Лекція 1 Вступ до комп’ютерного зору та математичні засади формування зображень

- Поняття та завдання комп’ютерного зору.
- Місце комп’ютерного зору в системах штучного інтелекту та цифрових двійниках.
- Математичний опис цифрового зображення як дискретної функції.
- Просторове та яскравісне квантування.
- Колірні моделі (RGB, HSV, Lab).
- Основні етапи обробки зображень.
- Огляд сучасних наукових і прикладних напрямів розвитку комп’ютерного зору.

Комп'ютерний зір (Computer Vision, CV) — це галузь штучного інтелекту (ШІ), яка навчає комп’ютери інтерпретувати та розуміти візуальний світ. Використовуючи цифрові зображення з камер, відео та глибоке навчання (Deep Learning), системи комп’ютерного зору можуть автоматично ідентифікувати, класифікувати об’єкти, а також реагувати на них, імітуючи людське зорове сприйняття.

Слайд 1. Поняття комп'ютерного зору

Мета: Надати машинам можливість "бачити" та аналізувати візуальні дані для отримання корисної інформації, аналогічно до того, як це робить людське око та мозок.

Принцип роботи: Алгоритми обробляють пікселі зображення, виділяють ознаки, розпізнають шаблони та приймають рішення.

Основа: Сучасний комп'ютерний зір базується на методах машинного навчання (Machine Learning) та, зокрема, згорткових нейронних мережах (CNN), які дозволяють автоматично навчатися на великих наборах даних.

https://www.nvidia.com/en-us/glossary/computer-vision/

Слайд 2. Основні завдання комп'ютерного зору

Комп'ютерний зір вирішує низку складних візуальних задач, серед яких:

1. **Класифікація зображень (Image Classification):** Присвоєння всьому зображенню однієї або кількох міток (наприклад, визначити, чи є на фото "кішка", "автомобіль" або "медичний знімок").
2. **Виявлення об'єктів (Object Detection):** Не просто визначення наявності об'єкта, а й його локалізація на зображенні за допомогою обмежувальної рамки (bounding box).
3. **Сегментація зображень (Image Segmentation):** Попіксельне виділення об'єктів.
    1. *Семантична сегментація:* Класифікація кожного пікселя (наприклад, усі пікселі дороги — одним кольором).
    2. *Екземплярна сегментація (Instance Segmentation):* Виділення кожного окремого об'єкта (наприклад, розрізнення автомобіля №1 від автомобіля №2).

