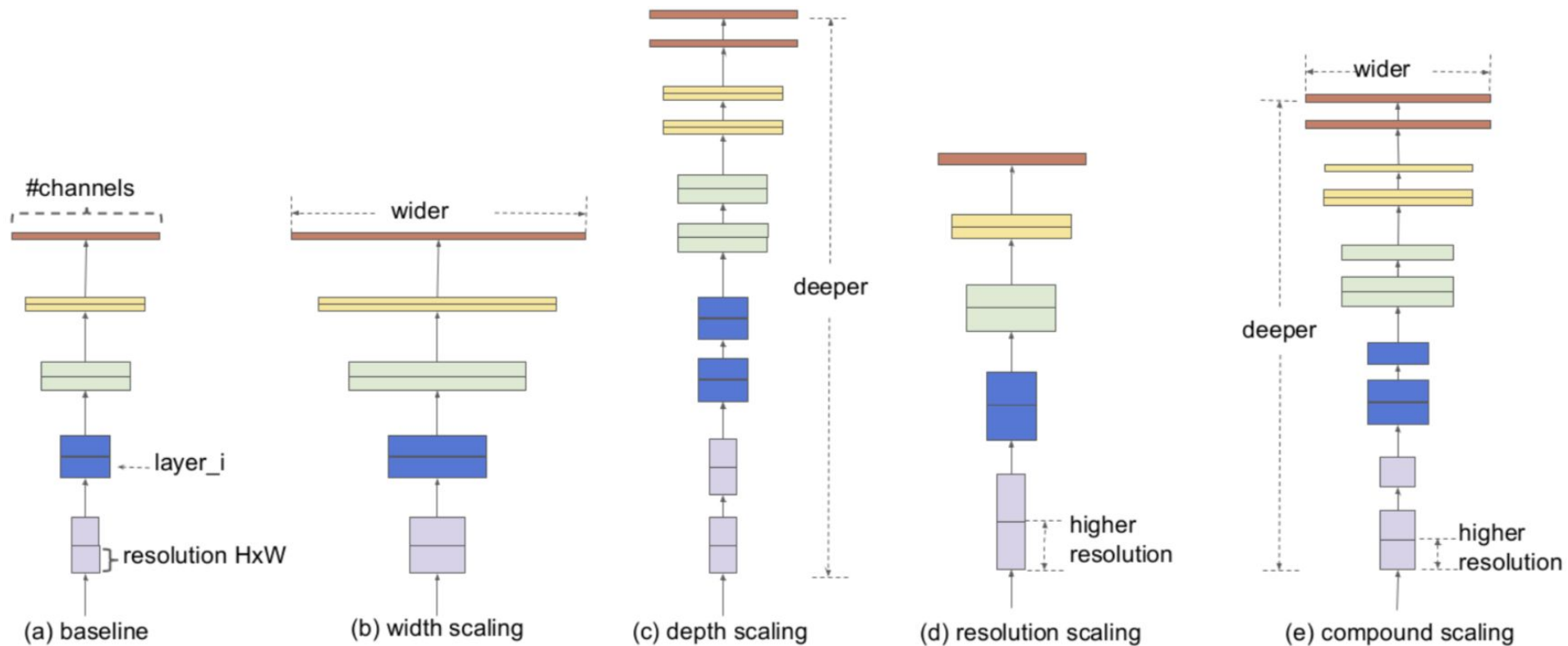


EfficientNet



EfficientNet

$$\mathcal{N} = \bigodot_{i=1\dots s} \mathcal{F}_i^{L_i} (X_{\langle H_i, W_i, C_i \rangle})$$

$$\mathcal{N}(d, w, r) = \bigodot_{i=1\dots s} \hat{\mathcal{F}}_i^{d \cdot \hat{L}_i} (X_{\langle r \cdot \hat{H}_i, r \cdot \hat{W}_i, w \cdot \hat{C}_i \rangle})$$

$$\text{depth: } d = \alpha^\phi$$

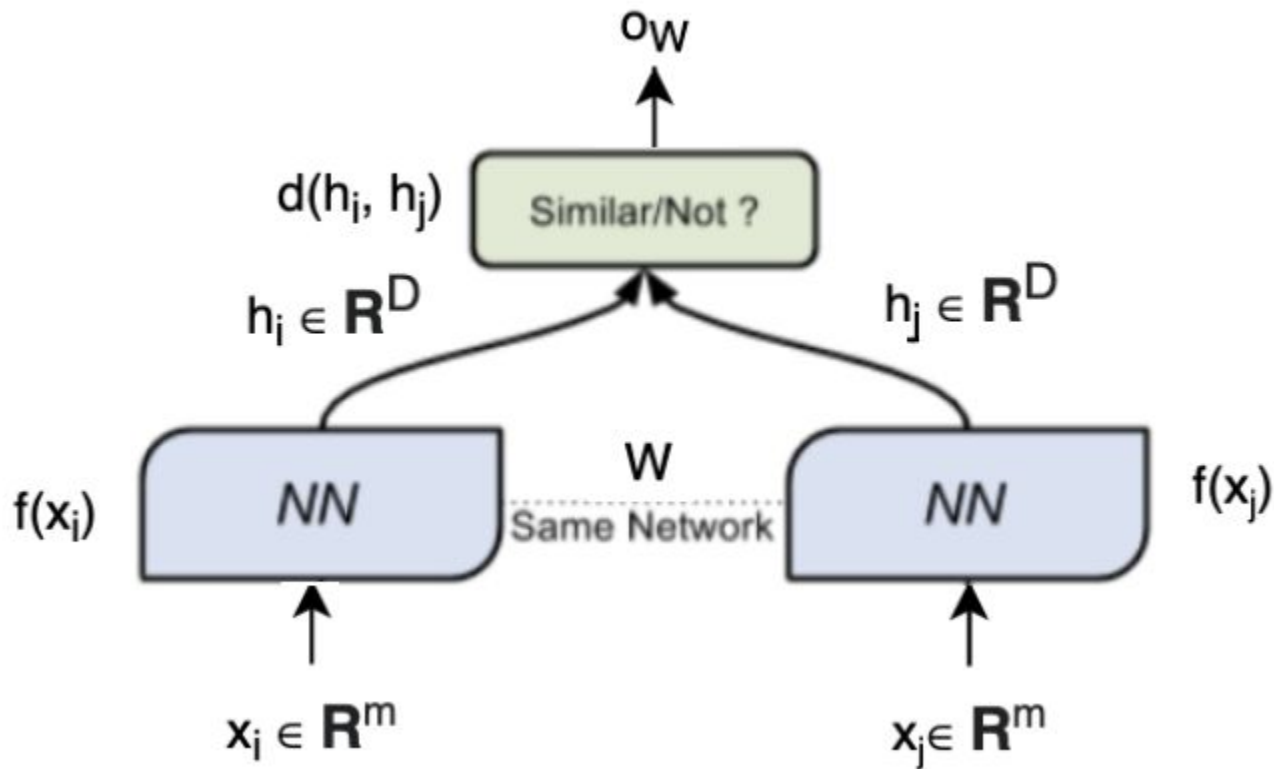
$$\text{width: } w = \beta^\phi$$

$$\text{resolution: } r = \gamma^\phi$$

$$\text{s.t. } \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$$

$$\alpha \geq 1, \beta \geq 1, \gamma \geq 1$$

Siamese Network



Siamese Network

$$S = \{(x_i, x_j, z_{ij}), i = 1, \dots, n; j = 1, \dots, n\},$$

$$z_{ij} = 0, \quad \text{при} \quad y_i = y_j$$

$$z_{ij} = 1, \quad \text{при} \quad y_i \neq y_j$$

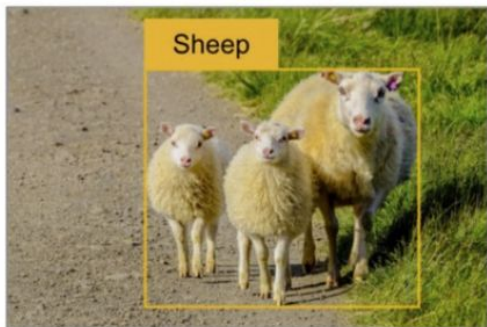
contrastive loss function

$$l(x_i, x_j, z_{ij}) = (1 - z_{ij})||h_i - h_j||_2^2 + z_{ij} \max(0, \tau - ||h_i - h_j||_2^2), \text{ де } \tau — \text{ заздалегіть задане}$$

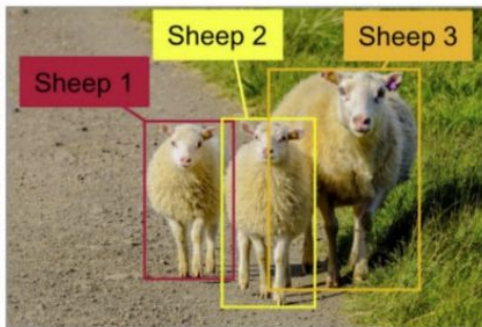
$$l(x_i, x_j, x_k) = \max(0, ||h_i - h_j||_2^2 - ||h_i - h_k||_2^2 + \alpha), \text{ де } \alpha — \text{ заздалегіть задане}$$

Системи розпізнавання образів

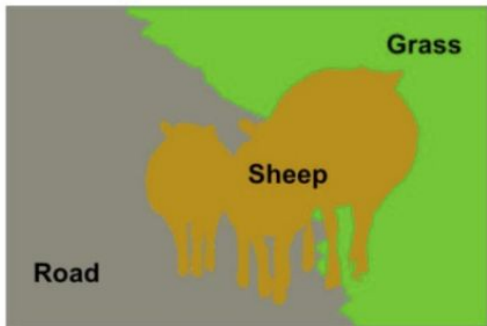
Згорткові нейронні мережі для задачі сегментації



