

КОМПЛЕКСНІ ОНТОЛОГІЧНІ ТА НЕЙРОМЕРЕЖЕВІ МОДЕЛІ ФОТОГРАФІЧНИХ ОБРАЗІВ

Галушка О.В., Шинкаренко В.І.

Український державний університет науки і технологій, Україна, Дніпро

Анотація. Ця стаття досліджує інноваційний підхід до розробки систем комп'ютерного зору, заснований на глибокій інтеграції онтологічних та нейромережових моделей для ефективного аналізу фотографічних зображень. У цій роботі пропонується новий гібридний метод у контексті сучасних викликів, пов'язаних з розпізнаванням зображень, таких як висока різноманітність класів та проблеми оклюзії об'єктів на фотографіях низької якості. Цей метод надає моделям комп'ютерного зору здатність до глибшого розуміння контексту та семантики зображень, використовуючи онтологічні структури для представлення візуального вмісту. Аналітичний процес включає ідентифікацію ключових об'єктів на зображенні та визначення їхнього онтологічного контексту, що дозволяє нейронним мережам ефективно виконувати градієнтну трансформацію вхідних даних для більш точного розпізнавання та класифікації. Запропонована модель демонструє потенціал для отримання переваги над традиційними методами у задачах комп'ютерного зору, відкриваючи нові можливості для розширення застосувань комп'ютерного зору у науковій, промисловій та побутовій сферах.

Ключові слова: Штучний інтелект, машинне навчання, глибоке навчання, нейронна мережа, онтологія, граф знань.

Основний матеріал. Системи та додатки, які використовують моделі машинного зору, відіграють все більшу роль в усіх аспектах життєдіяльності людини: в побуті, виробництві, науці, мистецтві та т.і.

Розробка та навчання моделей, здатних забезпечувати необхідну точність своєї роботи, ставлять перед інженерами штучного інтелекту серйозні виклики. Оглядаючи існуючі методики [1] в задачах комп'ютерного зору [2], ми стикаємося з рядом проблем, які впливають на ефективність та точність моделей.

Важливим аспектом є кількість та ступінь різноманітності класів, які модель може навчитись ефективно розпізнавати. Зі збільшенням кількості класів зростає складність моделі та потреба в більш різноманітних та об'ємних навчальних даних, що може ускладнити процес навчання. Досягнення високої

точності роботи моделі часто вимагає значних обчислювальних ресурсів. Це особливо актуально при роботі з великими наборами даних та складними моделями глибокого навчання [3].

Ще однією значною проблемою є частково закриті об'єкти. Коли об'єкти на зображенні перекриваються або закриваються іншими елементами сцени, це може значно знизити здатність моделі правильно їх розпізнати. Це вимагає від моделей глибокого розуміння контексту зображення та вміння виділяти корисну інформацію навіть з частково видимих об'єктів.

Фотографії поганої якості часто не можуть бути ефективно оброблені існуючими моделями. Це вимагає від розробників впровадження додаткових методів попередньої обробки зображень [4] або розробки більш розширених моделей, які здатні ефективно працювати з низькоякісними даними. Архітектурним прикладом таких моделей може бути сімейство нейронних мереж автоенкодерів [5].

Для підвищення ефективності моделей в задачах комп'ютерного зору, пропонується використання нового альтернативного підходу, який базується на гібридному методі, що інтегрує онтології [6, 7] з нейронними мережами. Цей підхід використовує онтологічну структуру як основу для семантичної контекстної структуризації зображення, що дозволяє моделі розуміти та інтерпретувати візуальний зміст на більш глибокому рівні.

Ініціалізація процесу аналізу починається з поділу вхідного зображення на регіони або субрегіони. Ці регіони називають «регіональними пропозиціями» або «регіональними кандидатами». Ця фаза аналізу відповідає за створення набору потенційних регіонів на зображенні, які ймовірно містять об'єкти. Гібридний підхід на основі онтологій та нейронних мереж сам не створює набір регіональних кандидатів, натомість він покладається на сторонні методи, такі як Selective Search [8] або Edge Boxes [9]. Виділення набору регіонів створює достатні умови для переходу до наступної фази аналізу – виявлення якірних об'єктів, як відправних точок для формування онтологічного контексту зображення.

Наступна фаза аналізу зображення полягає у попередній активації та виявленні об'єктів, які можуть бути ідентифіковані одразу без додаткових кроків. Їх розпізнавання дозволяє локалізувати онтологію та виділити певний її сегмент. Такий сегмент називається когнітивним контекстом зображення, який використовується для фокусування на концептах описаних в його межах. Цей контекст виступає базисом для нейронно-градієнтної трансформації зображення.