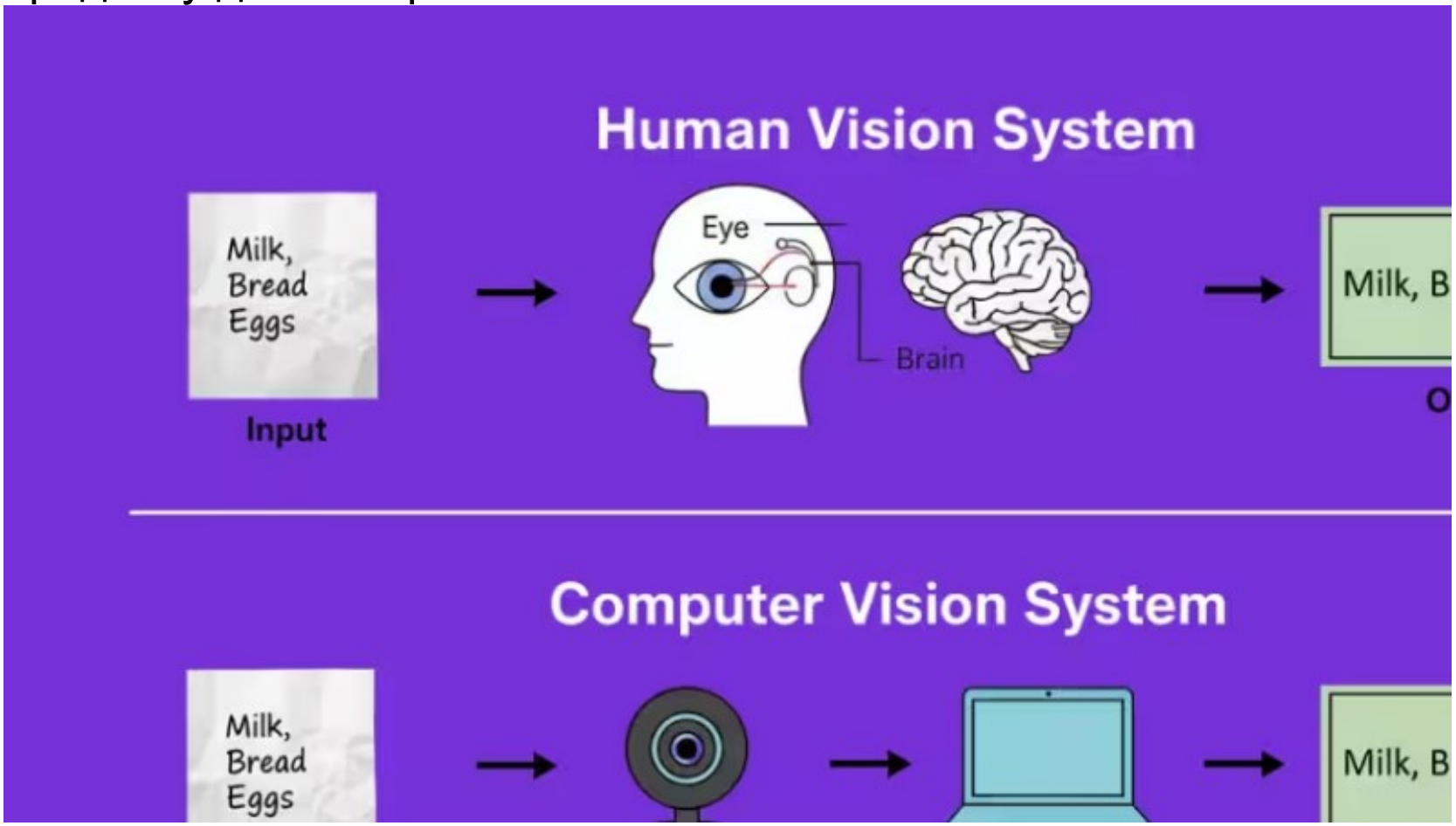
Слайд 3. Основні завдання комп'ютерного зору

4. **Відстеження об'єктів (Object Tracking)**: Навчання моделі стежити за рухом об'єкта у відеопотоці.
5. **Розпізнавання облич та людей (Face/Person Recognition)**: Ідентифікація людей на основі візуальних характеристик.
6. **Оптичне розпізнавання символів (OCR)**: Перетворення зображень тексту (сканованих документів, номерів авто) у машинозчитуваний текст.
7. **Оцінка пози (Pose Estimation)**: Визначення положення тіла людини або об'єкта на основі ключових точок.
8. **Реконструкція сцени (Scene Reconstruction)**: Створення 3D-моделі на основі 2D-зображень.

https://www.geeksforgeeks.org/computer-vision/computer-vision-tasks/
https://riseuplabs.com/computer-vision/

Слайд 4. Як працює комп'ютерний зір?

Системи комп'ютерного зору працюють за структурованим алгоритмом, який перетворює необроблені візуальні дані на значущу інформацію, придатну для використання.



Слайд 5.Алгоритм роботи систем комп’ютерного зору

Крок 1. Збір зображень
Процес починається із збору візуальних даних. Цифрові датчики, такі як камери, сканери, дрони або інші пристрої для створення зображень, надають нерухомі зображення або безперервні відеопотоки у форматах JPEG, PNG або MP4. Ці необроблені дані становлять вхідні дані, від яких залежить кожен наступний етап.

Крок 2. Попередня обробка та покращення якості зображення
Після збору даних вони стандартизуються та очищаються, щоб система могла надійно з ними працювати. Попередня обробка зазвичай включає такі операції, як зміна розміру, нормалізація, зменшення шуму, перетворення колірного простору та регулювання контрасту. Покращення якості зображення може підкреслити важливі структури та зменшити нерелевантні варіації, забезпечуючи, що моделі нижчого рівня отримують послідовні візуальні дані високої якості.