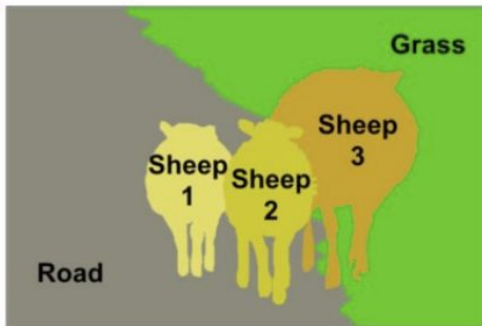
Classification + Localization



Object Detection

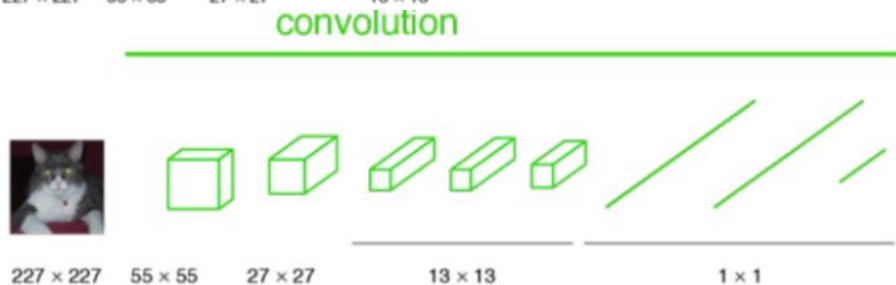
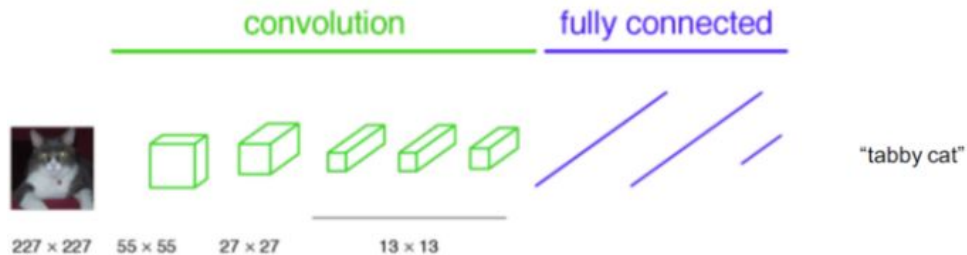


Semantic Segmentation

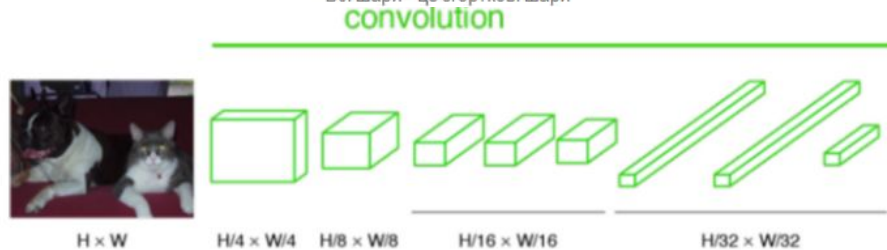


Instance Segmentation

Fully convolutional network



Всі шари - це згорткові шари

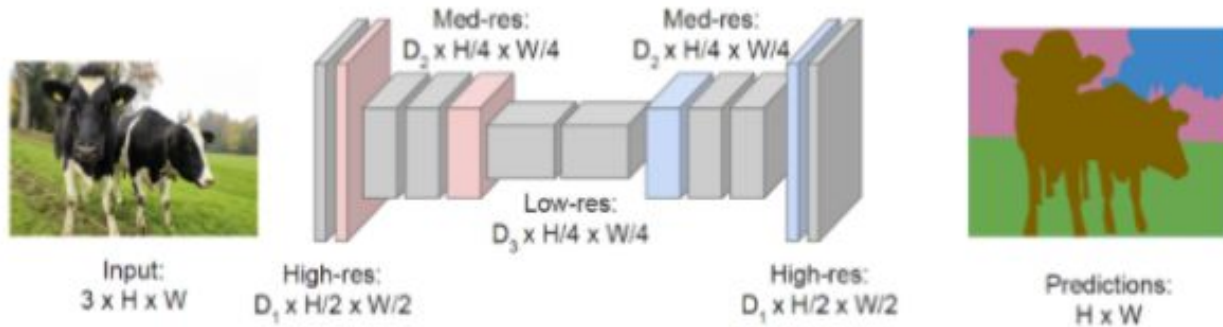
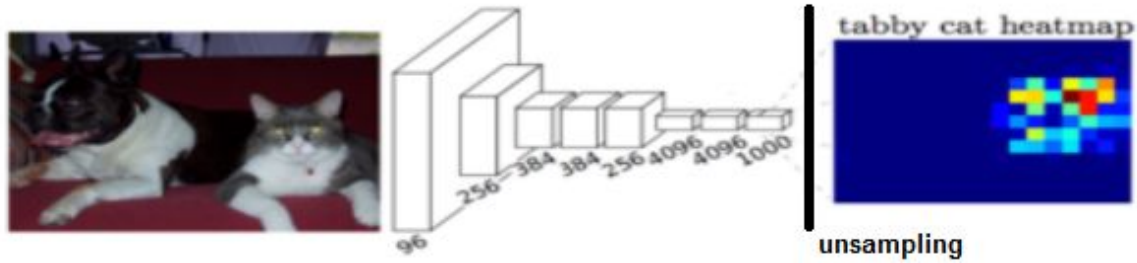


Всі шари - це згорткові шари

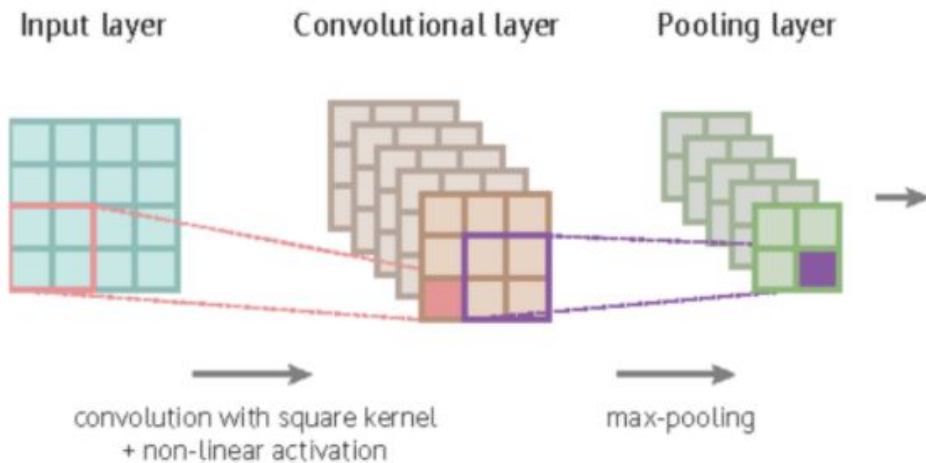
tabby cat heatmap



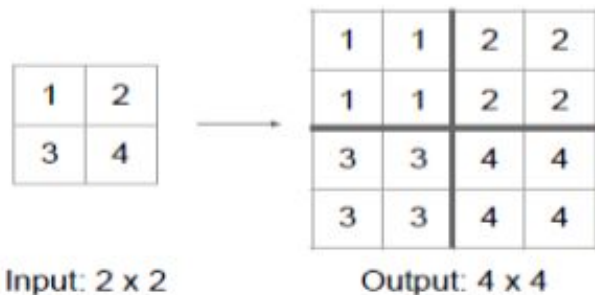
Fully convolutional network



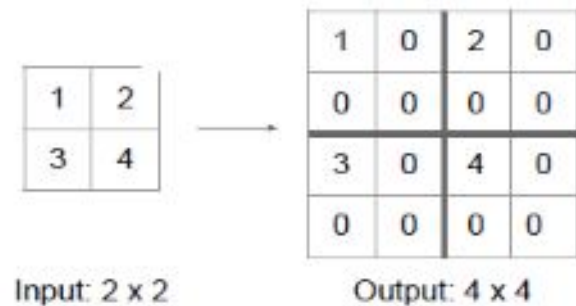
Upsampling



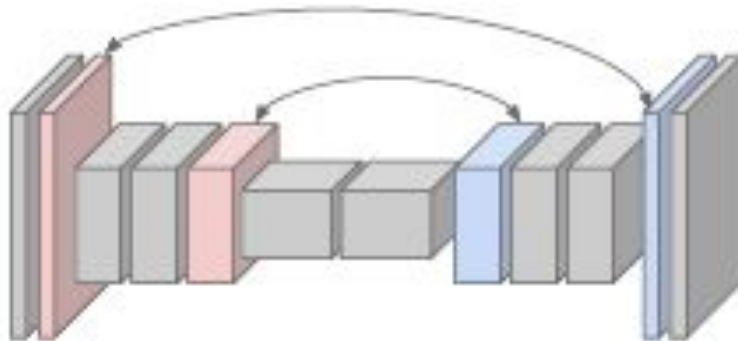
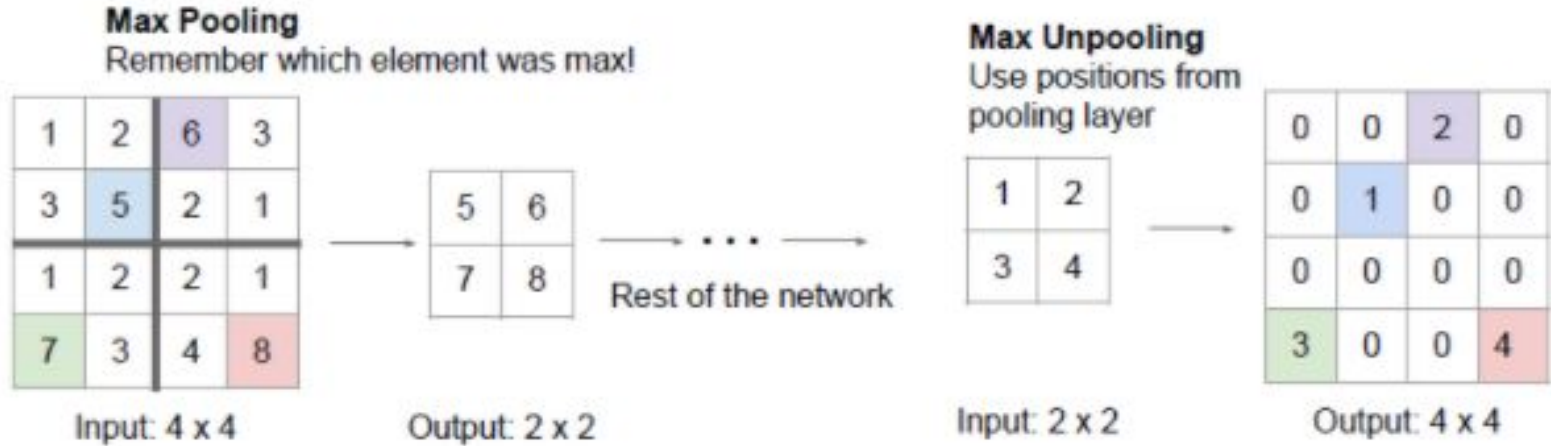
Метод найближчого сусіда



