Polytechnic State University

Chudnivska str., 103, Zhytomyr, Ukraine

orcid.org/0000-0001-8675-8047

kipz_mdk@ztu.edu.ua

Key words: *neural network, Mask R-CNN, adaptive neuro-fuzzy network, training, transfer learning.*

A modern necessity for almost all large cities is the availability of a parking management system that can locate available parking spaces. Such a system allows users to easily and quickly determine whether a particular parking space is available and receive information about its location. Mask R-CNN is a deep learning model that can be used to recognize and segment objects in images. It can be used to create a parking space recognition system that can help drivers find a parking space. The training process can be viewed as defining the network architecture, adjusting the weighting coefficients, or applying various retraining methods to effectively perform the task. In this research, we investigate transfer learning for the Mask R-CNN model. As a result of the study, a transfer learning model for recognizing free parking spaces was proposed. The model consists of three stages. The first stage is training on a large SOSO dataset. Training on a large dataset allows the model to generalize visual features of vehicles and parking spaces. The second stage is retraining on a small dataset created from video stream frames. Training on a small dataset allows the model to adapt to the specifics of parking spaces in a particular city. The third stage is post-training using an adaptive neuro-fuzzy network. This stage uses an adaptive neuro-fuzzy network to refine the basic characteristics. The general learning model allows you to effectively recognize free parking spaces. The use of transfer learning makes this model more efficient and productive. Transfer learning can be useful in many cases. When limited data is available to train a model for a specific task. In this case, transfer learning can help the model achieve better accuracy than if it were trained from scratch. When you need to quickly train a model for a new task. Transfer learning can help reduce model training time. When you need to train a model for a task for which no data is available.

Вступ. Традиційні моделі машинного навчання вимагають навчання з нуля, що потребує великої кількості даних для досягнення високої продуктивності. Методи, які використовуються в такому випадку, мають ізольований підхід до навчання, коли кожна модель незалежно навчається для певної мети, без будь-якої залежності від минулих знань. Однак бувають випадки, коли збирати навчальні дані важко, дорого, а подеколи недоцільно. Таким чином, існує потреба створити модель, яка отримує дані з різних областей. Ця методологія називається трансферним навчанням.

Трансферне навчання (Transfer Learning) – це метод машинного навчання, котрий дозволяє моделям штучного інтелекту використовувати знання, набуті у процесі виконання одного завдання, для покращення виконання іншого, але пов'язаного завдання. Цей метод корисний, коли доступні

обмежені дані для навчання моделі для конкретного завдання. Метод дозволяє моделі використовувати знання, отримані з подібного завдання, яке має більший обсяг даних. Наприклад. Модель, навчена на наборі даних ImageNet, може бути доопрацьована для класифікації типів автомобілів (легкові, вантажівки, автобуси тощо) на основі зображень із СОСО. Це може бути корисно для таких завдань, як автоматичне сортування транспортних засобів на парковках або розпізнавання типу автомобіля на камерах спостереження.

Аналіз попередніх досліджень і публікацій. Глибоке навчання привертає дедалі більше уваги дослідників та успішно застосовується у багатьох галузях, проте у деяких сферах створення великих анотованих наборів даних є складним завданням через високу вартість збору й анотації даних, що, у свою чергу, обмежує розвиток цих

галузей. Завдяки широким перспективам застосування трансферне навчання стало популярним і перспективним напрямком машинного навчання. У дослідженні [1] розглядається понад 40 репрезентативних підходів до перенесення навчання, особливо однорідних, із погляду даних і моделі. У роботі представлено застосування системи трансферного навчання. Щоб продемонструвати ефективність перенесення навчання, в експериментах використовується понад 20 репрезентативних моделей. Моделі навчаються на трьох різних наборах даних, тобто Amazon Reviews, Reuters-21578 та Office-31. Статті [2–4] зосереджені на огляді поточних досліджень трансферного навчання у різних галузях. Дослідження [5–8] висвітлюють різноманітні методи й алгоритми трансферного навчання. Останніми роками дослідницьке співтовариство з питань трансферного навчання переважно зосереджене на таких двох аспектах: адаптація домену та перенесення доменів із кількох джерел. Завдяки глибокому навчанню за останні роки було досягнуто значних успіхів у різних сферах, однак багато моделей навчаються з нуля й уникають перенесення навчання через різний характер даних [9].