Слайд 6. Алгоритм роботи систем комп’ютерного зору

Крок 3. Витяг ознак

Після попередньої обробки система перетворює необроблені пікселі в більш інформативні представлення. На цьому етапі вона ідентифікує візуальні шаблони, такі як краї, текстури, форми та розподіл кольорів. У класичних підходах ці особливості обчислюються за допомогою явних алгоритмів; у сучасних системах глибокі нейронні мережі автоматично вивчають багаторівневі особливості, шар за шаром. Результатом є компактний опис зображення, що містить інформацію, необхідну для розпізнавання та прийняття рішень.

Слайд 7. Алгоритм роботи систем комп’ютерного зору

Крок 4. Навчання та аналіз моделі

Потім представлення ознак пов'язуються із семантичним значенням за допомогою моделей машинного навчання. Під час навчання система отримує марковані приклади, щоб навчитися пов'язувати шаблони ознак із категоріями або завданнями. Залежно від того, чи метою є класифікація, виявлення, сегментація або комбінація завдань, зазвичай використовуються конволюційні нейронні мережі, архітектури виявлення об'єктів та моделі на основі трансформаторів Під час виконання навчена модель аналізує вхідні ознаки для виявлення та розпізнавання об'єктів, віднесення їх до заздалегідь визначених класів та визначення взаємозв'язків між ними. Саме тут система переходить від низькорівневих описів «як виглядає зображення» до структурованої інформації про «що присутнє» та «що воно представляє».

Слайд 8. Алгоритм роботи систем комп’ютерного зору

Крок 5. Вихідні дані, інтерпретація та прийняття рішень
На останньому етапі система генерує вихідні дані на основі свого аналізу. Це може бути мітка, набір обмежувальних рамок, маска сегментації, текстовий опис або рішення вищого рівня. Внутрішньо цей результат також інтерпретується в контексті: виявлені об'єкти групуються за категоріями, пріоритетизуються та оцінюються відповідно до правил або логіки подальших дій, щоб вирішити, що слід робити далі.

У багатьох додатках ці результати надходять до більших систем, які реєструють результати, запускають автоматизовані дії або оновлюють моделі з часом. У сукупності ці етапи дозволяють системам комп'ютерного зору просуватися від необроблених пікселів до сприйняття, перетворюючи неструктуровані візуальні дані на структуровані, значущі результати, які інші компоненти можуть зрозуміти та на основі яких можуть діяти.

Слайд 9. Місце комп’ютерного зору в системах штучного інтелекту та цифрових двійниках

У сучасних технологічних екосистемах комп'ютерний зір (CV) виступає ключовим інструментом для збору та інтерпретації візуальних даних, поєднуючи фізичний світ із цифровим інтелектом.

Місце у системах штучного інтелекту (ШІ)

Комп'ютерний зір є однією з найважливіших областей штучного інтелекту, що відповідає за «сприйняття».

Сенсорний рівень: CV діє як «очі» для ШІ-систем, дозволяючи їм отримувати корисну інформацію з відеопотоків, зображень та 3D-сканувань.

Аналітична база: Алгоритми машинного та глибокого навчання обробляють ці дані для прийняття автономних рішень.

Інтелектуальна автоматизація: CV інтегрується в роботизовані системи, дозволяючи їм орієнтуватися в просторі та взаємодіяти з об'єктами без втручання людини.

Слайд 10. Місце комп’ютерного зору в системах штучного інтелекту та цифрових двійниках

Роль у цифрових двійниках (Digital Twins)
Цифровий двійник — це віртуальна копія фізичного об'єкта або процесу. Комп'ютерний зір є «містком», який підтримує актуальність цієї копії.