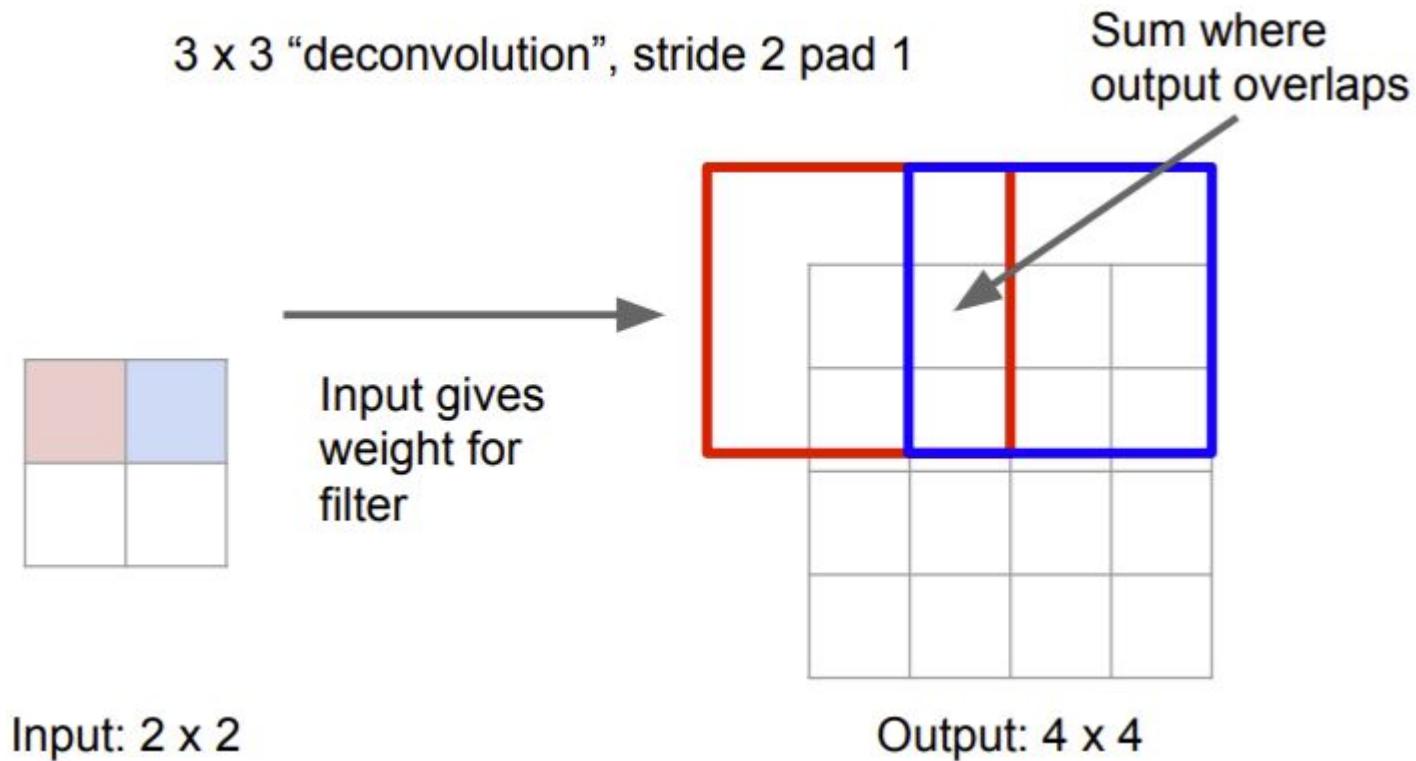
Bed of nails

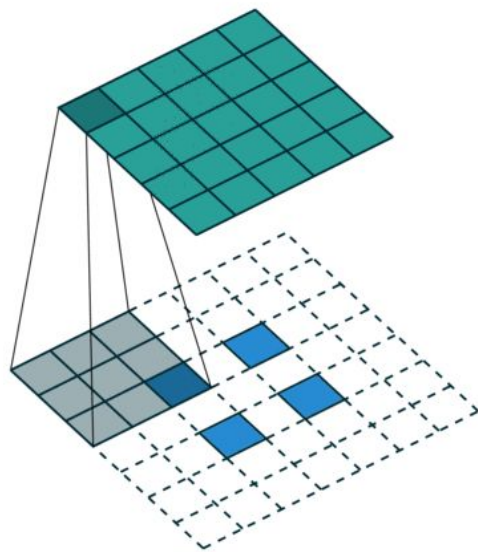
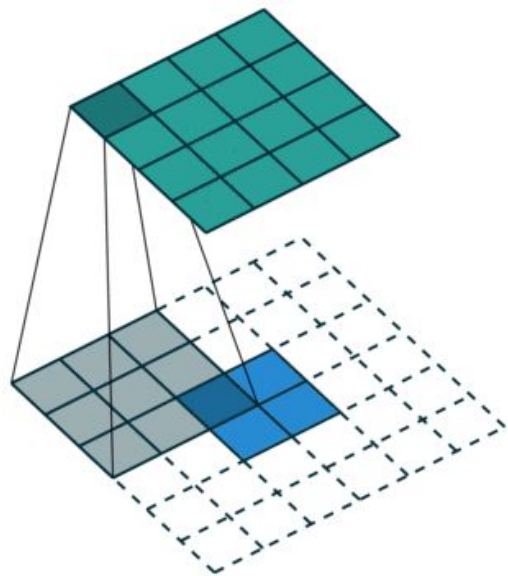


Unpooling

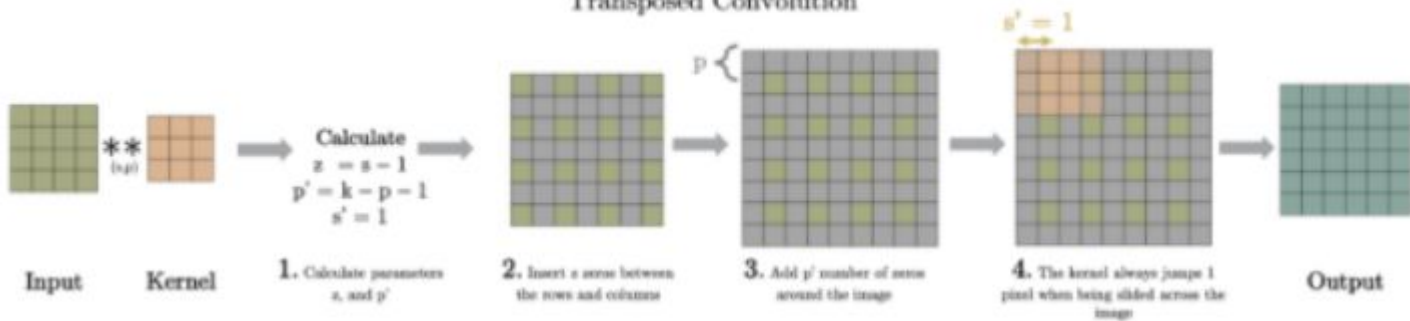


Транспонована згортка

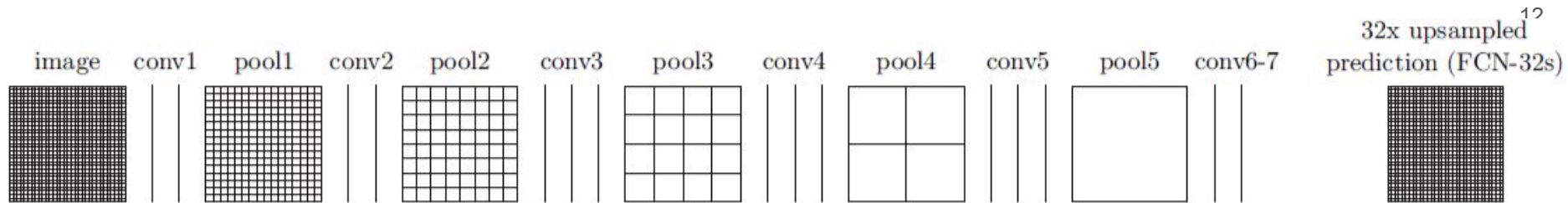




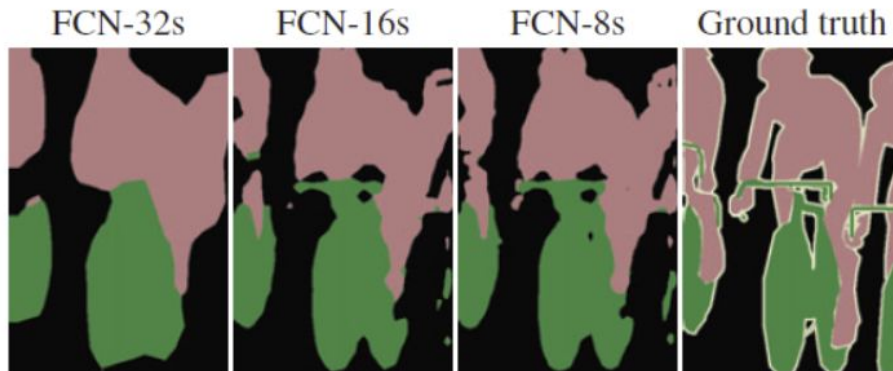
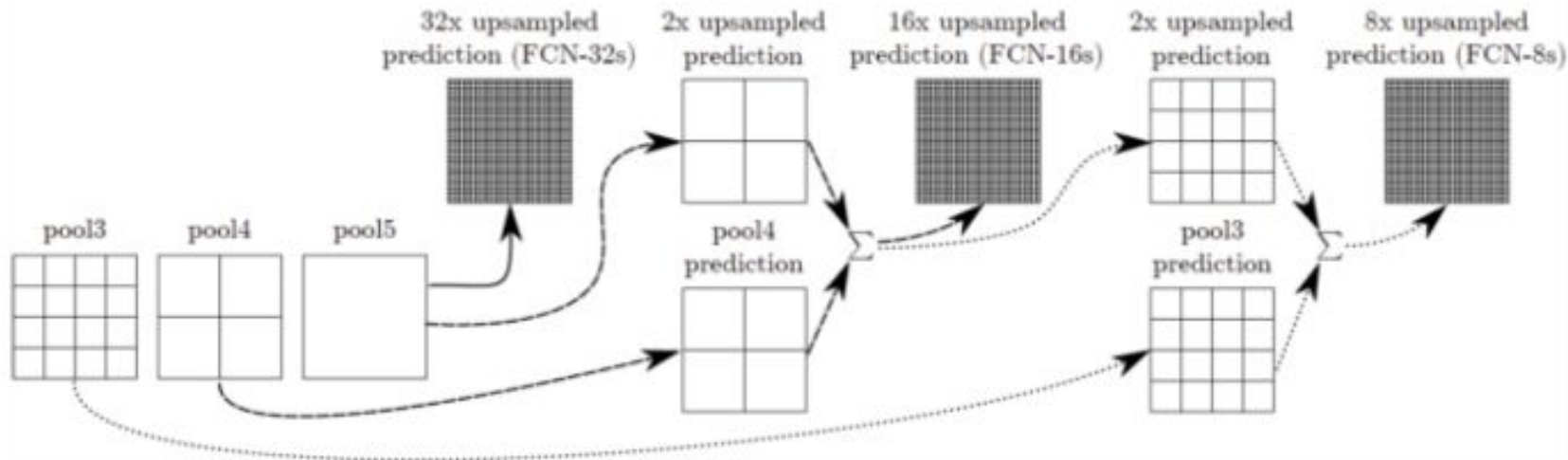
Transposed Convolution