Мета роботи. Ефективність глибокого навчання значною мірою залежить від наявності масштабних наборів даних, які забезпечують модель різноманітними зображеннями. Ці зображення містять інформацію, що допомагає моделі краще розуміти різні категорії та, як наслідок, покращувати результати класифікації, однак у реальних прикладних задачах часто складно зібрати велику кількість зображень, яких би вистачало для навчання моделі. Це обмежує можливість навченої моделі, тому стає актуальним питання про те, як максимально ефективно використовувати обмежені дані для навчання моделі для рішення задачі визначення вільного паркувального місця на парковках міста.

Викладення основного матеріалу й аналіз отриманих результатів. Трансферне навчання є потужним методом машинного навчання, який може бути використаний для покращення ефективності моделей штучного інтелекту в широкому спектрі завдань. Формально трансферне навчання можна описати таким чином: Нехай T – базове завдання, а S – нове завдання. Нехай D_T – набір даних для базового завдання, а D_S – набір даних для нового завдання. Нехай θ – параметри моделі.

Трансферне навчання – це метод машинного навчання, який дозволяє покращити точність моделі для нового завдання S , використовуючи набір даних для базового завдання T . Формально трансферне навчання можна описати як задачу оптимізації:

$$\min_{\theta} L(\theta, D_S), \quad (1)$$

де $L(\cdot)$ – функція втрат.

На практиці трансферне навчання виконується у два етапи:

– Навчання початкової моделі. На цьому етапі модель навчається на наборі даних для базового завдання D_T .

– Фінальне навчання. На цьому етапі модель адаптується до нового завдання D_S .

На етапі остаточного навчання модель перенавчається на наборі даних для нового завдання D_S . Можуть використовуватися різні методи адаптації моделі до нового завдання.

Система розпізнавання вільних місць для паркування прагне полегшити життя автовласникам у місті, автоматично знаходячи вільні місця для паркування. Ця система використовує камери, щоб сканувати парковки та визначати вільні місця. Інформація про вільні місця може бути доступна автовласникам через мобільний застосунок. Етапи роботи системи:

– Вибір потрібної парковки або найближчих за геопозицією.

– Камери на парковках фіксують вільні та зайняті місця.

– Система аналізує дані з обраних парковок і передає інформацію про вільні місця.

Головною складовою частиною програмної системи визначення вільних паркувальних місць є модуль розпізнавання. Для реалізації алгоритму розпізнавання вільних місць для паркування у режимі реального часу використана модель Mask R-CNN, модель глибокого навчання, яка може одночасно виявляти та сегментувати об'єкти на зображеннях.

Трансферне навчання можна реалізувати для Mask R-CNN кількома способами. Один із поширених способів – це заморозити нижні шари моделі та перенавчити лише верхні шари. Це дозволяє моделі зберегти базові поняття, які вона навчилася на базовому завданні, й адаптуватися до нового завдання, лише навчаючи шари, відповідальні за виявлення конкретних ознак для нового завдання.

Для реалізації цього способу можна використовувати такі кроки:

1. Завантажити та навчити модель Mask R-CNN на наборі даних для базового завдання.

2. Заморозити нижні шари моделі.

3. Перенавчити верхні шари моделі на наборі даних для нового завдання.

4. Оцінити отриманий результат. На цьому етапі модель оцінюється на наборі даних для нового завдання.

Для оцінки ефективності моделі після трансферного навчання використовуються такі метрики [10]:

– Accuracy – відсоток правильно класифікованих зображень.

- Precision – відсоток правильних позитивних результатів.
- Recall – відсоток правильних позитивних результатів, які були правильно виявлені.
- F1-score – середнє гармонічне відношення precision і recall.
- mAP (Mean Average Precision) – середнє значення середньої точності.

