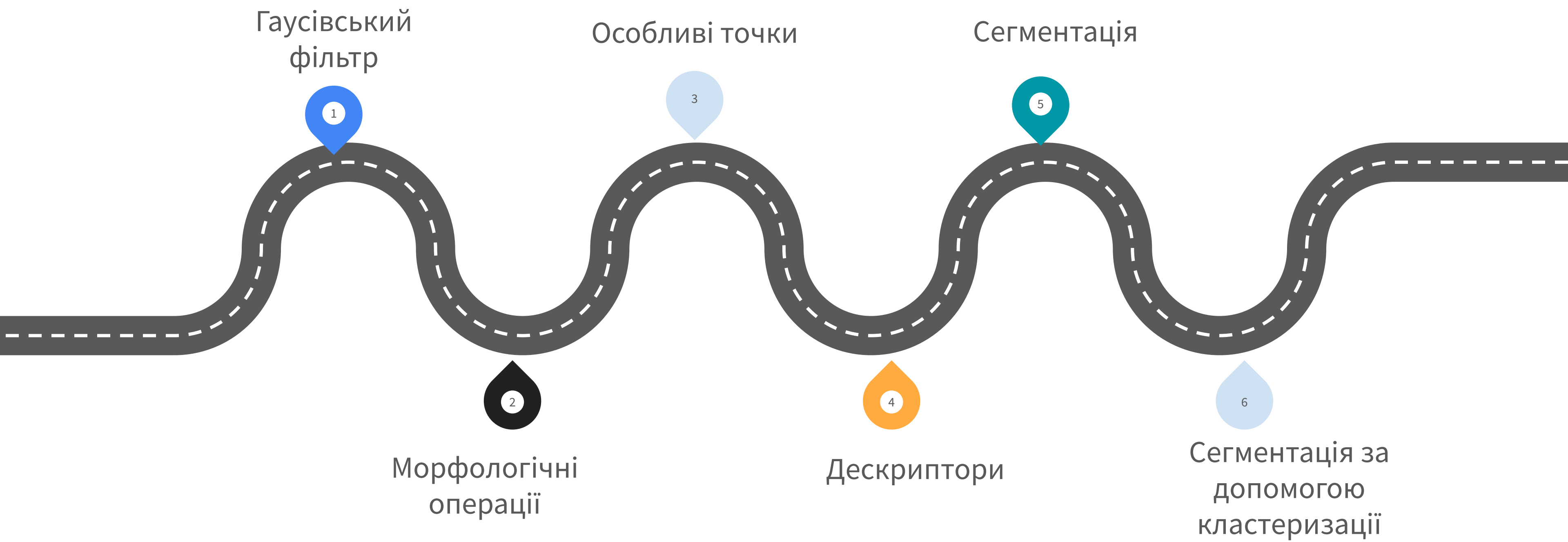
Розпізнавання образів

Колірні простори

Сьогодні на лекції



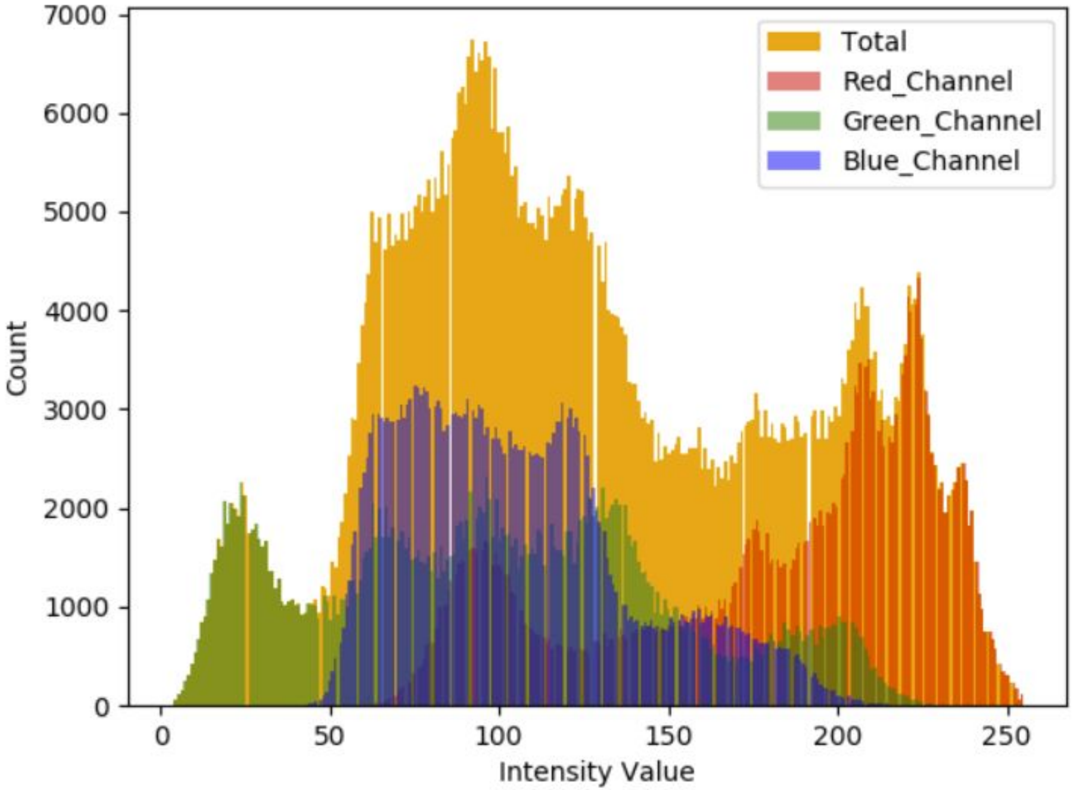
Гістограма цифрових зображень

Нехай є скалярне зображення I з пікселями (i,j,u) де 0≤u≤255. Ми визначаємо абсолютні частоти рахуючи скільки раз значення u зустрічається в носії Ω, який містить всі пікселі.

$$H_I(u) = \left|\{(x, y) \in \Omega : I(x, y) = u\}\right|,$$

Відносні частоти - значення між 0 і 1 - можна порівняти з функцією щільності ймовірності розподілу дискретної випадкової величини (p):

$$h_I(u) = \frac{H_I(u)}{|\Omega|}$$



Вирівнювання гістограми

I ⟶ Inew, для всіх u.

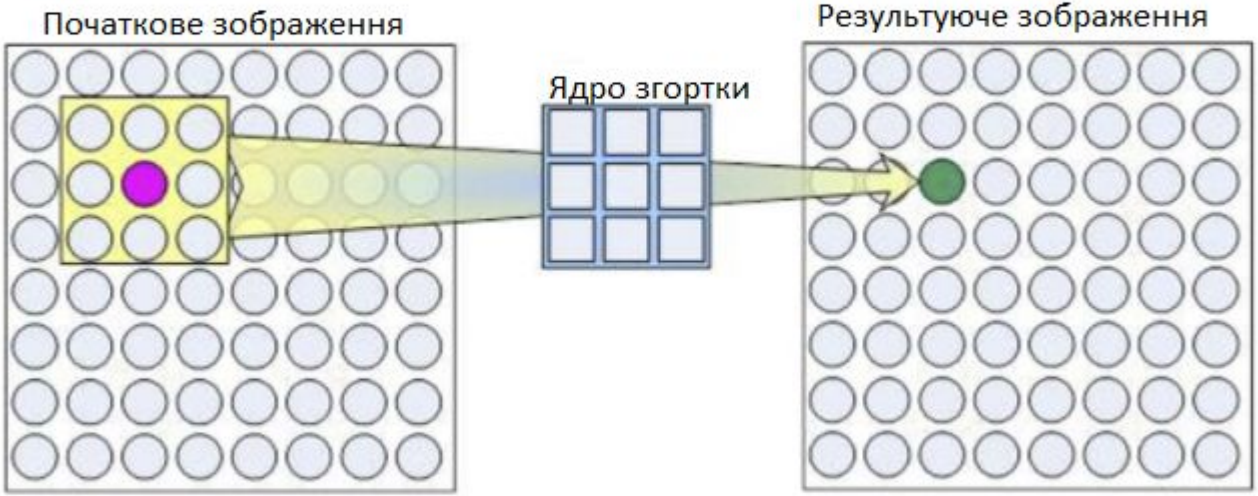
$$H_{I_{new}}(u) = \text{const} = \frac{N_{cols} \cdot N_{rows}}{G_{\max} + 1}$$