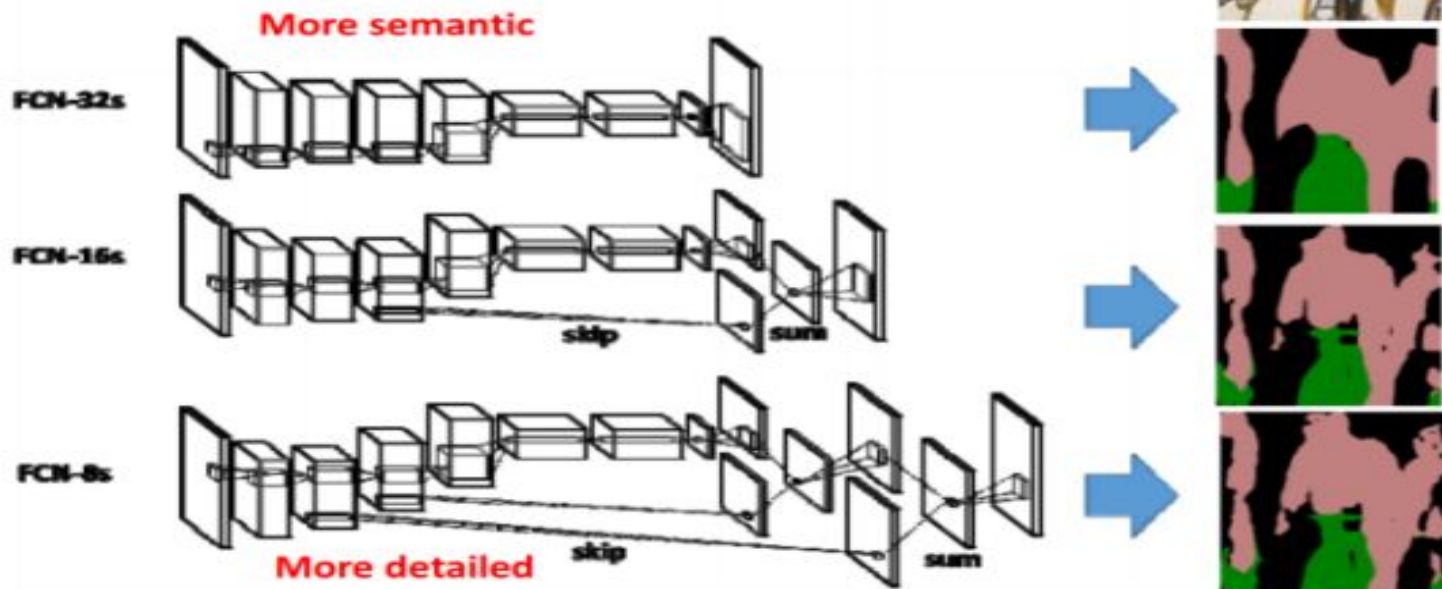


Fully convolutional network



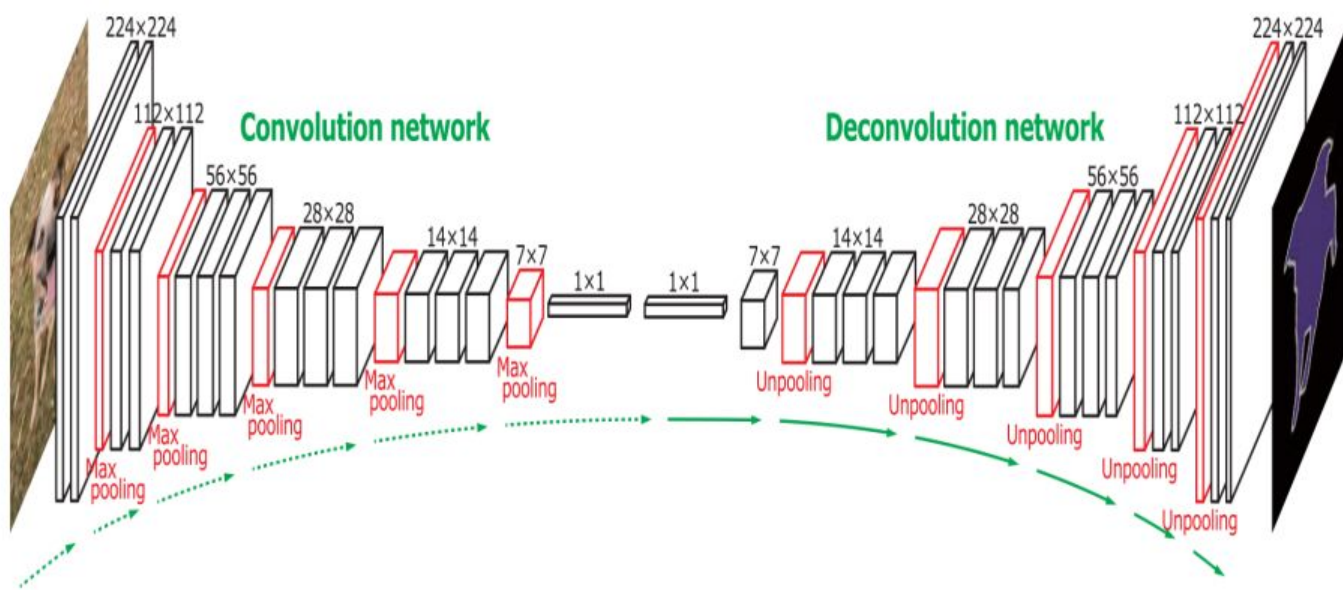
Skip connections



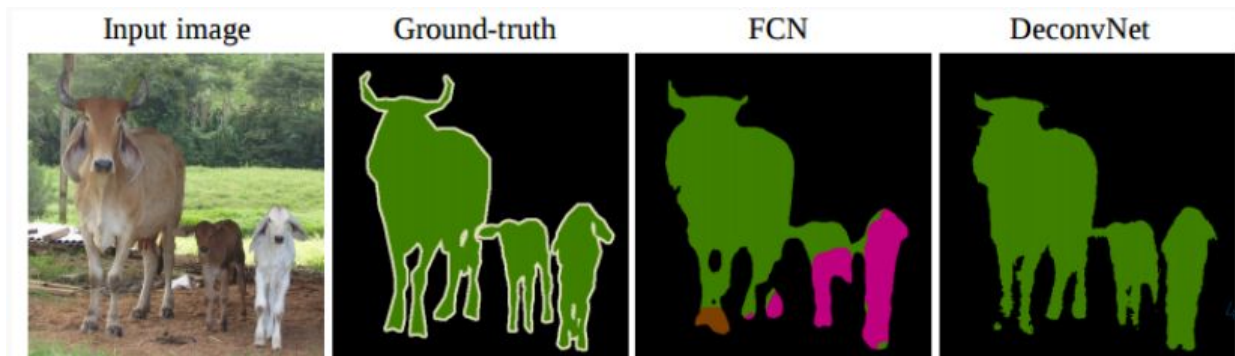


**Які ідеї застосовувалися для того щоб виконати
сегментацію зображення, а не просто його
класифікувати?**

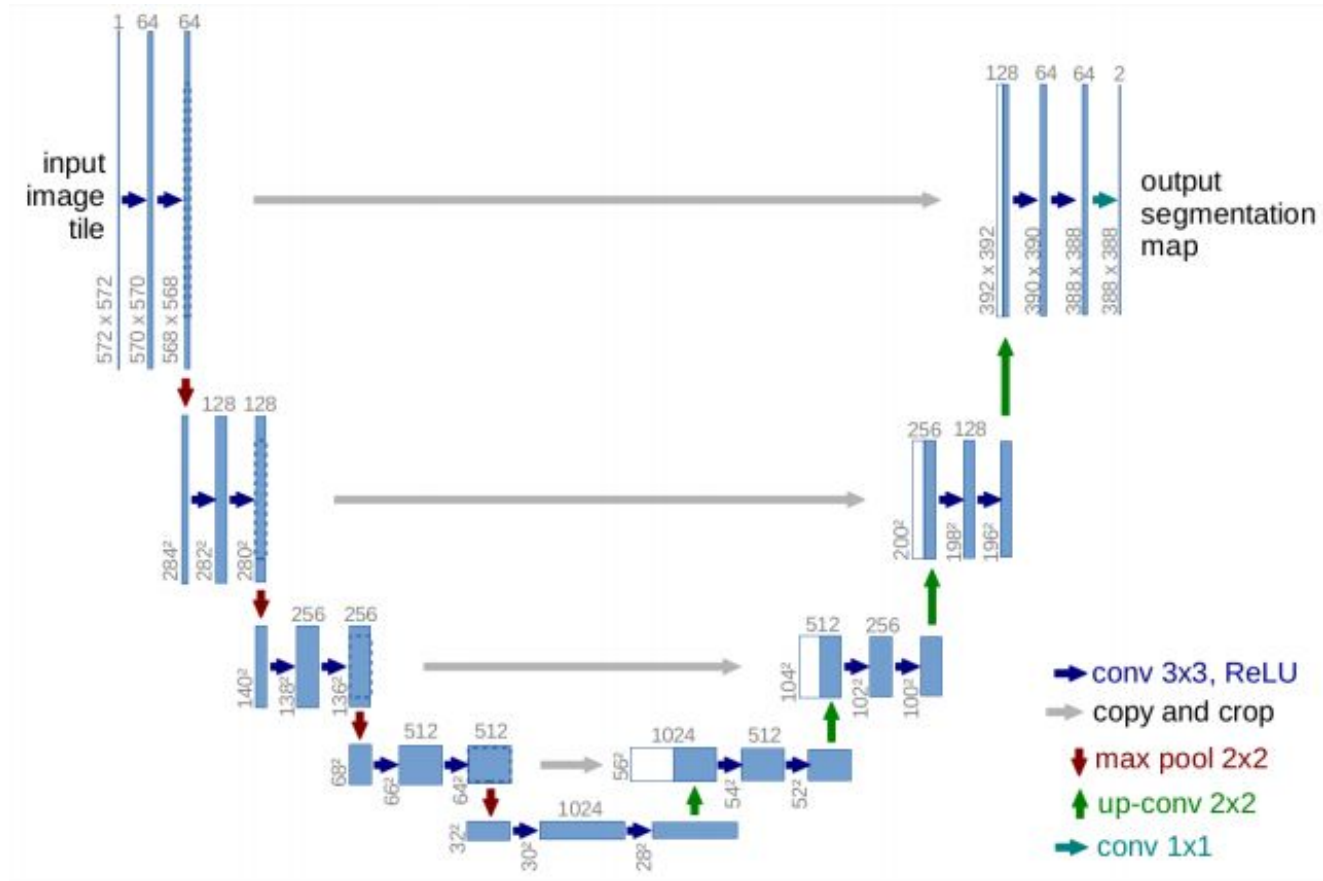




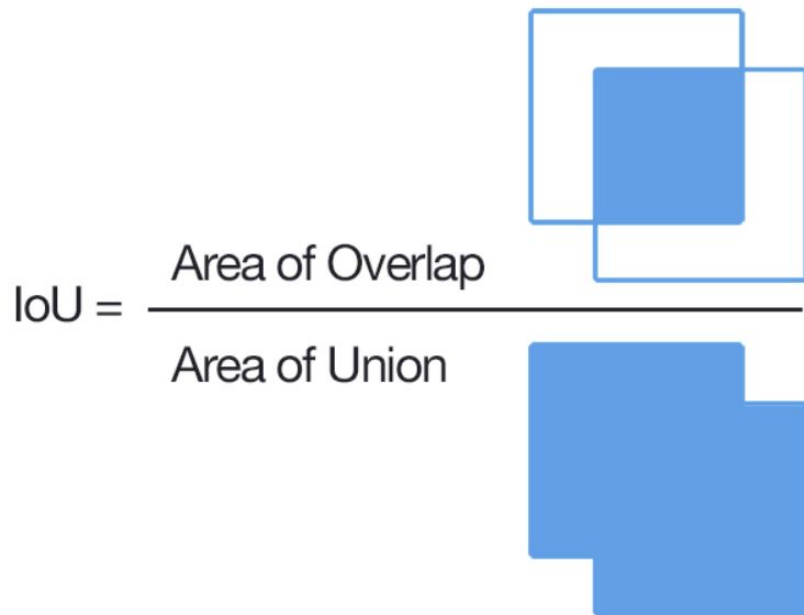
VGG 16



U-Net



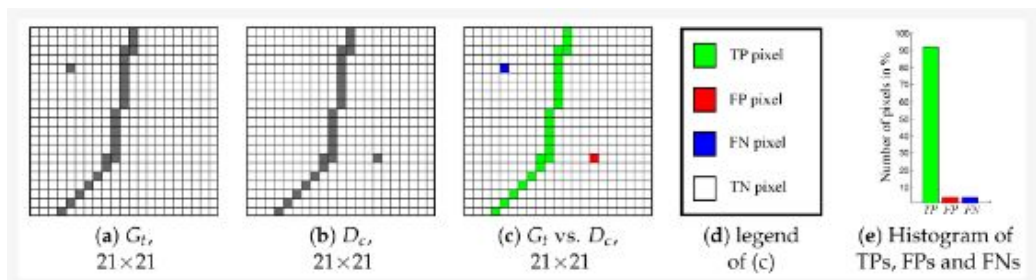
Оцінка сегментації



$$Jaccard = \frac{TP}{TP + TN + FP}$$

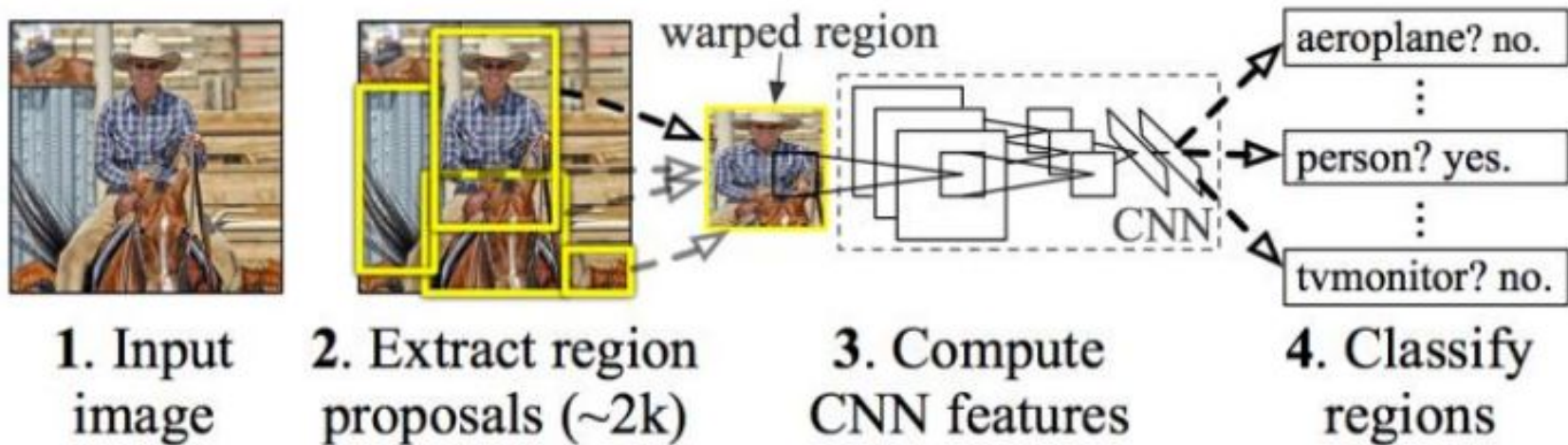
$$FL(p_t) = -(1 - p_t)^\gamma \log(p_t)$$