1. **Синхронізація в реальному часі**: CV здійснює моніторинг фізичного об'єкта та миттєво передає дані про його стан у цифрову модель.
2. **Виявлення аномалій та дефектів**: Система порівнює поточне візуальне зображення об'єкта з його еталонною цифровою моделлю, виявляючи зноси, поломки або відхилення від техпроцесу.
3. **Навчання в симуляції**: Цифрові двійники використовуються для генерації синтетичних даних, на яких навчаються моделі комп'ютерного зору. Це дозволяє безпечно тренувати ШІ для роботи в складних або небезпечних умовах.
4. **Прогнозне обслуговування**: Аналізуючи візуальні зміни об'єкта (наприклад, появу мікротріщин), ШІ всередині цифрового двійника може прогнозувати вихід з ладу обладнання.

Слайд 11. Математичний опис цифрового зображення як дискретної функції

Цифрове зображення математично описується як **двовимірна дискретна функція** $f(m, n)$ (або $I(x, y)$), яка задана на скінченній множині точок (пікселів), що утворюють прямокутну матрицю.

Ось детальний математичний опис:

1. Визначення функції

Зображення представляється як матриця розміром $M \times N$ (рядки та стовпці), де функція $f(m, n)$ повертає значення інтенсивності або кольору для пікселя з координатами (m, n):

$$f(m, n), \quad m \in \{0, 1, \ldots, M-1\}, \quad n \in \{0, 1, \ldots, N-1\}$$

- m, n — цілочисельні координати пікселя (рядки та стовпці).
- $f(m, n)$ — значення яскравості (інтенсивності) або кольору в цій точці.

Слайд 12. Математичний опис цифрового зображення як дискретної функції

2. Дискретизація (просторова та за рівнем)

Перехід від неперервного зображення до цифрового передбачає дв

- **Дискретизація простору (Sampling):** Неперервна сцена $f(x, y)$ на скінченну кількість пікселів.
- **Квантування яскравості (Quantization):** Значення інтенсивності набувають лише обмеженого набору цілочисельних значень.

3. Діапазон значень (Квантування)

Для монохромних (відтінки сірого) зображень значення функції $f(m,$ належать цілочисельному діапазону від 0 до $L-1$, де L — кількість квантування (найчастіше 256, тобто 8-бітне зображення):

Слайд 13. Математичний опис цифрового зображення як дискретної функції

4. Матрична форма

Цифрове зображення можна представити як матрицю:

$$\mathbf{I} = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N- \end{bmatrix}$$

5. Кольорові зображення

Кольорове зображення (наприклад, RGB) описується як векторна
набір із трьох матриць, де кожна матриця відповідає одному коль
(Red, Green, Blue):

https://gyansanchay.csjmu.ac.in/wp-content/uploads/2022/10/unit1.pdf

Слайд 14. Просторове та яскравісне квантування

Процес перетворення неперервного зображення в цифрову форму складаєтьс
двох основних етапів: **дискретизації** (визначення просторових координат) та
квантування (визначення значень яскравості).

1. Просторове квантування (Дискретизація)

Це процес розбиття неперервного зображення на окремі елементи —
пікселі.

- **Мета**: Визначити просторову роздільну здатність зображення.
- **Суть**: Неперервна площина розділяється сіткою на $M \times N$ комірок. Кожна к
(піксель) має фіксовані координати (x, y).
- **На що впливає**: Визначає рівень деталізації та чіткості. Більша кількість пікс
одиницю площі дозволяє відтворювати дрібніші деталі.

Слайд 15. Просторове та яскравісне квантування

2. Яскравісне квантування

Це процес перетворення неперервного значення інтенсивності світла пікселі у дискретне цифрове значення.

- **Мета**: Визначити глибину кольору або кількість рівнів сірого.
- **Суть**: Діапазон яскравості розбивається на скінченну кількість рівні Найчастіше це ступені двійки ($L = 2^k$), де k — кількість біт на піксел
  - **1-бітне** (бінарне): 2 рівні (чорний та білий).
  - **8-бітне**: 256 рівнів сірого (стандарт для більшості фото).
- **На що впливає**: Визначає плавність переходів між кольорами або в сірого.
- **Проблема**: При низькій кількості рівнів квантування виникає ефект

Слайд 16. Порівняльна таблиця

Параметр	Просторове квантування	Яскравісне квант
Що дискретизується	Координати (x, y)	Значення функції f
Одиниця виміру	Пікселі, DPI, PPI	Біти (bit depth), рів

Слайд 17. Колірні моделі (RGB, HSV, Lab)

Колірні моделі — це математичні способи представлення кольорів за
набору чисел. Вибір моделі залежить від завдання: відображення на
редагування чи точний аналіз кольору.