I розміром Ncols x Nrows з абсолютним частотами $H_I(u)$

Перетворення здійснюється за допомогою

градаційної функції: $g(u) = c_I(u) \cdot G_{\max}$

Згладжування. Фільтр Гауса



$$G_{\sigma,\mu_x,\mu_y}(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(x-\mu_x)^2+(y-\mu_y)^2}{2\sigma^2}\right) =$$

$$= \frac{1}{2\pi\sigma^2}\cdot e^{-\frac{(x-\mu_x)^2}{2\sigma^2}}\cdot e^{-\frac{(y-\mu_y)^2}{2\sigma^2}},$$

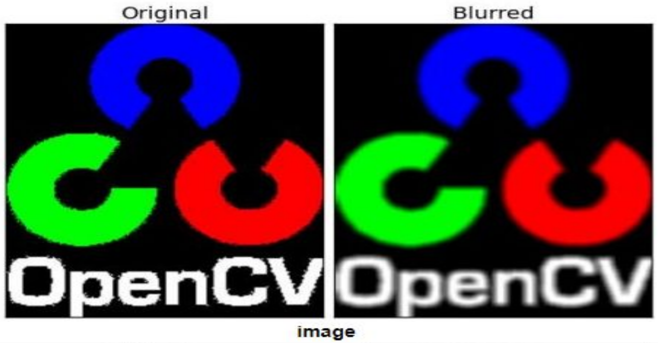
Гаусівський простір масштабів

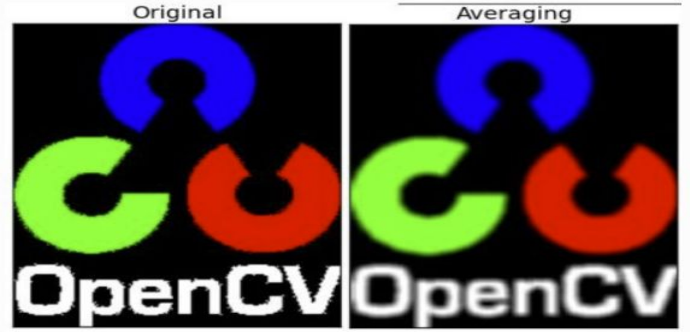
$L(x, y, a\sigma)$, $a^n \cdot \sigma$, $a > 1$

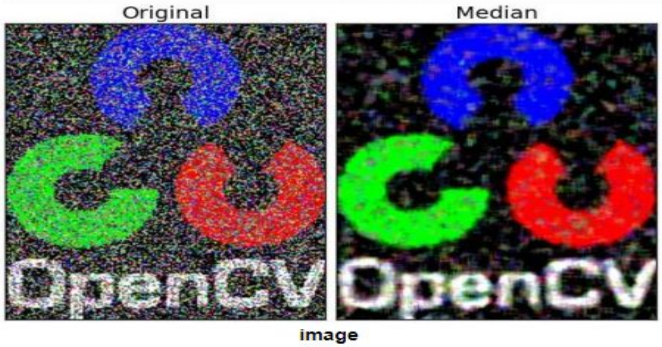
$n = 0, 1, \ldots, m$

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) = \frac{1}{\pi s}\cdot e^{-\frac{x^2}{2\sigma^2}}\cdot e^{-\frac{y^2}{2\sigma^2}}.$$

cv.GaussianBlur()







$\{4, 7, 3, 1, 8, 7, 4, 5, 2, 3, 8\} \longrightarrow$ Медіана 4

$1, 2, 3, 3, 4, 4, 5, 7, 7, 8, 8.$

LoG

Лапласіан гаусіана

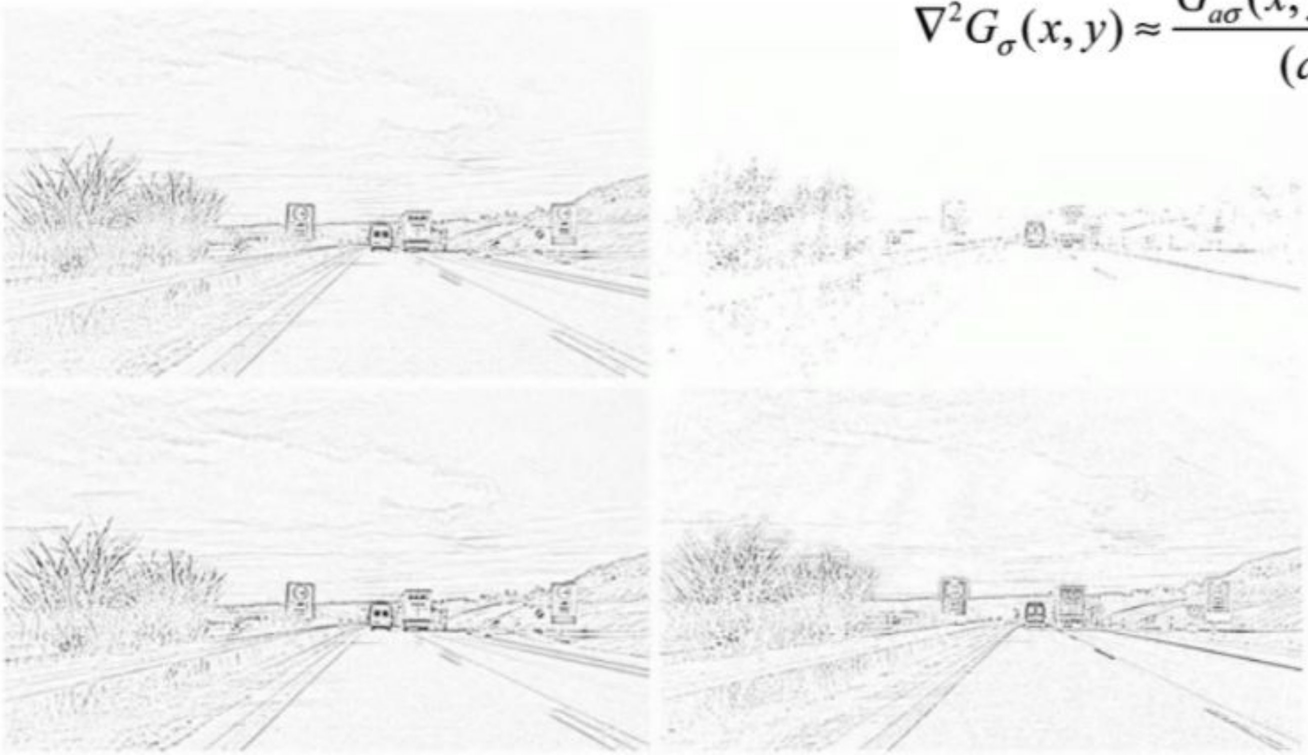
$$\nabla^2(G_\sigma * I) = I * \nabla^2 G_\sigma$$

DoG

Різниця гаусіан $D_{\sigma,a}(x, y) = L(x, y, \sigma) - L(x, y, a\sigma)$

a>1 - масштабний коефіцієнт

$$\nabla^2 G_\sigma(x,y) \approx \frac{G_{a\sigma}(x,y) - G_\sigma(x,y)}{(a-1)\sigma^2}, \text{ де } a = 1,6$$



```
Cv2.resize,
cv2.getRotationMatrix2D,
cv2.warpAffine
```