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

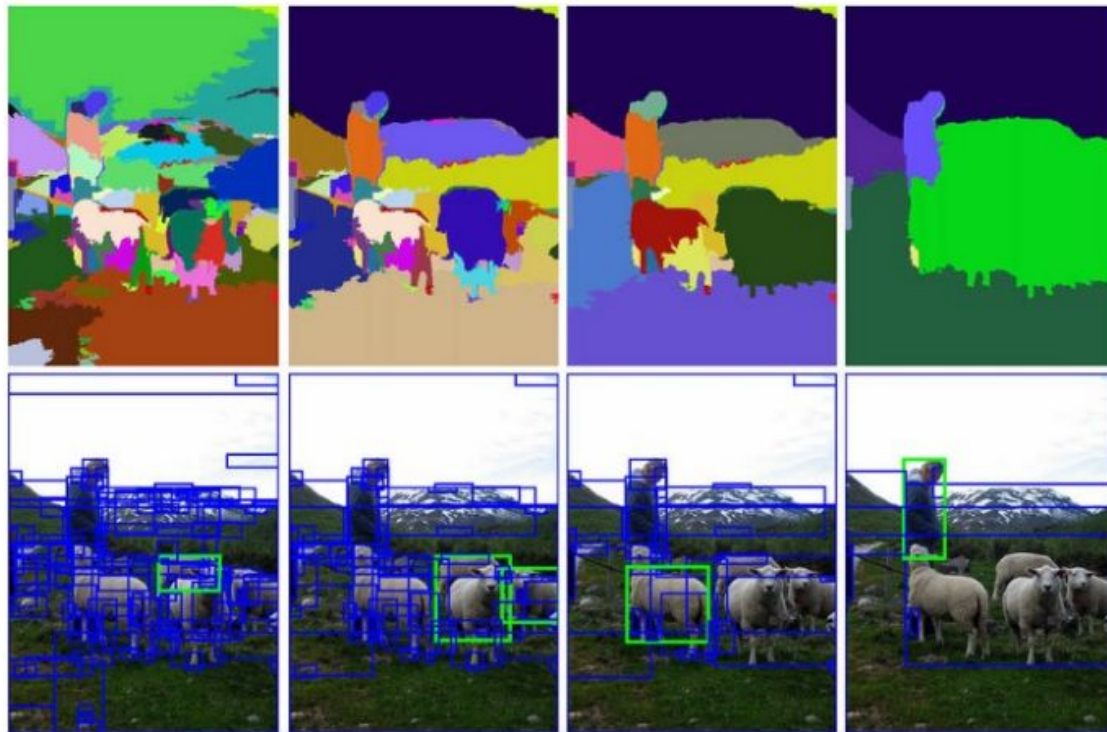


R-CNN (Region-Based Convolutional Neural Network)

R-CNN: *Regions with CNN features*



Модуль пропозиції регіонів

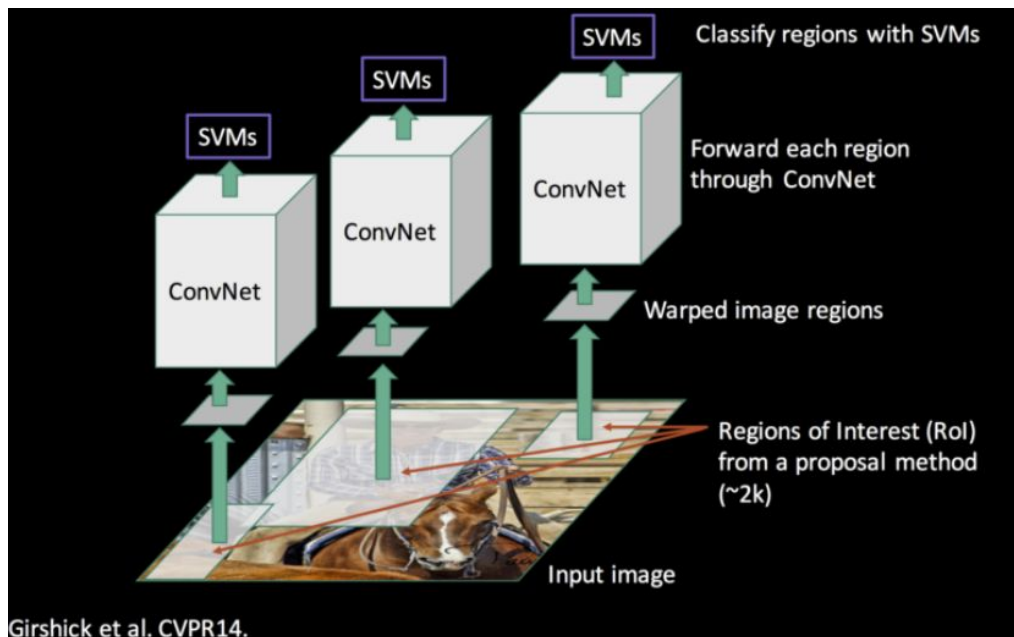


Вибірковий пошук:

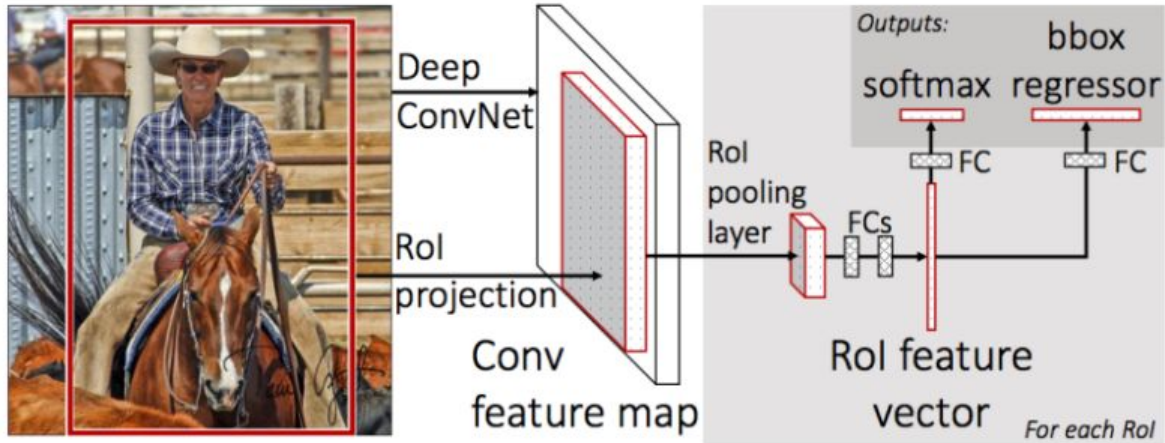
1. Генеруємо початкову субсегментацію, генеруємо багато регіонів-кандидатів
2. Використовуйте жадібний алгоритм для рекурсивного об'єднання подібних регіонів у більші
3. Використовуйте згенеровані регіони для створення пропозицій щодо остаточного регіону-кандидата