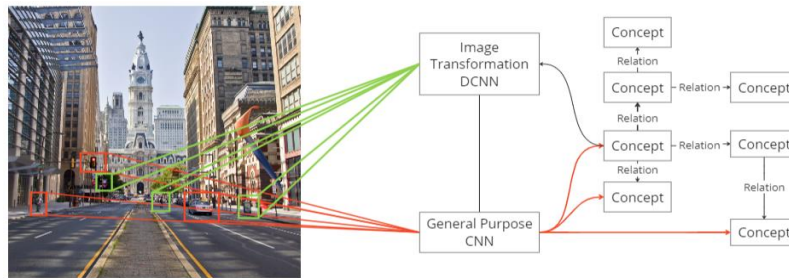


Рисунок 1 – Загальна структура онтологічної неймережевої моделі

Нейронно-градієнтна трансформація вхідного зображення – це наступна фаза аналізу, яку можна розуміти як складову механізму моделі. Суть методу полягає в частковій трансформації зображення використовуючи активації обраних нейронів моделі [10]. Механізм уяви призначений для комбінування та співставлення фактичних (зображених на вхідній фотографії) та еталонних образів об'єктів на основі локалізованого онтологічного сегменту.

Основна ідея такого підходу полягає у важливості врахування контексту того, що відбувається за аналогією з реальним світом. Кожна жива істота, яка сприймає цей світ за допомогою рецепторів будь-якого типу, завжди враховує контекст в комбінації з вхідними даними для формування остаточного виводу того, що вона зараз баче або чує, який смак або запах відчуває. При цьому контекст, як правило, не є однорідним за своєю природою, так як інформація поступає з різних органів чуття.

Висновок. Використання онтологічної структури в моделях глибокого навчання може значно покращити здатність моделей до розуміння контексту зображень, роблячи їх більш адаптивними та ефективними при роботі з різноманітними та складними візуальними даними. Онтології допомагають моделям встановлювати зв'язки між об'єктами та їх атрибутами, підтримуючи

більш глибоке семантичне розуміння зображень, що може виявитися особливо корисним у ситуаціях з перекритими об'єктами або нечіткими зображеннями. Таким чином, гібридний підхід, що поєднує онтології з нейронними мережами, відкриває нові перспективи для розвитку більш інтелектуальних та ефективних систем комп'ютерного зору.

ЛІТЕРАТУРА / REFERENCES

1. Siwei, C. Analysis of object recognition trends based on deep learning. ACE (2023) Vol. 5: 292-299. DOI: 10.54254/2755-2721/5/20230582.
2. Snyder WE, Qi H. Fundamentals of Computer Vision. Cambridge University Press; 2017.
3. Farsal, W., Anter, S. & Ramdani, M. Deep learning: An overview. in Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications 1–6 (ACM, 2018). DOI: <https://doi.org/10.1145/3289402.3289538>
4. Distantе A, Distantе C Image segmentation. In Handbook of image processing and computer vision. Springer, Cham (2020), pp 271–332. DOI: https://doi.org/10.1007/978-3-030-42374-2_5
5. Xinlin, Z., Yuanmeng, H., Li Z., Yajing, K., Xiang, G., Huajing, W. Review of Deep Neural Network Based on Auto-encoder. DEStech Transactions on Computer Science and Engineering. 2019, 10.12783/dtcse/iciti2018/29087.
6. Qiao, L., Yang, L., Hong, D., Yao, L., Zhiguang, Q. Knowledge Graph Construction Techniques[J]. Journal of Computer Research and Development, 2016, 53(3): 582-600. DOI: 10.7544/issn1000-1239.2016.20148228
7. Shynkarenko, V., Zhuchyi, L. Ontological suitability analysis of railway tracks for high-speed traffic. Computer Systems and Information Technologies, 2022, 3, 11-21. DOI: <https://doi.org/10.31891/csit-2022-3-2>
8. Uijlings, J. R., Van De Sande, K. E., Gevers, T., & Smeulders, A. W. Selective search for object recognition. International journal of computer vision, 2013, 104, 154-171.
9. Zitnick, C. L., & Dollár, P. Edge boxes: Locating object proposals from edges. In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13 (pp. 391-405). Springer International Publishing.
10. Zeiler, Matthew D., Fergus, R.: Visualizing and understanding convolutional networks. In: Fleet, David, Pajdla, Tomas, Schiele, Bernt, Tuytelaars, Tinne (eds.) ECCV 2014. LNCS, vol. 8689, pp. 818–833. Springer, Cham (2014). DOI: https://doi.org/10.1007/978-3-319-10590-1_53

COMPLEX ONTOLOGICAL AND NEURAL NETWORK MODELS OF PHOTOGRAPHIC IMAGES

Halushka O., Shynkarenko V.

Abstract. *This article explores an innovative approach to developing computer vision systems, based on the deep integration of complex ontological and neural network models for effective analysis of photographic images. This work proposes a novel hybrid method in the context of contemporary challenges associated with image recognition, such as a high diversity of classes and problems with the occlusion of objects in low-quality photos. This method provides computer vision models with a deeper understanding of the context and semantics of images, utilizing ontological structures to represent visual content. The analytical process includes the identification of key objects in the image and the determination of their ontological context, allowing neural networks to effectively perform a gradient transformation of the input data for more accurate recognition and classification. The proposed model demonstrates the potential for gaining an advantage over traditional methods in computer vision tasks, opening new possibilities for expanding computer vision applications in scientific, industrial, and domestic spheres.*

Keywords. *Artificial intelligence, machine learning, deep learning, neural network, ontology, knowledge graph.*