Аугментація

$$(x_t,\ y_t) = T[(x,\ y)]$$

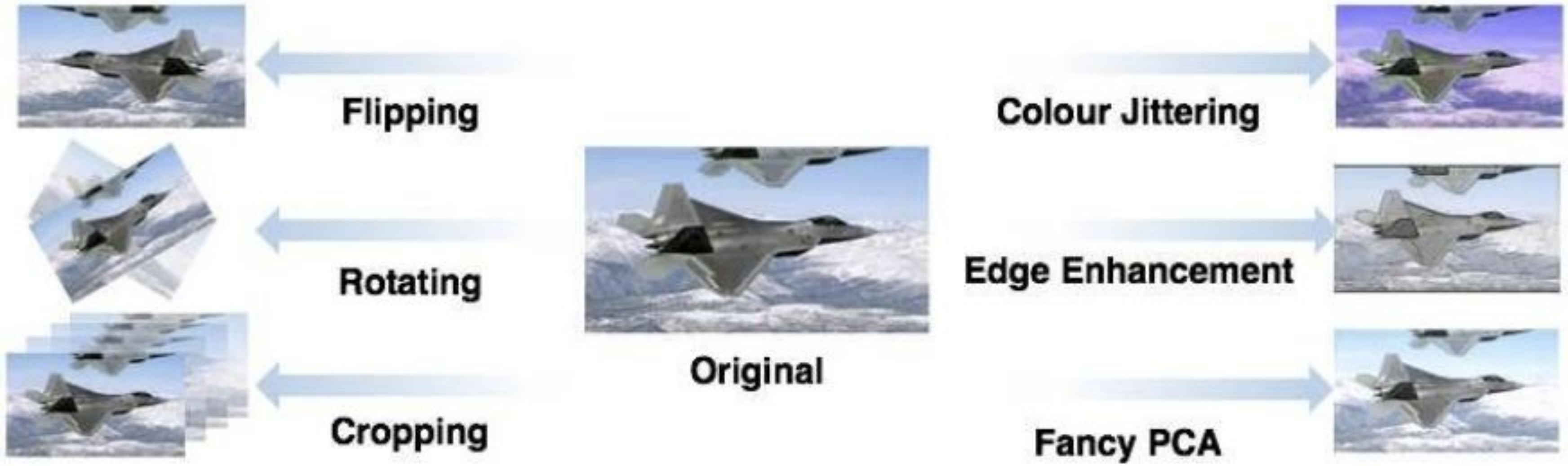
$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} cos(\Theta) & sin(\Theta) \\ -sin(\Theta) & cos(\Theta) \end{bmatrix} \begin{bmatrix} yx \end{bmatrix}$$

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & s_h \\ s_v & 1 \end{bmatrix} \begin{bmatrix} yx \end{bmatrix}$$

1. Масштабування
2. Обертання
3. Переміщення
4. Перекіс

$$(x_t,\ y_t) = T[(x,\ y)] = (c_x \cdot x,\ c_y \cdot y)$$

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} + \begin{bmatrix} t_x t_y \end{bmatrix}$$



Морфологічні операції

Морфологічні операції-це набір математичних функцій, відомих як нелінійні фільтри в обробці зображення, які обробляють зображення на основі морфології або форми.

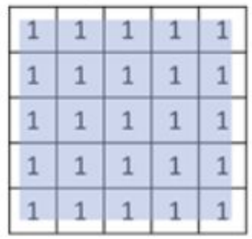
		A	B	C
fit	s_1	yes	no	no
	s_2	yes	yes	no
hit	s_1	yes	yes	yes
	s_2	yes	yes	no

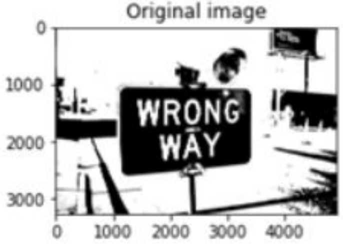
Fitting and hitting of a binary image with structuring elements s_1 and s_2.

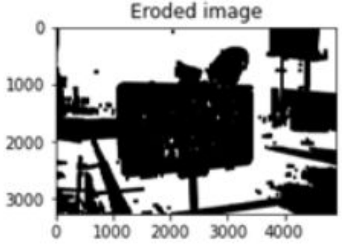
Ерозія (звуження)

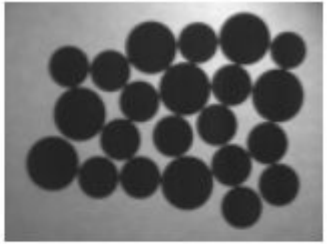
$g = f \ominus s$, $g(x,y) = 1$ is s fits f and 0 otherwise, repeating for all pixel coordinates (x,y).