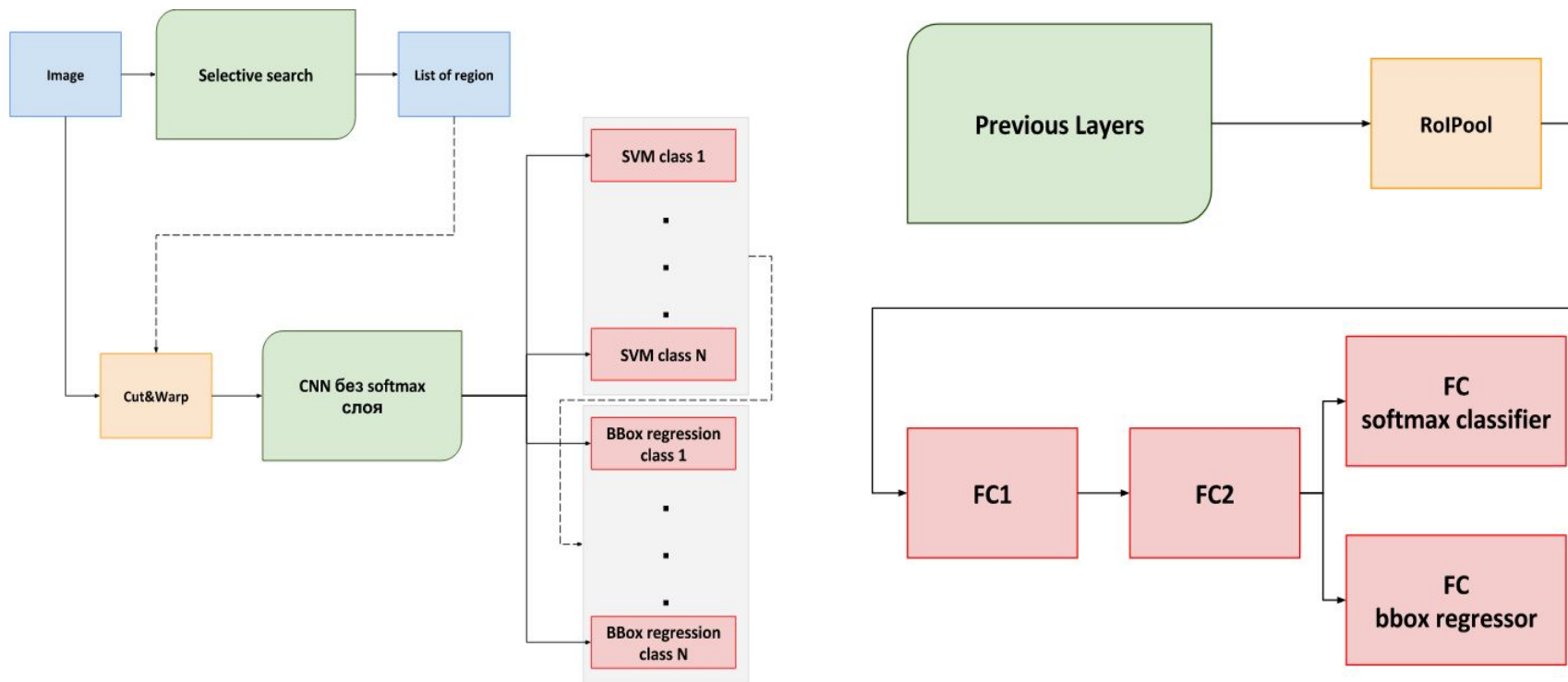
R-CNN

- 2000 пропозицій регіонів-кандидатів
- AlexNet створює 4096-мірний вихід
- вилучені ознаки подаються у SVM класифікатори
- Передбачення обмежувальної рамки