1. RGB (Red, Green, Blue)

Це **адитивна** модель, заснована на змішуванні трьох основних коль
світла. τομεας φυσικης +1

- **Принцип**: Колір утворюється шляхом додавання інтенсивностей
зеленого та синього каналів.
- **Застосування**: Монітори, телевізори, цифрові камери (все, що ви
світло).
- **Геометрія**: Представляється у вигляді куба, де чорний — точка (0
(255, 255, 255) (для 8-бітного кодування).

Слайд 18. Колірні моделі (RGB, HSV, Lab)

2. HSV (Hue, Saturation, Value)

Модель, орієнтована на **сприйняття людиною**, яка розділяє інформацію п
яскравість.

- **Компоненти**:
  - **Hue (Відтінок)**: Тип кольору (червоний, жовтий тощо), вимірюється в від 0° до 360°.
  - **Saturation (Насиченість)**: «Чистота» кольору; 0% — сірий, 100% — ч колір.
  - **Value (Значення/Яскравість)**: Світлота; 0% — чорний.
- **Застосування**: Обробка зображень (сегментація об'єктів), графічні ред

Слайд 19. Колірні моделі (RGB, HSV, Lab)

3. Lab (CIELAB)

Апаратно-незалежна та **перцептивно рівномірна** модель, що охоплює всі к які бачить людське око.

- **Компоненти**:
  - **L (Lightness)**: Яскравість (від 0 до 100).
  - **a**: Позиція між червоним (+) та зеленим (-).
  - **b**: Позиція між жовтим (+) та синім (-).
- **Застосування**: Професійна корекція кольору, поліграфія, наукові розрахун різниці кольорів.

Слайд 20. Порівняння основних характеристик

Модель	Тип	Для чого найкраща	Головна пе
RGB	Адитивна	Техніка, дисплеї	Рідна для з
HSV	Циліндрична	Сегментація, пошук кольору	Інтуїтивніст

Слайд 21. Математичні формули перетворення з RGB у HSV
Ось основні математичні перетворення між популярними колірними моделями. Для розрахунків значення RGB зазвичай нормалізують у діапазон (поділивши на 255).

Нехай $max = \max(R, G, B)$, $min = \min(R, G, B)$, а $\Delta = max -$

- **Відтінок (Hue):**

$$H = \begin{cases} 0^\circ, & \text{якщо } \Delta = \\ 60^\circ \times \left(\frac{G-B}{\Delta} \mod 6 \right), & \text{якщо } max \\ 60^\circ \times \left(\frac{B-R}{\Delta} + 2 \right), & \text{якщо } max \\ 60^\circ \times \left(\frac{R-G}{\Delta} + 4 \right), & \text{якщо } max \end{cases}$$

- **Насиченість (Saturation):**

$$S = \begin{cases} 0, & \text{якщо } max = 0 \\ \frac{\Delta}{max}, & \text{якщо } max \neq 0 \end{cases}$$

Слайд 22. Математичні формули перетворення з RGB у Lab

Це перетворення зазвичай робиться через проміжний простір **XYZ** (
CIE).