Rectangular element

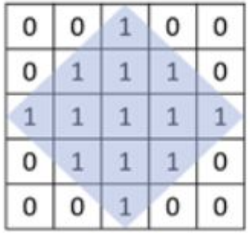


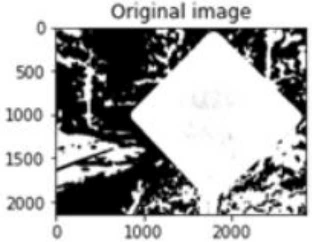


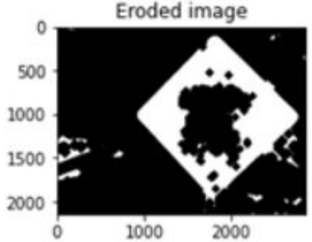


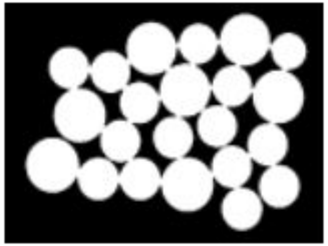


Diamond-shaped element

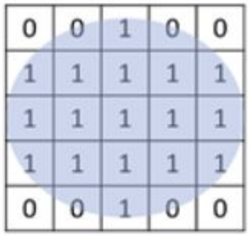


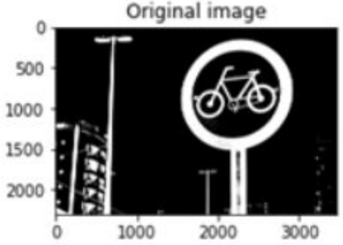


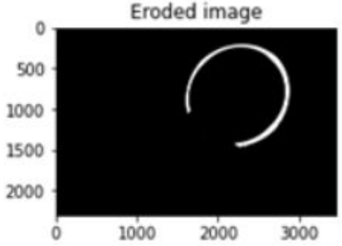


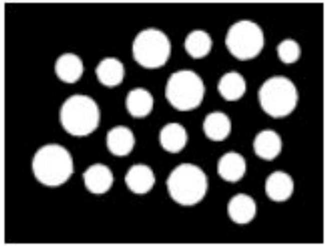


Elliptical element



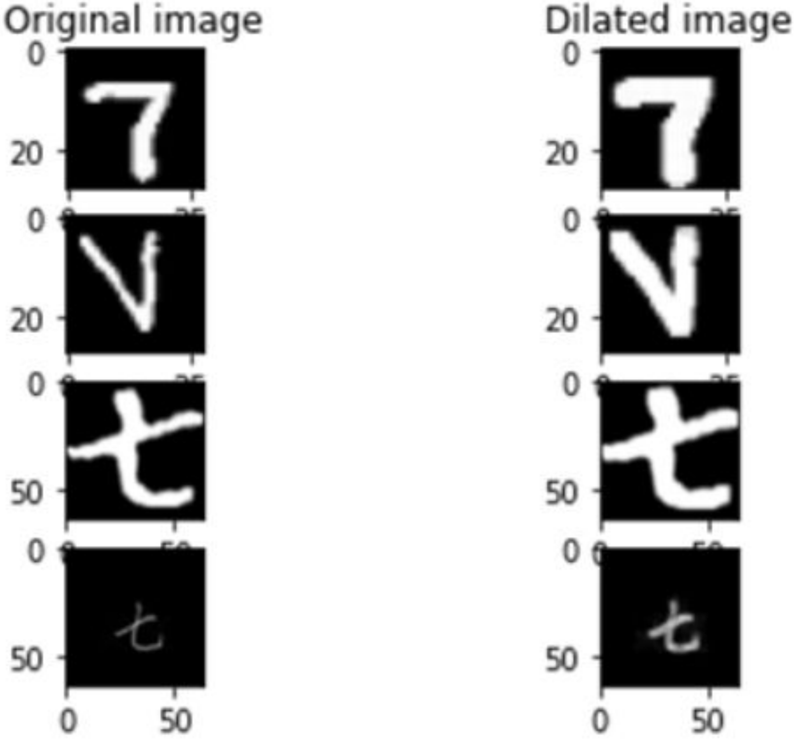




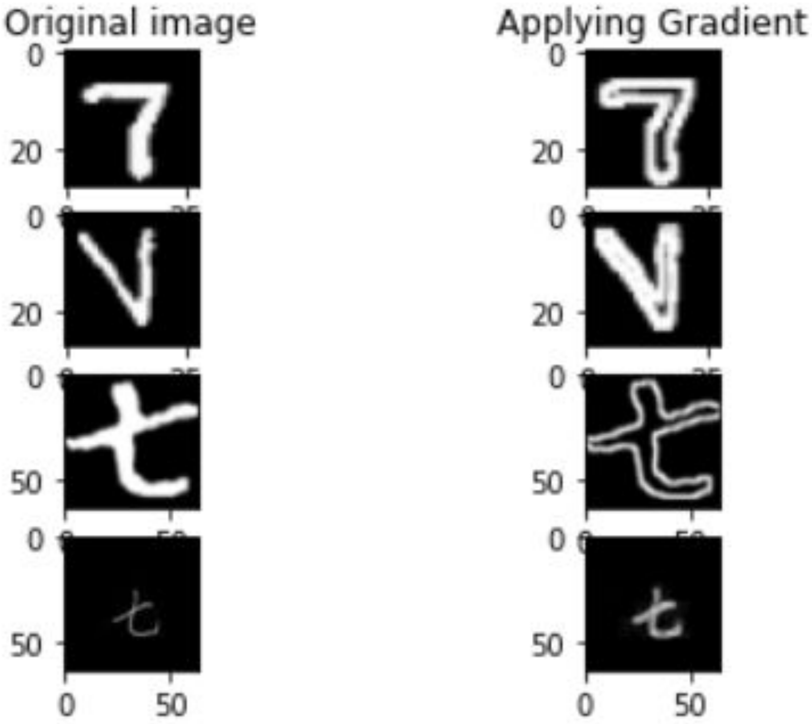


Розширення (розтягування)

g=
fo



Морфологічний градієнт



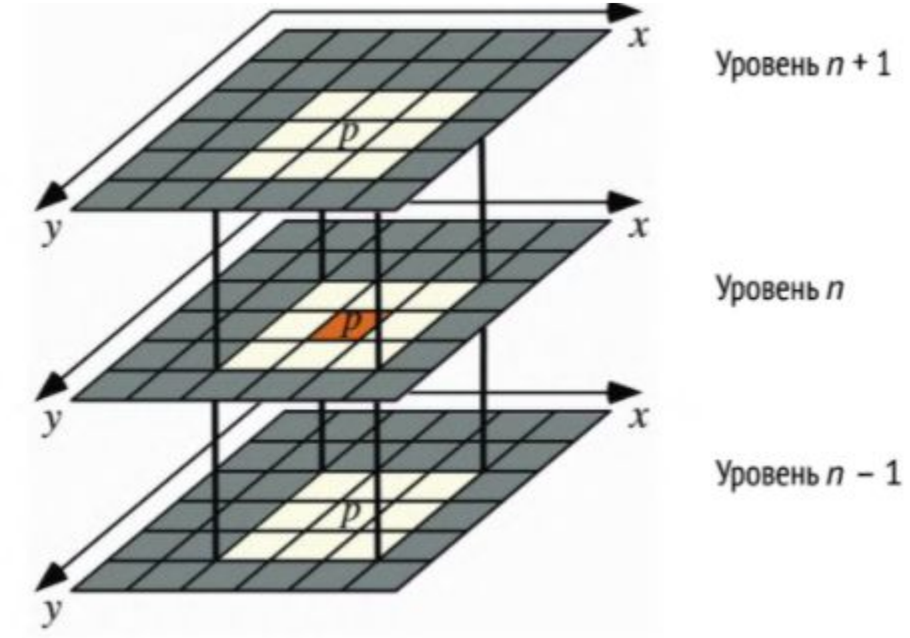
Як ми можемо виділити ознаки на зображенін?

Powered by Poll Everywhere

Особливі точки

- Особлива точка визначається характерною зміною яскравості оточуючих пікселів.

В позиції p(x,y) знаходиться особлива точка якщо існує рівень n, 0≤n≤m, такий що $D_{\sigma,a^n}(x, y)$ є локальним мінімумом або максимумом в трьохмірній околиці (x, y,n)

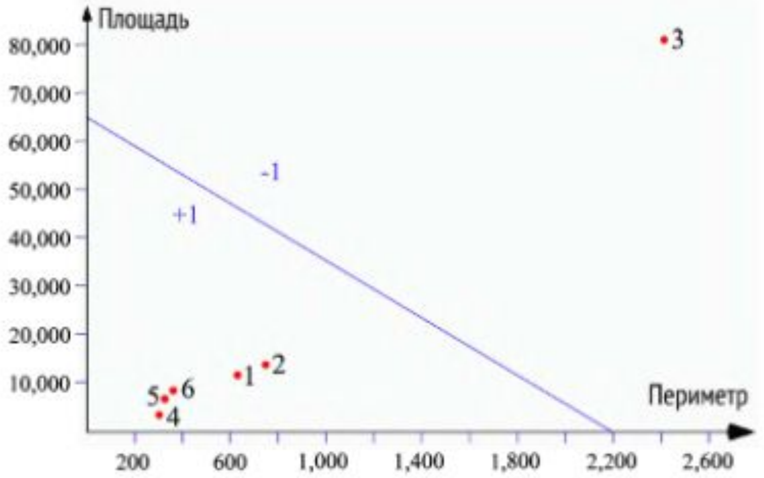


Дескриптор

Дескриптором x = (x1, ..., xn) називається точка в просторі R^n(простором дескрипторів), що представляє виміряні або обчислені значення властивостей в певному порядку.

Розглянемо випадок n=2. Наприклад, в цьому просторі дескрипторів з двома властивостями - «периметр» і «площа» - є дескриптор x1 = (621.605, 10 940), відповідний сегменту 1.

Разом особлива точка і дескриптор утворюють ознаку.



Дескриптор **SIFT** (scale-invariant feature transform)

1. Ключова точка шукаються як просторі DoG.

   Далі створюється дескриптор, який містить інформацію про візуальні характеристики навколо ключової точки, але не чутливий до обертання та зміни освітлення на зображенні.

