

ПРИКЛАДНІ ЗАСТОСУВАННЯ СИСТЕМ КОМП'ЮТЕРНОГО ЗОРУ

Ященко С.С.

e-mail: stanislav.iashchenko@nure.ua

Науковий керівник – канд. техн. наук, ст. викл. Путятіна О.Є.
Харківський національний університет радіоелектроніки, каф. ІНФ
м. Харків, Україна

This paper explores the applied applications of computer vision (CV) systems across various industries. It defines the fundamental principles of CV and highlights core technologies such as convolutional neural networks (CNNs) and deep learning. The study examines practical use cases, including biometric authentication in mobile devices, intelligent video surveillance, autonomous driving systems, and automated medical image analysis. Furthermore, the research addresses critical challenges hindering widespread adoption, notably the scarcity of high-quality annotated training data, the fatal consequences of recognition errors in mission-critical environments, and the profound ethical and privacy concerns associated with biometric surveillance and algorithmic bias.

Комп'ютерний зір (Computer Vision) – це міждисциплінарна галузь штучного інтелекту, яка розробляє теоретичні та алгоритмічні основи для вилучення, аналізу та інтерпретації високорівневої інформації з цифрових зображень або відеопотоків. Головна мета цієї технології полягає в автоматизації задач, які виконує зорова система людини, дозволяючи обчислювальним машинам «бачити» навколишній світ, ідентифікувати об'єкти та приймати рішення на основі візуального контексту [1]. В умовах стрімкого зростання обсягів візуальних даних, прикладне застосування систем комп'ютерного зору стає одним із головних драйверів технологічного прогресу.

Основні технології Фундаментом сучасних систем комп'ютерного зору є методи глибинного навчання (Deep Learning). Переважна більшість задач розв'язується за допомогою згорткових нейронних мереж (Convolutional Neural Networks, CNN), які здатні автоматично виділяти ієрархічні візуальні ознаки з пікселів – від простих ліній до складних форм [1]. До ключових технологічних процесів належать: класифікація зображень, виявлення об'єктів (Object Detection, наприклад, архітектура YOLO), семантична сегментація (попиксельне розділення об'єктів) та відстеження об'єктів у реальному часі.

Приклади застосування. Сьогодні алгоритми комп'ютерного зору глибоко інтегровані в різні сфери людської життєдіяльності та бізнес-процеси. До найбільш репрезентативних прикладів належать:

Біометрична ідентифікація та безпека: Технології розпізнавання облич (наприклад, FaceID) у сучасних смартфонах здійснюють побудову точної

3D-карти обличчя користувача для безпечного розблокування пристрою та авторизації фінансових транзакцій [2]. Сучасні системи також використовують алгоритми Liveness Detection для захисту від спроб підміни обличчя (спуфінгу). Крім того, інтелектуальні камери відеоспостереження у рамках концепції «Розумного міста» (Smart City) здатні в режимі реального часу аналізувати багатотисячні натовпи, виявляти підозрілу або агресивну поведінку, фіксувати правопорушення та автоматично ідентифікувати осіб, що перебувають у розшуку, співставляючи їхні візуальні ознаки з національними базами даних.

Автономні транспортні засоби та робототехніка: Системи автопілота (наприклад, в автомобілях Tesla) повністю спираються на комплексний аналіз відеопотоку з масиву бортових камер. Комп'ютерний зір забезпечує високоточне розпізнавання дорожньої розмітки, знаків, пішоходів та інших учасників руху, дозволяючи системі орієнтуватися у складному динамічному середовищі [3]. Алгоритми семантичної сегментації в режимі реального часу розділяють кадр на класи об'єктів, що дозволяє бортовому комп'ютеру миттєво прогнозувати траєкторії руху, уникати зіткнень та приймати рішення щодо маневрування без участі водія.