Для реалізації алгоритму розпізнавання вільних місць для паркування у режимі реального часу потрібно реалізувати такі дії (рис. 1):

1. Провести попереднє навчання на наборі даних COCO [11].
2. Адаптувати модель до нового завдання шляхом точного налаштування ваг вищих рівнів.
3. Для подальшої роботи потрібно використати кадри з камер автостоянок. Для цього необхідно створити набір зображень із відеокадрів обраних парковок міста.
4. Провести другий етап навчання на відеокадрах обраних парковок.

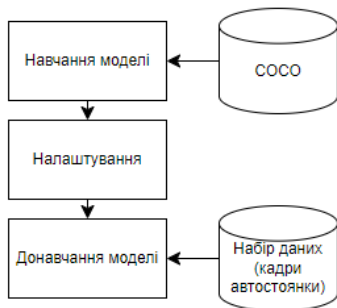


Рис. 1. Трансферне навчання для вирішення задачі розпізнавання вільного паркувального місця

Для створення набору даних зображень автостоянок міста використаємо відеозаписи з автостоянок, які надалі будуть розбиватися на кадри-зображення. Набір даних буде складатися з розмічених зображень чотирьох найбільших парковок міста. У зображеннях представлені різні часові проміжки (день, вечір, ніч, ранок). Як результат, для донавчання було відібрано близько 1 000 зображень.

Після того, як пройдено повний курс навчання, потрібно провести тестування моделі. Для цього необхідно встановити зв'язок із відеокамерою; отримати кадри з відеопотоку; розпізнати всі автомобілі на поточному кадрі; отримати вільні та зайняті місця на парковці.

Інтерфейс користувача реалізовано за допомогою мобільного застосунку. На рис. 2 зображено модель розпізнавання вільних місць для паркування.

У реалізованій системі можуть бути не всі парковки міста, тому треба ще раз пройти донавчання. Для цього можна використати адаптивну нейро-нечітку мережу (АННМ). Адаптивна нейро-нечітка мережа – це тип нейронної мережі, яка поєднує принципи нейронних мереж і нечіткої логіки.

Переваги використання АННМ для донавчання моделі:

- Ефективність. АННМ може навчатися на невеликих наборах даних більш ефективно, ніж традиційні нейронні мережі.
- Точність. АННМ може давати більш точні результати, ніж традиційні нейронні мережі, завдяки своїй здатності враховувати невизначеність.
- Інтерпретованість. АННМ легше інтерпретувати, ніж традиційні нейронні мережі, що робить їх більш зручними для використання у задачах, де важлива пояснюваність.

Для АННМ було виведено 18 правил на основі системи нечіткого логічного висновку типу Такагі-Сугено. Наприклад: Якщо <транспортний засіб займає менше 10% паркувального місця> То <місце сильно вільне> F1.

Для навчання та тестування було використано два набори даних і база нечітких правил (таблиця 1).

Результати. Для експериментального дослідження було використано модель Mask R-CNN від Matterport – це офіційна реалізація Mask R-CNN із відкритим вихідним кодом, побудована на Python, Keras і TensorFlow. Усі експерименти проводилися на комп'ютерах, оснащених процесором Intel Xeon(R) E2678 v3@2,50 ГГц і графічним процесо-

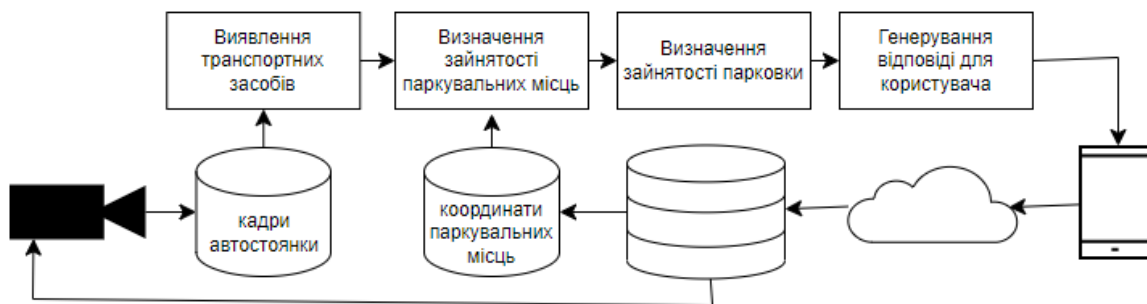


Рис. 2. Модель розпізнавання вільних і зайнятих місць на обраній парковці

Таблиця 1

Склад навчальних і тестових даних

	Навчальний набір (кількість об'єктів)	Набір для тестування (кількість об'єктів)
СОСО	56000	14000
Спеціальний набір	700	300
База нечітких правил	18	18