2. Рахуємо модуль та напрямок градієнта

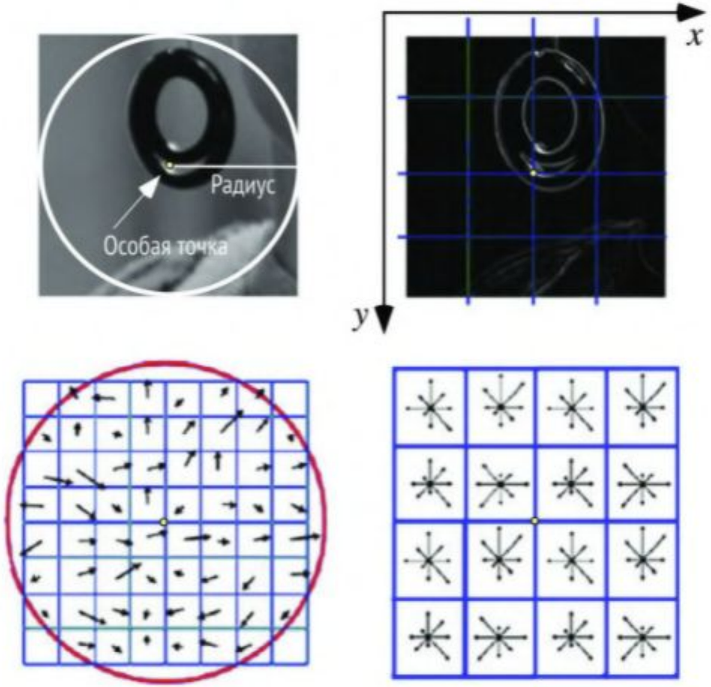
$$m(x, y) = \sqrt{[L(x, y+1) - L(x, y-1)]^2 + [L(x+1, y) - L(x-1, y)]^2},$$

$$\theta(x, y) = \text{atan } 2([L(x, y+1) - L(x, y-1)], [L(x+1, y) - L(x-1, y)])$$

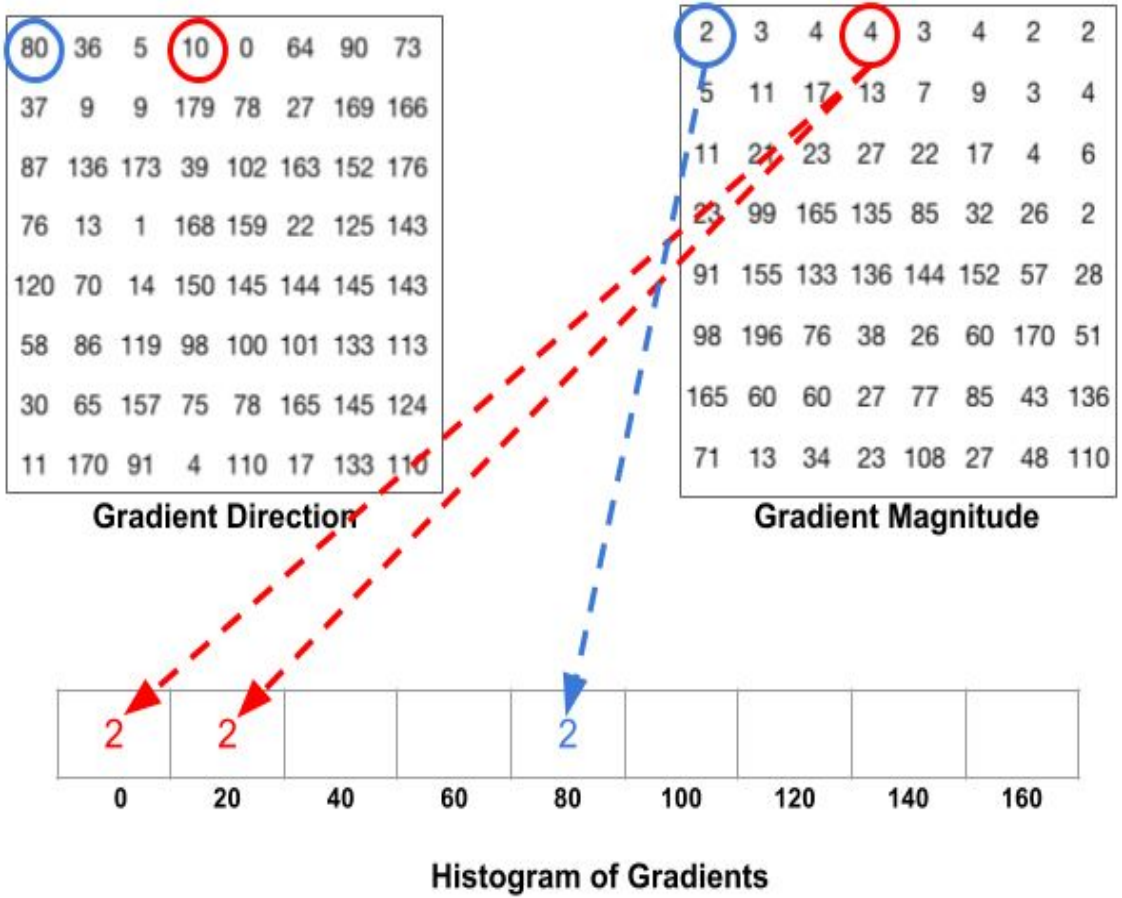
atan2(b, a) b/a

Дескриптор **SIFT**

3.Знаходимо масив гістограм - який і буде дескрптором **SIFT**



Дескриптор **SIFT**

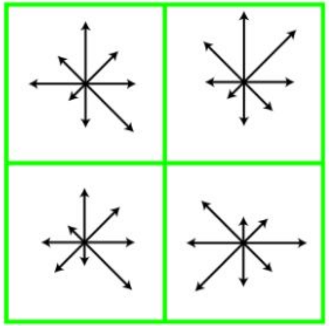


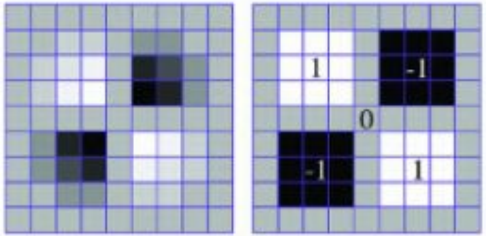
Що таке дескриптор?

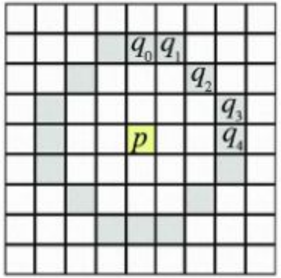
Дескриптор

- Дескриптор **SIFT** - scale-invariant

  feature transform (градієнт)
- Дескриптор **SURF -** speeded-up robust features (хаароподібні ознаки)
- Дескриптор **ORB** (oriented robust binary features)







supervised learning - є входи та правильні виходи, треба навчити систему співставляти входи на правильні виходи.

semi-supervised learning - є невеликий набір входів і правильних виходів і великий набір лише входів. Налаштовуємо систему.

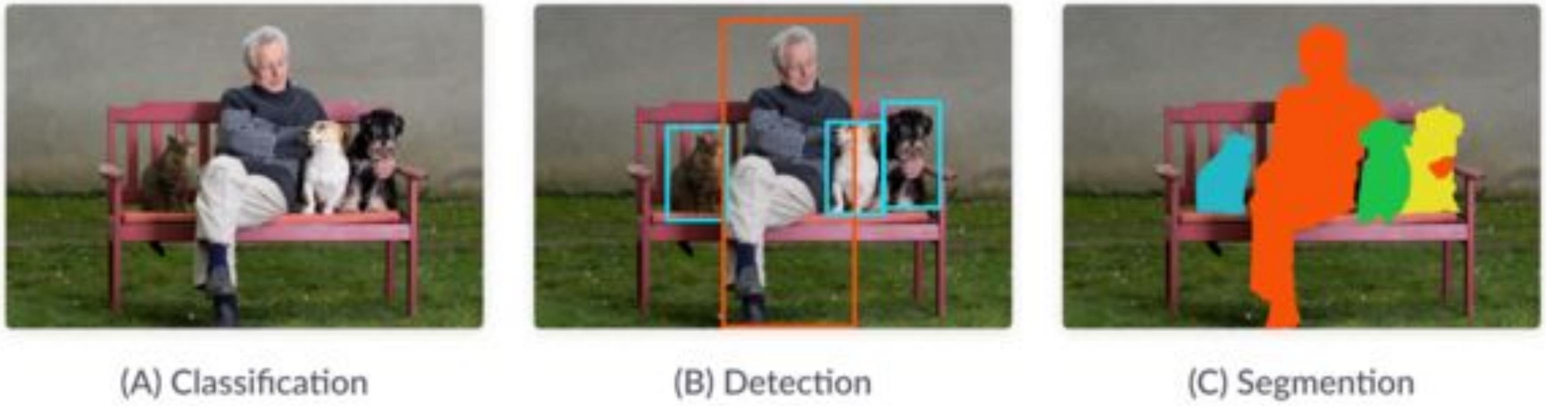
unsupervised learning - є вхідні дані, міток немає. Групуємо дані - кластеризація.

self-supervised learning - є вхідні дані, міток немає. Але намагаємося налаштувати систему.