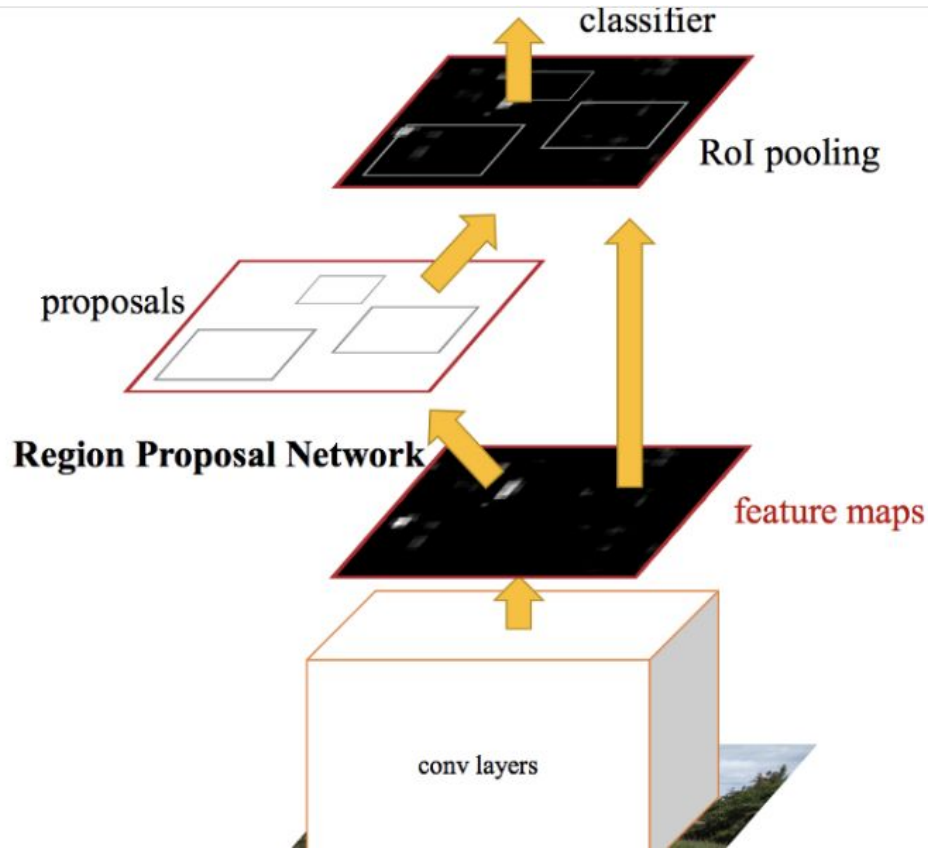
Fast R-CNN

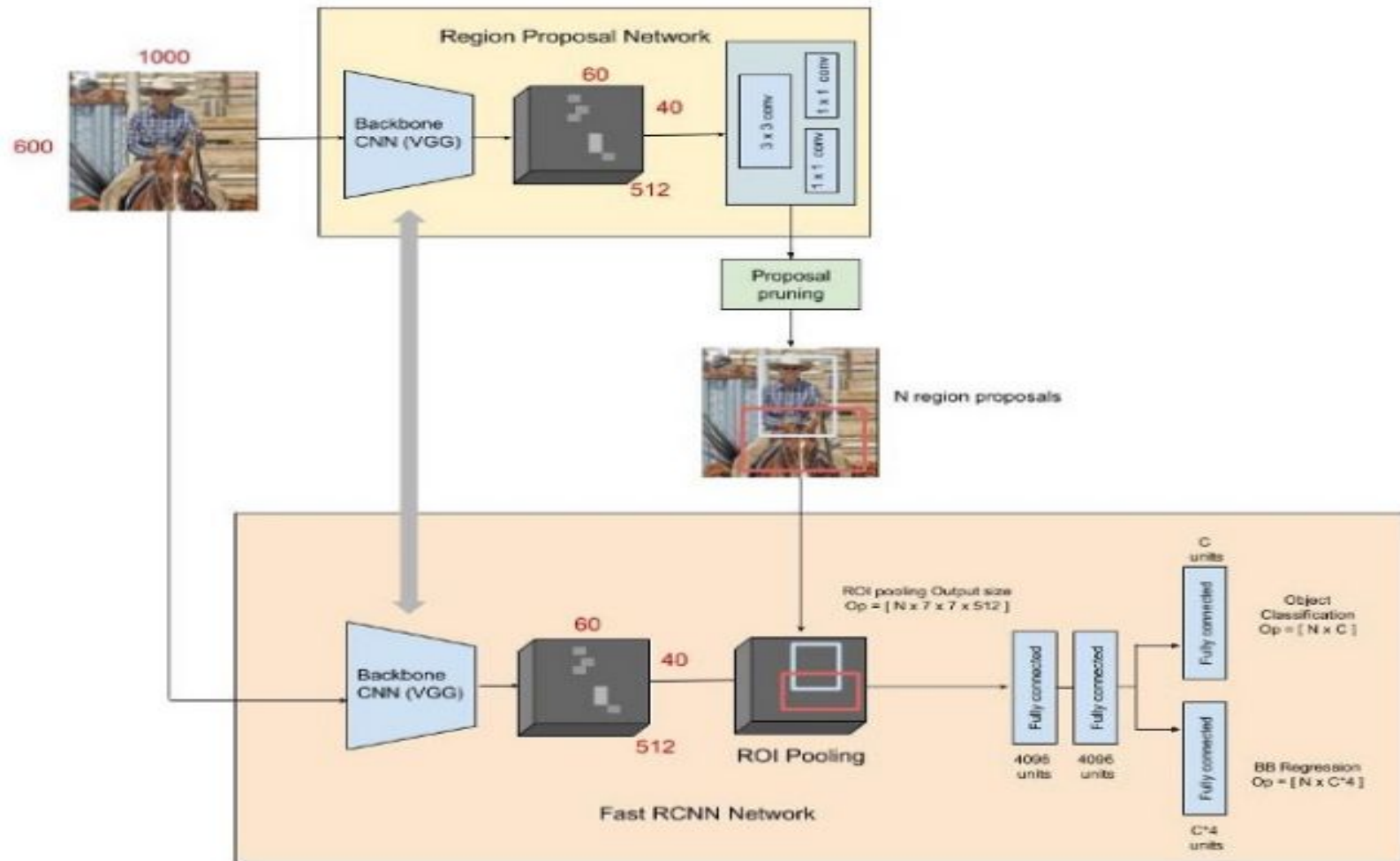




Вузьке місце – генерація регіонів кандидатів

Faster R-CNN







Виберіть твердження яке відповідає дійсності щодо розглянутих варіацій R CNN

Мережа сегментує зображення
класифікуючи кожен піксель

Мережа будує обмежувальні рамки, в
яких з високою ймовірністю є об'єкт

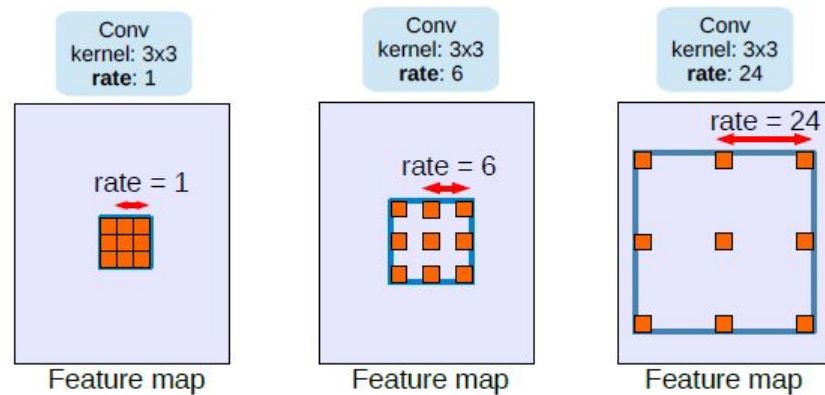
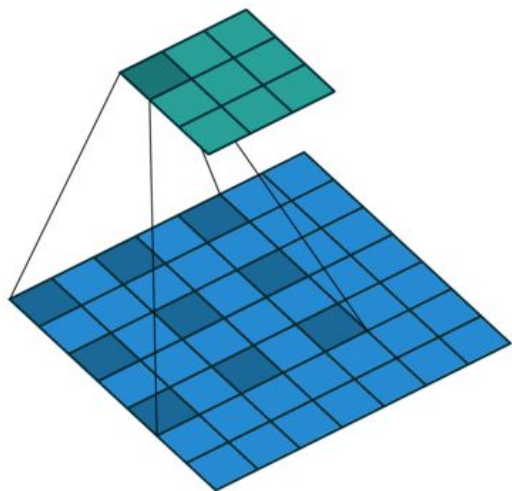
Мережа відносить зображення до
того чи іншого класу

В якості базової мережі в ній може
бути використана лише AlexNet

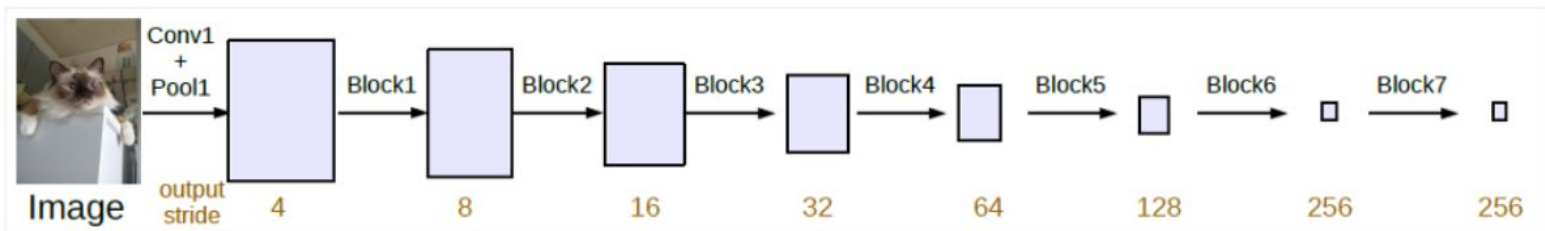


DeepLab

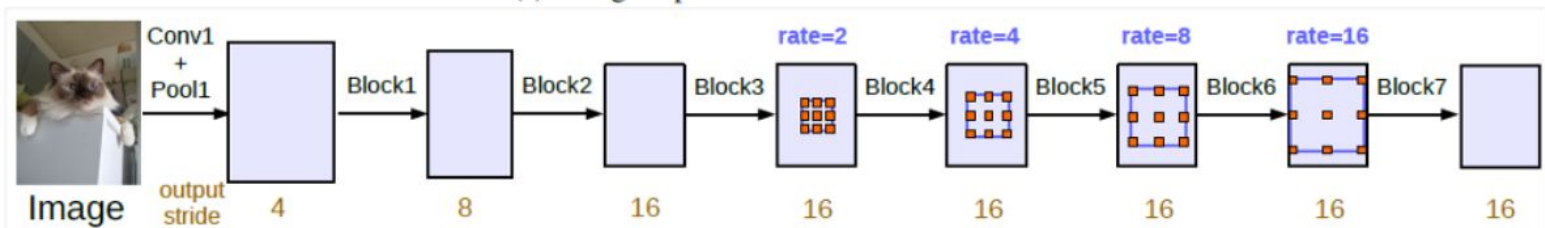
- 1. atrous convolutions



$$y[i] = \sum_k x[i + r \cdot k] w[k]$$

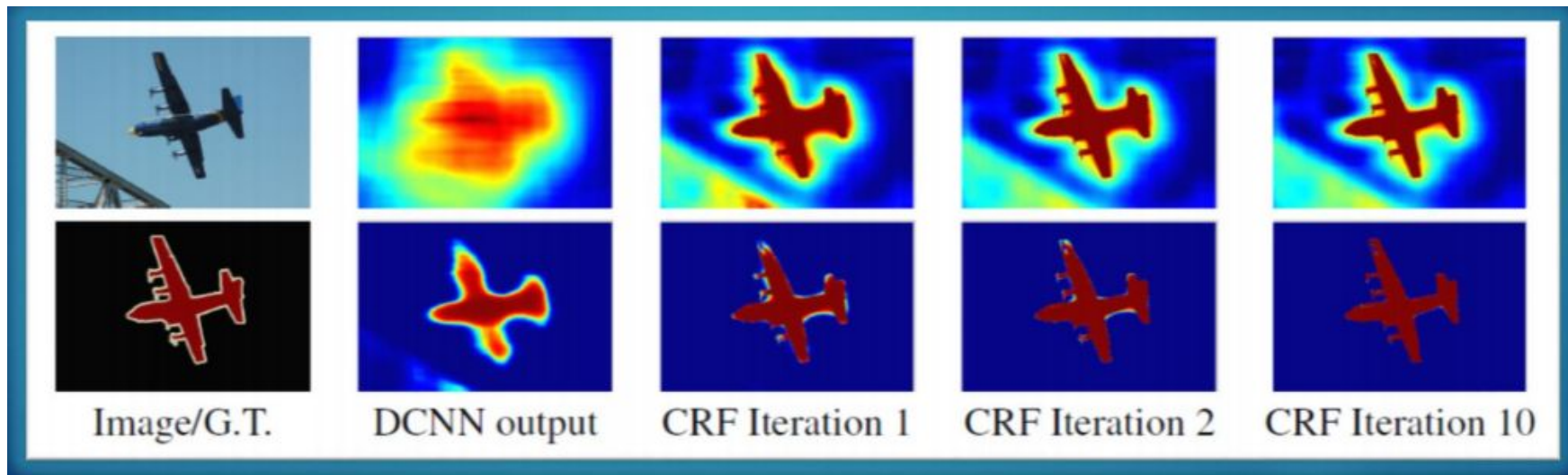


(a) Going deeper without atrous convolution.



DeepLab

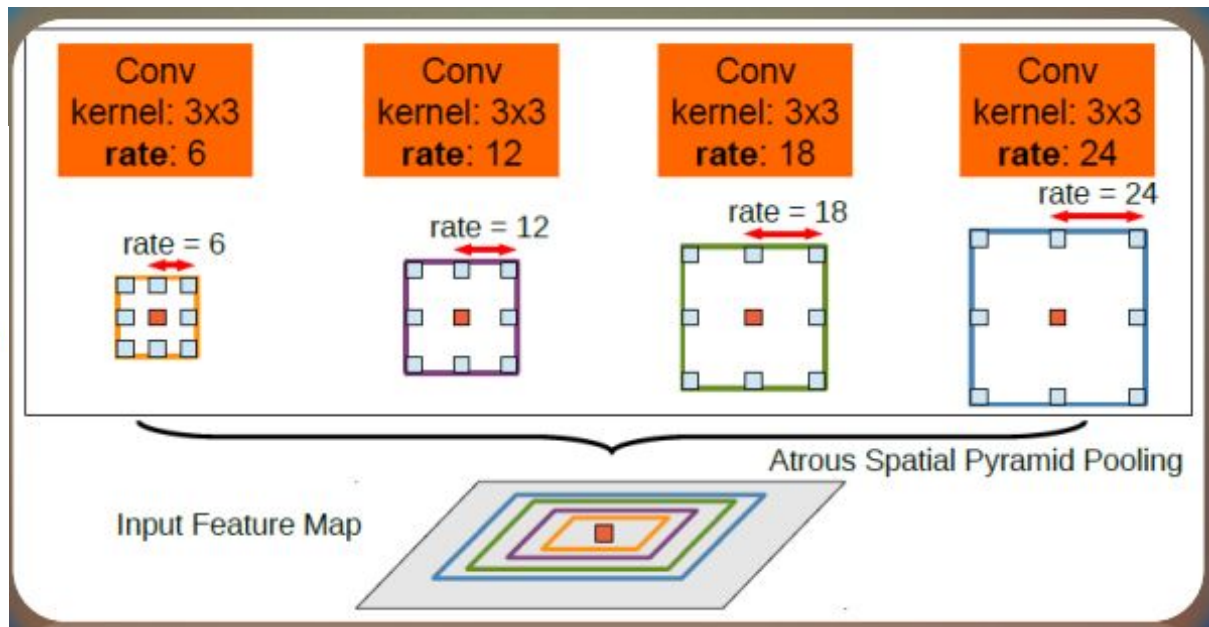
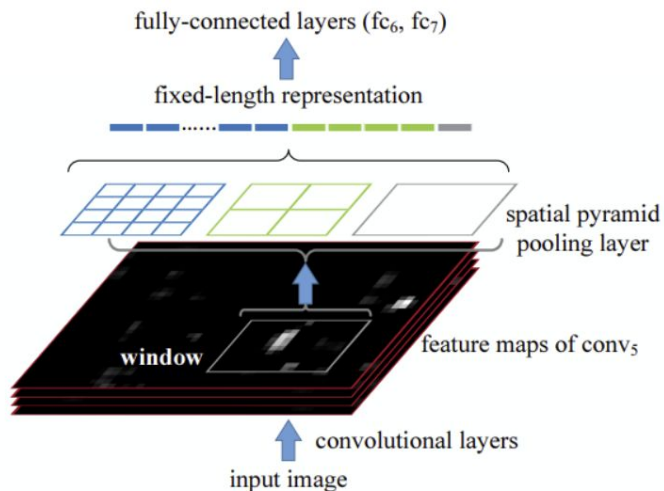
- 2. Conditional Random Field (CRF)