Таблиця 2

Порівняння mAP

	AP	AP ₅₀	AP ₇₅	основа
Після першого донавчання	60,53	82,48	64,53	Mask R-CNN
Після адаптивного навчання	61,94	91,11	69,68	Mask R-CNN +ANFIS

ром Nvidia RTX 3060ti через FiftyOne, реалізовані CUDA 11.8 і CUDNN 11.8.

Параметри, які було використано для тренування: 50 епох. Навчання моделі проходило на моделі СОСО, починаючи зі швидкості навчання 0,001. Було використано розпад ваги 0,0001 та імпульс 0,9. Це зайняло вісім годин навчання на одній машині з 1 графічним процесором. Доновчання моделі проходило на спеціальному наборі та зайняло дві години.

За результатами дослідження можна зробити висновки, що mAP [0.5;0.95] за хороших погодних умов у день mAP= 92,23. У таблиці 2 наведено порівняльні дані щодо класичної моделі Mask R-CNN. AP, AP50 та AP75 надають детальнішу інформацію про продуктивність моделі об'єктного виявлення. AP (Average Precision) – це середня величина точності (precision) по всім рівням відкликання (recall) для цього класу об'єктів. AP50 – це середня точність при рівні відкликання 50%. Це означає, що з усіх об'єктів певного класу, присутніх на зображеннях, модель повинна правильно класифікувати 50% із них. AP75 – це середня точність при рівні відкликання 75%. Це означає, що модель повинна правильно класифікувати 75% об'єктів певного класу з усіх присутніх на зображеннях.

Виявлення паркувальних місць і їх класифікацію за критерієм зайнятості автомобілями

ускладнює низка чинників: розташування камер; різний рівень освітленості; погодні умови; природне стирання ліній паркувальної розмітки або взагалі парковка без розмітки; перекриття ліній розмітки транспортними засобами на стоянці або тими, що проїжджають мимо; некоректне паркування тощо.

Висновки. Було запропоновано модель трансферного навчання, яка включає в себе спочатку навчання на великому наборі даних СОСО, донавчання на маленькому наборі даних, створеному з кадрів із відеопотоку, взятих із камер, розташованих на автостоянках міста. Другий етап донавчання використовує адаптивну нейро-нечітку мережу. Ця мережа може динамічно адаптуватися до нових даних, що покращує її ефективність і продуктивність порівняно із традиційними нейронними мережами. Завдяки використанню заздалегідь навчених моделей адаптивна нейро-нечітка мережа може почати навчання з кращої точки. Це скорочує час навчання та покращує кінцевий результат.

Проведено опис етапів створення моделі навчання для системи розпізнавання вільних місць для паркування. Визначено методи підготовки вхідних даних і загальна модель навчання.

Використання трансферного навчання покращує ефективність моделі та зменшує час і ресурси, необхідні для її навчання.

ЛІТЕРАТУРА

1. A Comprehensive Survey on Transfer Learning / F. Zhuang et al. *Proceedings of the IEEE*. 2021. Vol. 109. № 1. P. 43–76. URL: <https://doi.org/10.1109/jproc.2020.3004555>
2. A survey on deep transfer learning / Tan, Chuanqi, et al. *In Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4–7, 2018, Proceedings, Part III* 27. P. 270–279. Springer International Publishing.
3. Raffel, Colin, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research* 21. 2020. № 1. 2020. P. 5485–5551.
4. Liu, Ruijun, Yuqian Shi, Changjiang Ji, and Ming Jia. A survey of sentiment analysis based on transfer learning. *IEEE access*. 2019. № 7. P. 85401–85412.