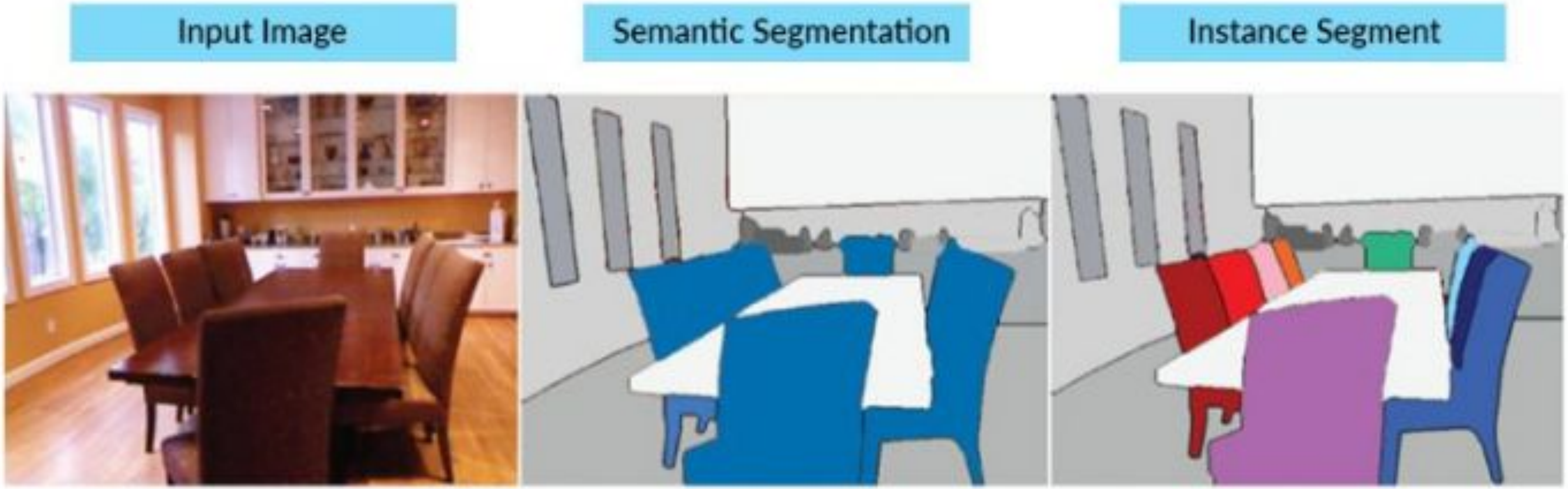
active learning - запитуємо користувача а як же ж позначити входи.

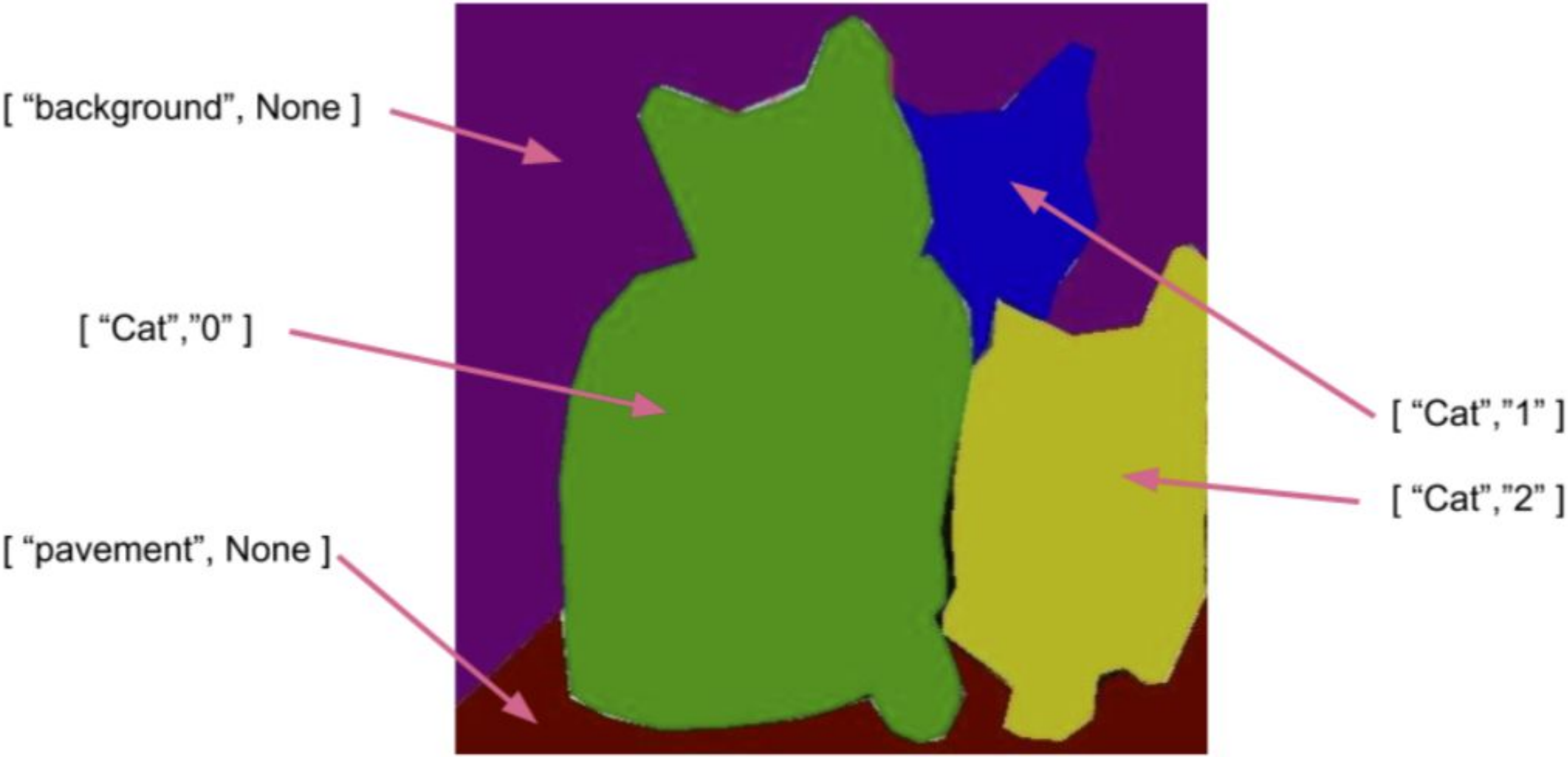
Розпізнавання образів

Сегментація зображень



Методи сегментації




[“background”, None]
[“Cat”,”0”]
[“pavement”, None]
[“Cat”,”1”]
[“Cat”,”2”]

Методи сегментації

- Підхід на основі виявлення подібності - виявлення подібних пікселів на зображенні
- Підхід на основі виявлення розривів (границь)