Медична діагностика та аналіз зображень: Алгоритми ШІ демонструють результати, що часто перевищують точність лікарів-людей при аналізі складних медичних знімків (рентгенографії, КТ, МРТ та гістологічних препаратів). Вони здатні аналізувати об'ємні дані на піксельному рівні та виявляти мікроскопічні патології – наприклад, ранні стадії онкологічних захворювань, мікроаневризми при діабетичній ретинопатії чи специфічні ураження легень [4]. Впровадження таких систем значно прискорює процес постановки діагнозу в ургентних станах, знижує рутинне навантаження на радіологів та слугує надійним інструментом «другої думки» (second opinion) для мінімізації лікарських помилок.

Переваги та проблеми впровадження Основні переваги таких систем очевидні: безперервність роботи (24/7), здатність обробляти терабайти візуальної інформації за мілісекунди та мінімізація людського фактора у рутинних задачах.

Однак, впровадження комп'ютерного зору стикається з низкою критичних проблем:

Нестача даних для навчання: Для досягнення високої точності нейромережі потребують колосальних масивів якісно розмічених даних (анотованих датасетів). Вузькоспеціалізовані галузі часто страждають від дефіциту таких даних [1].

Фатальні наслідки помилок розпізнавання (False Positives / False Negatives): У критичних системах помилка алгоритму має катастрофічну ціну. Наприклад, хибне розпізнавання об'єкта (або його ігнорування) системою автопілота на високій швидкості призводить до летальних ДТП, а помилка в медичній діагностиці може коштувати пацієнту життя [3].

Етичні та правові колізії: Масове використання систем розпізнавання облич камерами спостереження порушує фундаментальне право на приватність. Крім того, виникають питання алгоритмічної упередженості (bias), коли через незбалансовані навчальні дані системи гірше розпізнають людей певних етнічних груп, що породжує серйозні соціальні та етичні конфлікти [5].

Підсумовуючи, слід зазначити, що системи комп'ютерного зору пройшли стрімкий шлях від експериментальних лабораторних розробок до фундаментальної технології прикладного рівня. Інтеграція методів глибинного навчання дозволила обчислювальним машинам досягти, а в багатьох вузькоспеціалізованих задачах і перевершити, людський рівень візуального аналізу [1, 4]. Сьогодні ця технологія виступає потужним каталізатором інновацій, кардинально трансформуючи підходи до забезпечення громадської безпеки, управління автономним транспортом та надання високотехнологічної медичної допомоги.

Водночас, глобальне масштабування систем комп'ютерного зору генерує нові технологічні та соціальні виклики. Критична залежність моделей від величезних масивів розмічених даних, надвисока ціна алгоритмічної помилки у системах життєзабезпечення (автопілоти, медицина) та глибокі етичні протиріччя, пов'язані з масовим біометричним стеженням, вимагають негайного вирішення [3, 5]. Подальший успішний розвиток цієї галузі залежатиме не лише від вдосконалення математичних моделей у напрямку їхньої прозорості (Explainable AI) та надійності, але й від створення чітких міжнародних законодавчих рамок. Лише такий комплексний підхід гарантуватиме безпечне, неупереджене та контрольоване використання технологій візуального штучного інтелекту в інтересах усього суспільства.

Список використаних джерел:

1. Yann LeCun, Yoshua Bengio, Geoffrey Hinton Deep learning // Nature. – 2015. – Vol. 521, No. 7553. – P. 436–444.
2. Florian Schroff, Dmitry Kalenichenko, James Philbin FaceNet: a unified embedding for face recognition and clustering // IEEE Conference on Computer Vision and Pattern Recognition 2015. – 2015. – P. 815–823.
3. Sorin Grigorescu, Bogdan Trasnea, Tiberiu Cocias, Gheorghe Macesanu A survey of deep learning techniques for autonomous driving // Journal of Field Robotics. – 2020. – Vol. 37, No. 3. – P. 362–386.
4. Pranav Rajpurkar та ін. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning // arXiv. – 2017. – arXiv:1711.05225.
5. Brent Daniel Mittelstadt та ін. The ethics of algorithms: mapping the debate // Big Data & Society. – 2016. – Vol. 3, No. 2.