5. Label efficient learning of transferable representations across domains and tasks / Luo, Zelun, et al. *Advances in neural information processing systems*. 2017. № 30.
6. Balanced distribution adaptation for transfer learning / J. Wang et al. *2017 IEEE International Conference on Data Mining (ICDM)*, New Orleans, LA, 18–21 November 2017. URL: <https://doi.org/10.1109/icdm.2017.150>
7. Mukhlif A.A., Al-Khateeb B., Mohammed M.A. Incorporating a Novel Dual Transfer Learning Approach for Medical Images. *Sensors*. 2023. Vol. 23. № 2. P. 570. URL: <https://doi.org/10.3390/s23020570>
8. Hussain, M., Bird, J.J., & Faria, D.R. (2018). A Study on CNN Transfer Learning for Image Classification. In *UK Workshop on Computational Intelligence*. 2018. № 840. P. 191–202 (Advances in Computational Intelligence Systems). https://doi.org/10.1007/978-3-319-97982-3_16
9. Свідоцтво про реєстрацію авторського права на твір № 109822, Комп'ютерна програма «Алгоритмічно-програмне забезпечення обробки та аналізу потоку кадрів відеоданих що надходять з камер міста» / Автори: Левківський В.Л., Марчук Г.В., Ципоренко В.В., Марчук Д.К. Дата реєстрації: 25.11.2021
10. Марчук Д.К., Граф М.С. Методи оцінки ефективності моделей виявлення об'єктів у комп'ютерному зорі. *Вісник Херсонського національного технічного університету*. 2023. № 2 (85). С. 181–186. URL: <https://doi.org/10.35546/kntu2078-4481.2023.2.25>
11. *COCO Common Objects Context*. URL: <https://cocodataset.org/#home>.

REFERENCES

1. A Comprehensive Survey on Transfer Learning / F. Zhuang et al. *Proceedings of the IEEE*. 2021. Vol. 109, no. 1. P. 43–76. URL: <https://doi.org/10.1109/jproc.2020.3004555>
2. A survey on deep transfer learning / Tan, Chuanqi, et al. In *Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4–7, 2018, Proceedings, Part III 27*, p.270-279. Springer International Publishing.
3. Raffel, Colin, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research* 21, no. 1 (2020): 5485–5551.
4. Liu, Ruijun, Yuqian Shi, Changjiang Ji, and Ming Jia. A survey of sentiment analysis based on transfer learning. *IEEE access* 7 (2019): 85401–85412.
5. Label efficient learning of transferable representations across domains and tasks / Luo, Zelun, et al. *Advances in neural information processing systems*. 2017. 30.
6. Balanced distribution adaptation for transfer learning / J. Wang et al. *2017 IEEE International Conference on Data Mining (ICDM)*, New Orleans, LA, 18–21 November 2017. 2017. URL: <https://doi.org/10.1109/icdm.2017.150>
7. Mukhlif A.A., Al-Khateeb B., Mohammed M.A. Incorporating a Novel Dual Transfer Learning Approach for Medical Images. *Sensors*. 2023. Vol. 23, no. 2. P. 570. URL: <https://doi.org/10.3390/s23020570>
8. Hussain, M., Bird, J.J., & Faria, D.R. (2018). A Study on CNN Transfer Learning for Image Classification. In *UK Workshop on Computational Intelligence*. 840 : Springer, 2018. pp. 191–202 (Advances in Computational Intelligence Systems). https://doi.org/10.1007/978-3-319-97982-3_16
9. Svidotstvo pro reiestratsiiu avtorskoho prava na tvir № 109822, Kompiuterna prohrama “Alhorytmichno-programne zabezpechennia obrobky ta analizu potoku kadriv videodanykh shcho nadkhodiat z kamer mista”. Avtory: Levkivskiy V.L., Marchuk H.V., Tsyporenko V.V., Marchuk D.K. Data reiestratsii: 25.11.2021 (in Ukrainian)
10. Marchuk D.K., Hraf M.S. Metody otsinky efektyvnosti modelei vyivlennia ob'ektiv u kompiuternomu zori. *Visnyk Khersonskoho natsionalnoho tekhnichnoho universytetu*. 2023. № 2(85). S. 181–186. URL: <https://doi.org/10.35546/kntu2078-4481.2023.2.25> (in Ukrainian)
11. *COCO Common Objects Context*. URL: <https://cocodataset.org/#home>