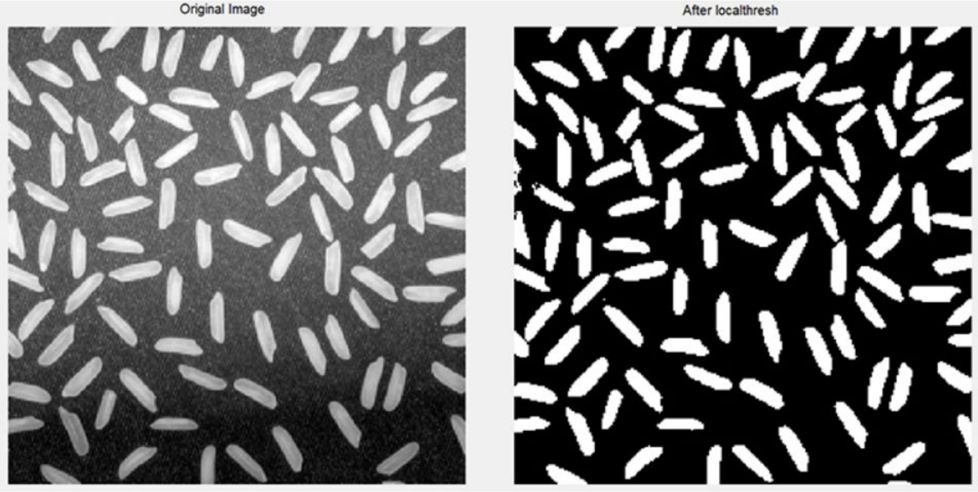
Edge VS Region Methods

Порогові методи

- Простий поріг

$$J(x, y) = \begin{cases} 0 & \text{якщо } I(x, y) < T \\ 1 & \text{інакше} \end{cases}$$

- Адаптивний поріг - бінаризація Оцу



Бінаризація Оцу

В алгоритмі Оцу вибирається поріг, який доставляє максимум міжкласової дисперсії σ_b^2, яка визначається як звичайна дисперсія, обчислена для середніх значень класів. У разі двох класів:

$$\sigma_b^2 = P_1(\mu_1 - \mu)^2 + P_2(\mu_2 - \mu)^2 = P_1P_2(\mu_1 - \mu_2)^2,$$

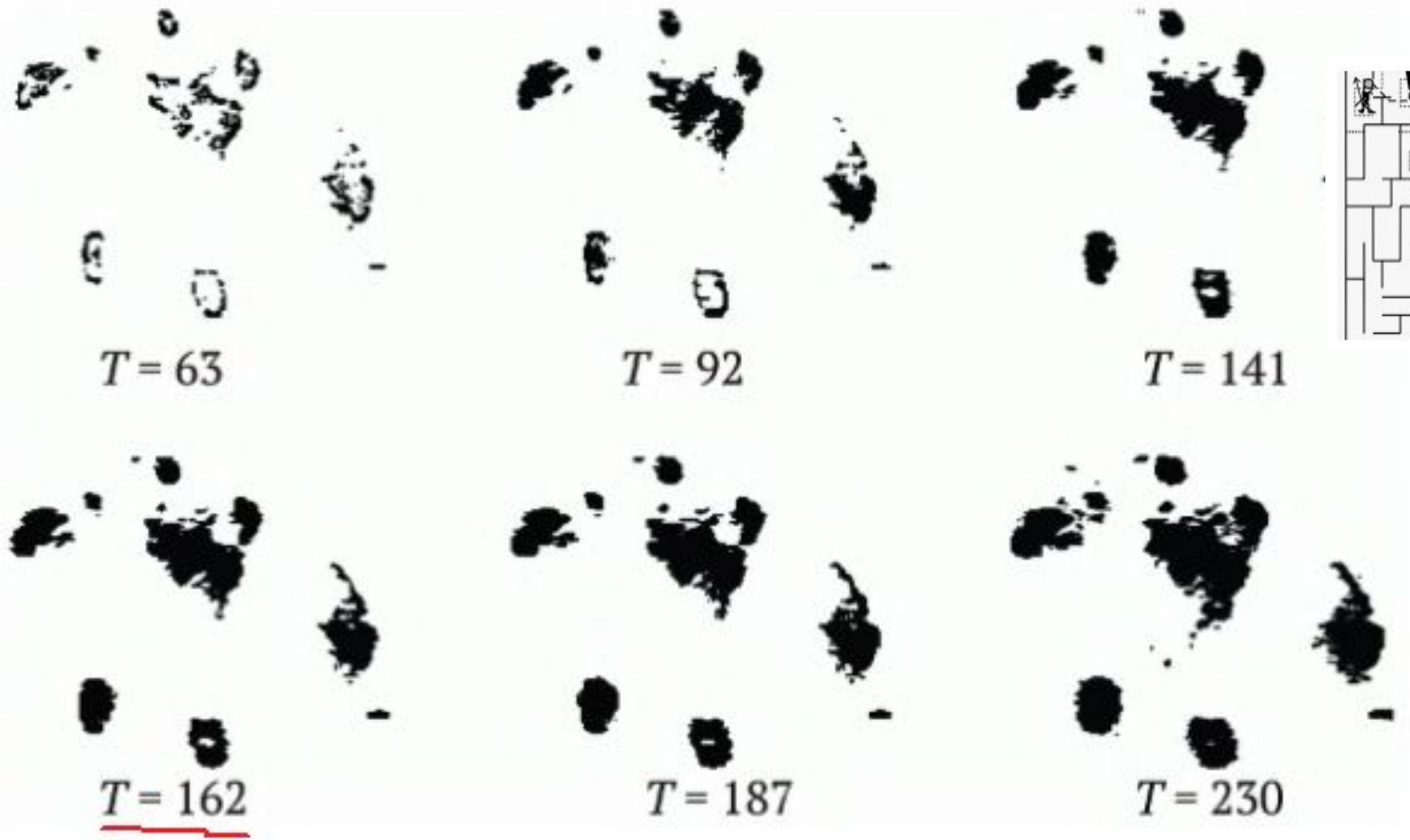
де P1 і P2- ймовірності класів.

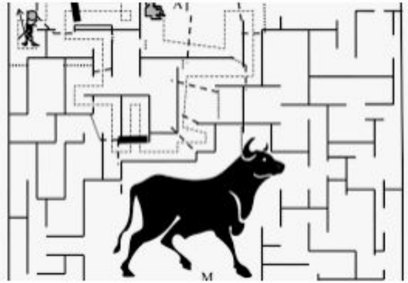
Обраний поріг u, 0 < u <Gmax, визначає «темні» пікселі об'єктів, для яких I(p) <= u, і «світлі» пікселі фону, для яких u <I(p).

Позначимо μ(i), i = 1, 2, середні по класах об'єктів і фону. Нехай C_I гістограма відносних кумулятивних частот зображення /.

$$c_I(u) = \sum_{v=0}^{u} h_I(v) \qquad h_I(u) = \frac{H_I(u)}{|\Omega|}.$$

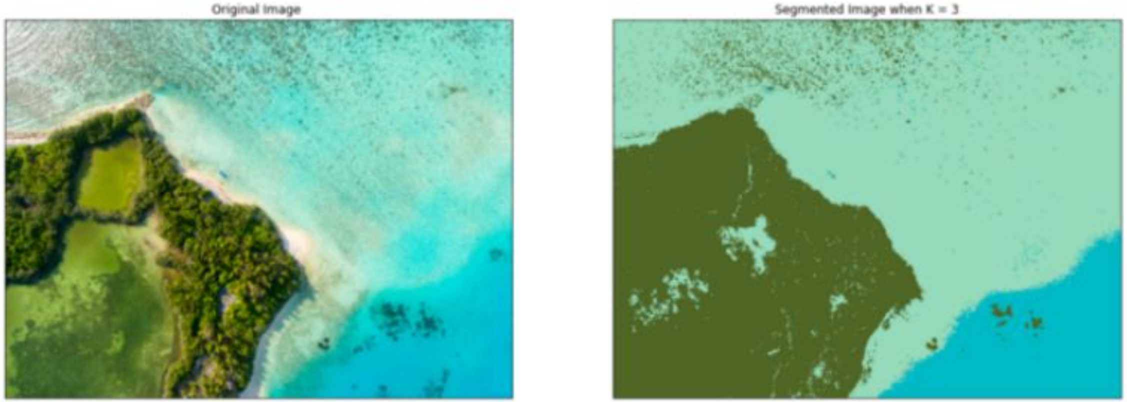
Ймовірності P1 і P2 апроксимуються значеннями з C_I (u) і 1 - C_I(u) відповідно; це загальне число пікселів в кожному класі, поділене на потужність множини | Ω|.