Етап А: Перехід до XYZ

Використовується матричне множення (для стандарту D65):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Етап Б: Перехід до Lab

$$L = 116 \cdot f(Y/Y_n) - 16$$

$$a = 500 \cdot [f(X/X_n) - f(Y/Y_n)]$$

Слайд 23. Математичні формули перетворення з RGB у Чорно-біле (Grayscale) та RGB у CMYK (для друку)

RGB у Чорно-біле (Grayscale)

Найпростіша функція яскравості (Y), що враховує фізіологію ока (чутлив зеленого):

$$Y = 0.299R + 0.587G + 0.114B$$

RGB у CMYK (для друку)

Спочатку обчислюється допоміжна компонента K (Black):

$$K = 1 - \max(R, G, B)$$

Якщо $K = 1$, то $C = M = Y = 0$. Якщо $K < 1$:

$$C = \frac{1 - R - K}{} \qquad M = \frac{1 - G - K}{} \qquad Y$$

Слайд 24. Алгоритм перетворення HSV в RGB

Перетворення з **HSV** назад у **RGB** є складнішим, оскільки воно залежи
який «сектор» колірного кола потрапляє значення відтінку (H).

Алгоритм перетворення HSV → RGB

Нехай $H \in [0, 360°)$, $S \in [0, 1]$, $V \in [0, 1]$.

1. **Обчислення допоміжних значень**:

   - $C = V \times S$ (Хрома — інтенсивність кольору)

Слайд 25. Алгоритм перетворення HSV в RGB

2. **Визначення проміжних значень (R', G', B') залежно від сектора H**

- Якщо $0° \leq H < 60°$: $(R', G', B') = (C, X, 0)$
- Якщо $60° \leq H < 120°$: $(R', G', B') = (X, C, 0)$
- Якщо $120° \leq H < 180°$: $(R', G', B') = (0, C, X)$
- Якщо $180° \leq H < 240°$: $(R', G', B') = (0, X, C)$
- Якщо $240° \leq H < 300°$: $(R', G', B') = (X, 0, C)$
- Якщо $300° \leq H < 360°$: $(R', G', B') = (C, 0, X)$

3. **Фінальний розрахунок**:

Слайд 26. Алгоритм перетворення Lab в RGB

Це двоступеневий процес, що потребує повернення до апарат простору XYZ.

Крок 1: Lab → XYZ

- $P = (L + 16)/116$
- $X = X_n \cdot f^{-1}(P + a/500)$
- $Y = Y_n \cdot f^{-1}(P)$
- $Z = Z_n \cdot f^{-1}(P - b/200)$

  (Де $f^{-1}(t) = t^3$, якщо $t > 0.206$, інакше $f^{-1}(t) = \frac{1}{7.787}(t -$

Крок 2: XYZ → RGB

Використовується зворотна матриця:

Слайд 27. Алгоритм сегментації за кольором

Алгоритм сегментації за кольором найчастіше реалізують у просторі HSV, оскільки він дозволяє відокремити «чистий» колір від впливу тіней та освітлення (на відміну від RGB). Алгоритм складається з наступних кроків:

1. Підготовка даних

- Зображення зчитується в форматі **RGB**.
- Виконується перетворення в **HSV** (за формулами, які ми розгляда
  ключовий крок, щоб яскравість (V) не заважала розпізнавати колі

2. Визначення колірних меж (Thresholding)

Для виділення конкретного об'єкта (наприклад, червоного яблука) з
пороги: **нижній** та **верхній** вектори $[H_{min}, S_{min}, V_{min}]$ та $[H_{max}, S_{ma}$

- **Hue (H)**: Визначає сам колір (наприклад, синій — це діапазон нав

Слайд 28. Алгоритм сегментації за кольором

3. Створення бінарної маски

Програма проходить по кожному пікселю. Якщо значення пікселя по
заданий діапазон, він стає **білим** (1), якщо ні — **чорним** (0).

$$M(x,y) = \begin{cases} 1, & \text{якщо Lower} \le f(x,y) \le \text{Upper} \\ 0, & \text{інакше} \end{cases}$$

4. Морфологічна фільтрація (Очищення)

Після сегментації часто залишається «шум» (поодинокі білі пікселі) а
об'єкті. Застосовуються операції:

- **Ерозія (Erosion)**: «З'їдає» дрібний шум.
- **Диляція (Dilation)**: Розширює межі об'єкта, заповнюючи порожне

Слайд 29. Основні етапи обробки зображень

Конвеєр обробки зображень зазвичай поділяють на рівні: від простого до складного інтелектуального аналізу.