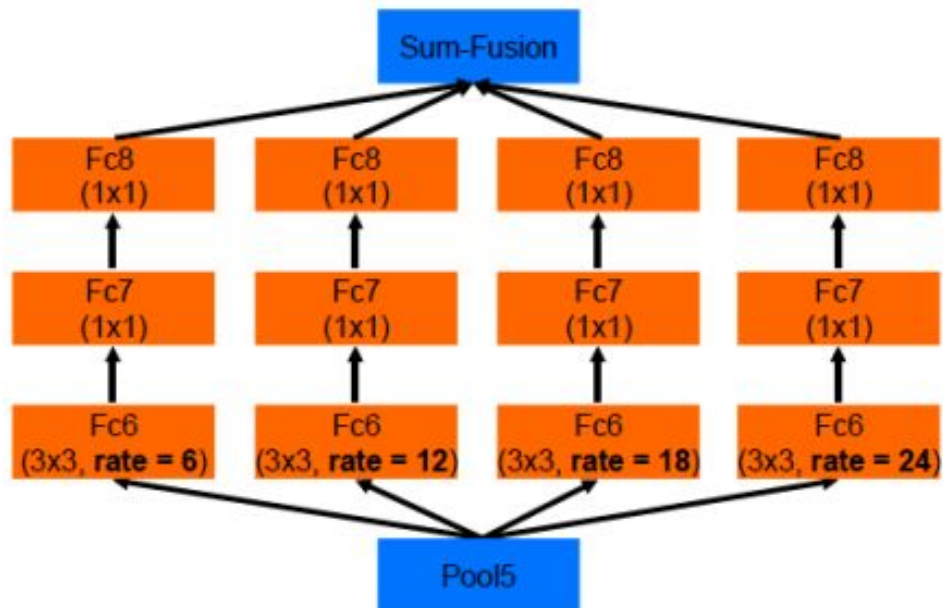
DeepLab

- 3. Atrous Spatial Pyramid Pooling (ASPP)

SPP



DeepLab

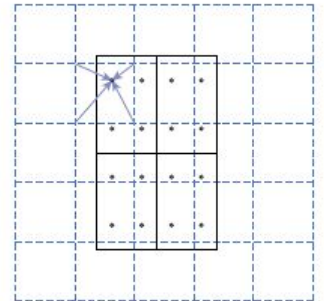
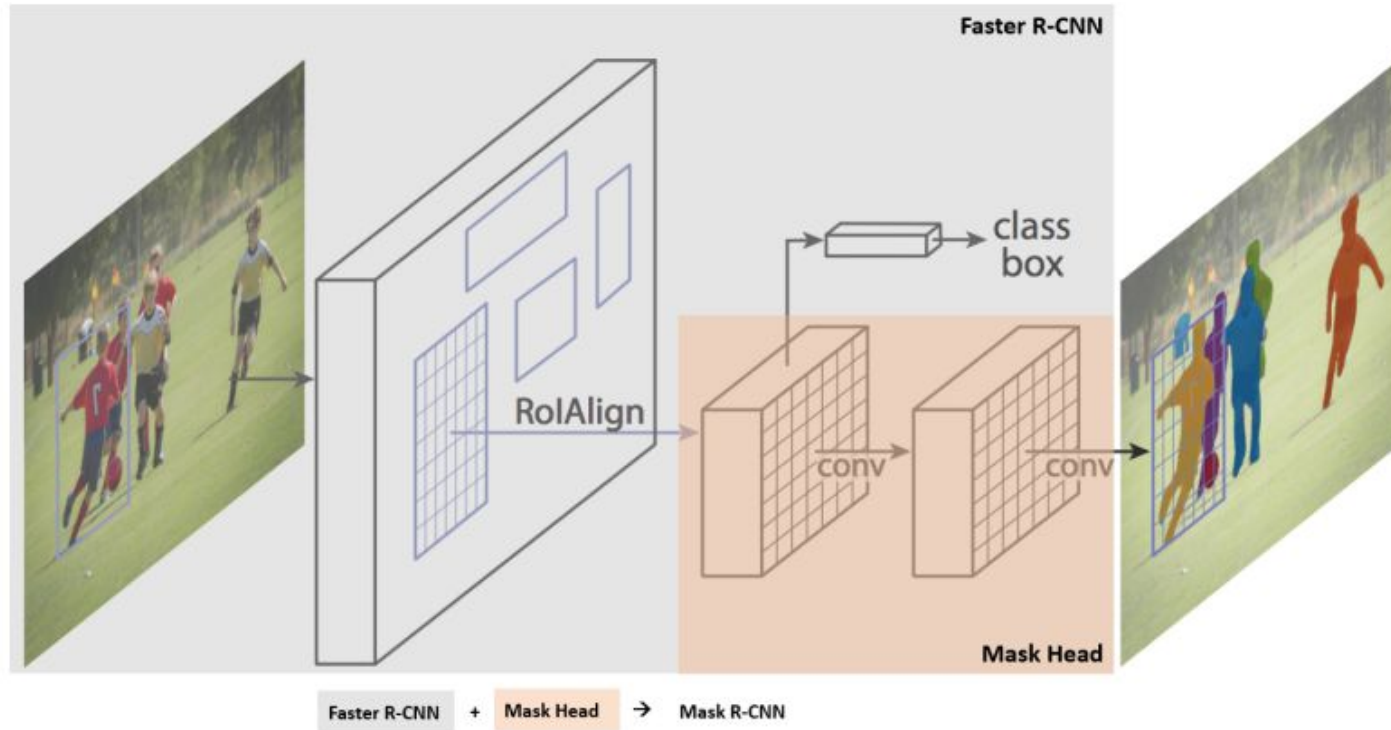


- DeepLabV1: використовує розширювальну згортку та Fully Connected Conditional Random Field (CRF) для управління роздільною здатністю, при якій обчислюються характеристики зображення.
- DeepLabV2: використовує Atrous Spatial Pyramid Pooling (ASPP) щоб розглянути об'єктів у різних масштабах та сегментації з набагато кращеною точністю.
- DeepLabV3: Окрім використання Atrous Convolution, DeepLabV3 використовує вдосконалений модуль ASPP, включаючи пакетну нормалізацію та функції рівня зображення. Не використовується CRF, як використовується у V1 та V2.

В чому первага використання atrous (dilated) згортки?



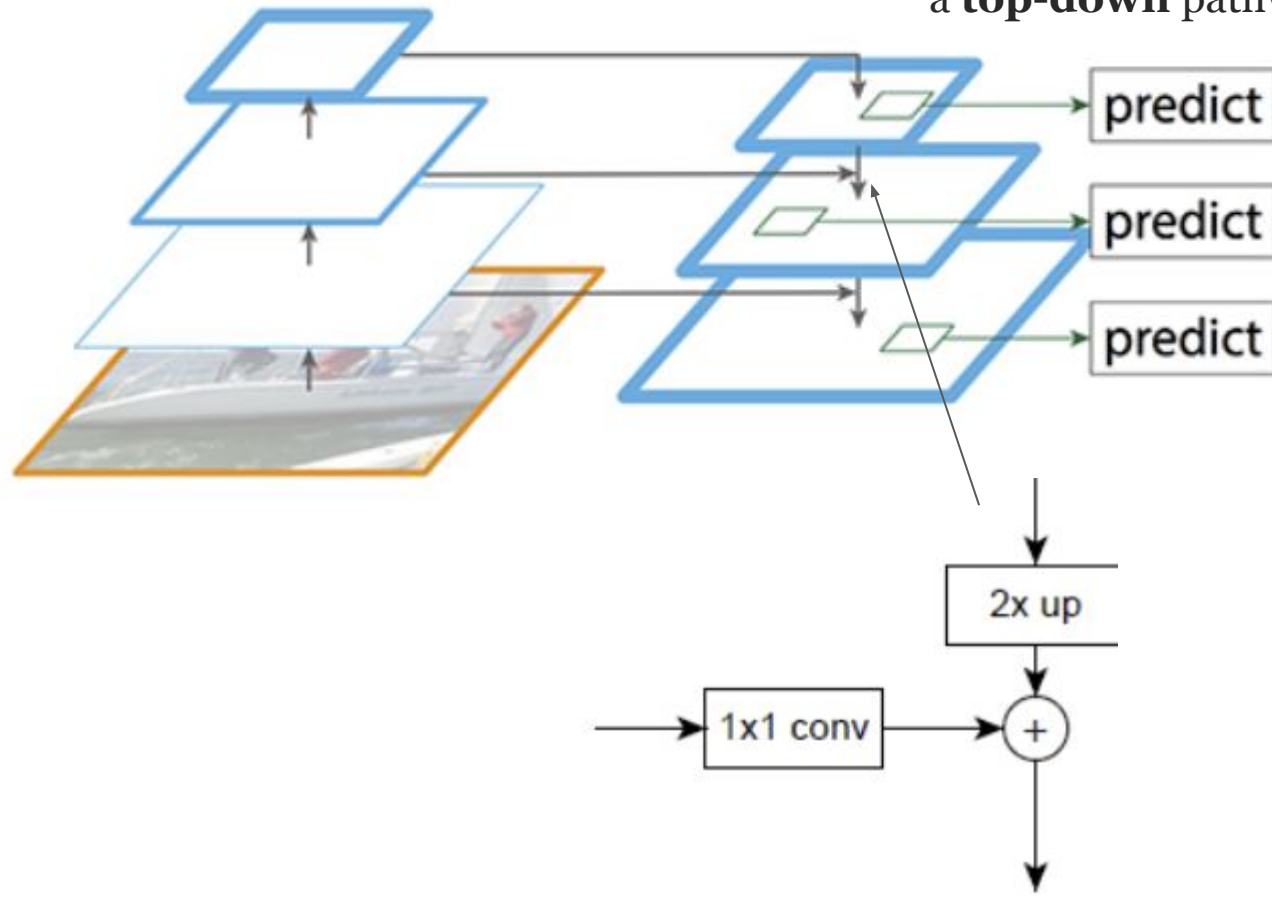
Mask R-CNN



Feature Pyramid Network (FPN)

a **bottom-up** pathway

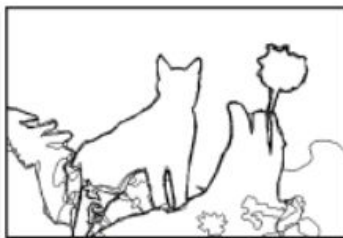
a **top-down** pathway



Holistically-Nested Edge Detection



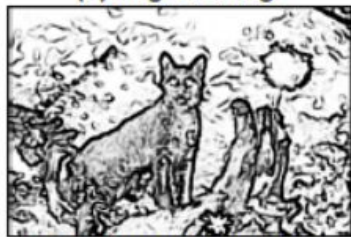
(a) original image



(b) ground truth



(c) HED: output



(d) HED: side output 2



(e) HED: side output 3



(f) HED: side output 4



(g) Canny: $\sigma = 2$



(h) Canny: $\sigma = 4$



(i) Canny: $\sigma = 8$