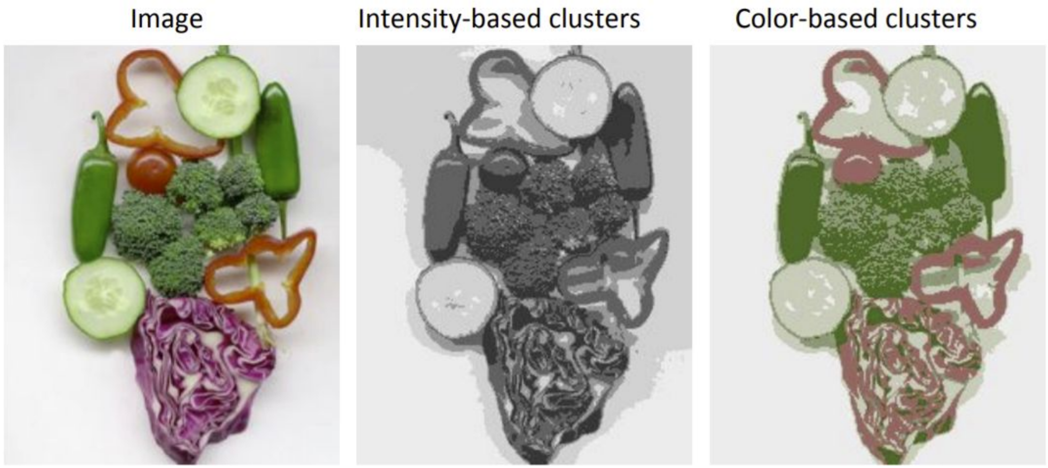

T = 63
T = 92
T = 141
T = 162
T = 187
T = 230



Методи кластеризації

- K-means, fuzzy k-means
- Spectral clustering
- Mean shift
- EM





EM

- Ініціалізація параметрів для кожного Ск кластера: мат.сподівання $\mu_к$, матриця коваріацій $\Sigma_к$, ваги wk=p(Ск)
- Е крок. Співставлення пікселів кластерам з ймовірністю: $p(C_к|x_к) = p(x_i|C_к) * p(C_к) / p(x_i)$

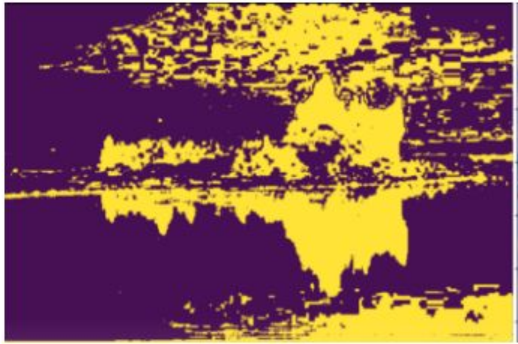
  где $p(x_i) = \sum_k p(C_k|x_i) * p(x_i)$

$$p(x_i|C_к) = \frac{1}{\sqrt{2\pi|\Sigma_k|}} e^{-\frac{1}{2}(V(x_i)-\mu_k)^T \Sigma^{-1}(V(x_i)-\mu_k)}$$

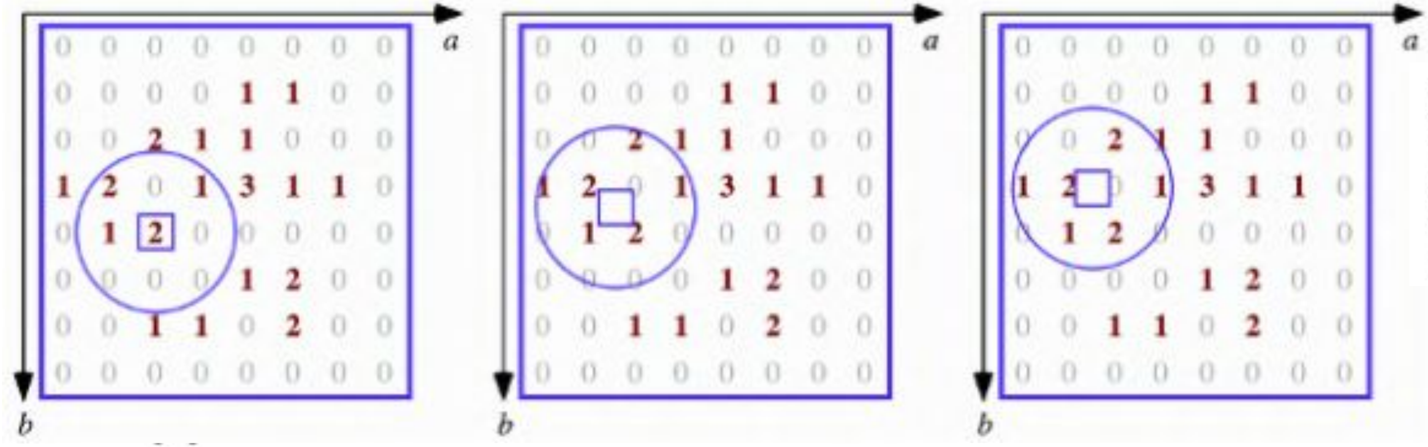
$$\mu_k = \frac{\sum_i p(C_k|x_i) * V(x_i)}{\sum_i p(C_k|x_i)}$$

- Оновлення параметрів

$$\Sigma_k = \frac{\sum_i p(C_k|x_i) * (V(x_i)-\mu_k)(V(x_i)-\mu_k)^T}{\sum_i p(C_k|x_i)}$$



Mean shift



[(5, 2), (6, 2), (3, 3), (3, 3), (4, 3), (5, 3), (1, 4), (2, 4), (2, 4), (4, 4), (5, 4), (5, 4), (5, 4), (6, 4), (7, 4), (2, 5), (3, 5), (3, 5), (5, 6), (6, 6), (6, 6), (3, 7), (4, 7), (6, 7), (6, 7)]

$$\mathbf{D} = \begin{bmatrix} 5\ 6\ 3\ 3\ 4\ 5\ 1\ 2\ 2\ 4\ 5\ 5\ 5\ 6\ 7\ 2\ 3\ 3\ 5\ 6\ 6\ 3\ 4\ 6\ 6 \\ 2\ 2\ 3\ 3\ 3\ 3\ 4\ 4\ 4\ 4\ 4\ 4\ 4\ 4\ 4\ 5\ 5\ 5\ 6\ 6\ 6\ 7\ 7\ 7\ 7 \end{bmatrix}. \qquad \mathbf{M} = \begin{bmatrix} 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3\ 3 \\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5 \end{bmatrix}.$$

$$\mathbf{D} - \mathbf{M} \Rightarrow \mathbf{E} = \begin{bmatrix} 4\ 9\ 0\ 0\ 1\ 4\ 4\ 1\ 1\ 1\ 4\ 4\ 4\ 9\ 16\ 1\ 0\ 0\ 4\ 9\ 9\ 0\ 1\ 9\ 9 \\ 9\ 9\ 4\ 4\ 4\ 4\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 1\ 1\ 4\ 4\ 4\ 4 \end{bmatrix}.$$

[13 18 4 4 5 8 5 2 2 2 5 5 5 10 17 1 0 0 5 10 10 4 5 13 13].

$S_{<r\text{^}2} = \{(2, 4), (2, 4), (4, 4), (2, 5), (3, 5), (3, 5)\}$ $\quad r^2 = 2{,}56.$ $\qquad \mathbf{u}_S = (2{,}67, 4{,}5)$