1. Попередня обробка (Pre-processing)

Мета — покращити візуальну якість та підготувати дані до аналізу.

- **Усунення шумів**: Використання фільтрів (медіанний, Гаусів) для очи «зернистості».
- **Корекція яскравості та контрасту**: Гістограмна еквалізація, зміна є

Слайд 30. Основні етапи обробки зображень

2. Сегментація (Segmentation)

Це процес розділення зображення на складові частини або виділення ок об'єктів.

- **Порогова обробка (Thresholding)**: Перетворення на чорно-біле за п яскравості.
- **Виділення контурів (Edge Detection)**: Пошук меж об'єктів (наприкла Кенні або Собеля).

Слайд 31. Основні етапи обробки зображень

2. Сегментація (Segmentation)

Це процес розділення зображення на складові частини або виділення с об'єктів.

- **Порогова обробка (Thresholding)**: Перетворення на чорно-біле за г яскравості.
- **Виділення контурів (Edge Detection)**: Пошук меж об'єктів (наприкла Кенні або Собеля).
- **Кластеризація**: Об'єднання пікселів за схожістю кольору чи текстур

3. Представлення та опис (Representation & Description)

Перетворення виділених областей у зручну для комп'ютера форму.

Слайд 32. Основні етапи обробки зображень

4. Розпізнавання та інтерпретація (Recognition)

Найвищий рівень, де системі надається зміст.

- **Класифікація**: Присвоєння об'єкту певної категорії (наприклад, «ав
  «людина»).
- **Інтерпретація**: Розуміння сцени в цілому (наприклад, аналіз медичн
  виявлення патологій).

5. Стиснення та зберігання (Compression)

Слайд 33. Огляд сучасних наукових і прикладних напрямів розвитку комп’ютерного зору

Сучасний комп'ютерний зір (Computer Vision, CV) переживає етап стрімкої трансформації, переходячи від простого розпізнавання об'єктів до глибокого розуміння контексту та генерації контенту. Розвиток галузі визначається кількома ключовими науковими та прикладними векторами

Слайд 34. Наукові напрями розвитку

- **Vision Transformers (ViT)**: Поступове витіснення традиційних згорткових нейромереж (CNN) архітектурами трансформерів, які краще моделюють глобальні залежності в зображеннях.
- **Мультимодальний ШІ**: Поєднання візуальних даних із текстовими або аудіоданими для глибшого розуміння сцен (наприклад, моделі типу GPT-4V).
- **Генеративний комп'ютерний зір**: Використання дифузійних моделей для створення реалістичних зображень, редагування відео та синтезу 3D-сцен.
- **Explainable AI (XAI)**: Розробка методів, що роблять рішення моделей прозорими та зрозумілими для людини, що критично для медицини та безпеки.
- **Edge AI (Периферійні обчислення)**: Оптимізація алгоритмів для роботи безпосередньо на пристроях (смартфони, дрони) без звернення до хмарних серверів.

Слайд 35. Прикладні напрями застосування

- **Автономний транспорт**: Використання CV для детектування пішо планування траєкторій та побудови HD-карт у реальному часі.
- **Медична діагностика**: Аналіз рентгенівських знімків, МРТ та КТ д. ранніх стадій захворювань (рак, діабетична ретинопатія) з точніст перевищує людську.
- **Робототехніка та хірургія**: Роботизовані системи для проведення малоінвазивних операцій з доповненою реальністю (AR) та автома відстеженням інструментів.
- **Розумне сільське господарство**: Моніторинг стану посівів за доп виявлення шкідників та автоматизація збору врожаю.
- **Промисловість (Індустрія 4.0)**: Автоматизований контроль якост лініях, що дозволяє виявляти до 90% дефектів продукції.

Слайд 36. Порівняння ключових трендів

Напрям	Основна технологія	Ключовий ефект
Генерація	Дифузійні моделі, GAN	Створення контенту, синтетичні д
Аналітика	Transformers, Multimodal AI	Розуміння складних сцен та конте
Медицина	Deep Learning (CNN/ViT)	Рання діагностика та точна хірург