

Bài 8:

Một số ứng dụng học sâu trong thị giác máy (Phần 2)

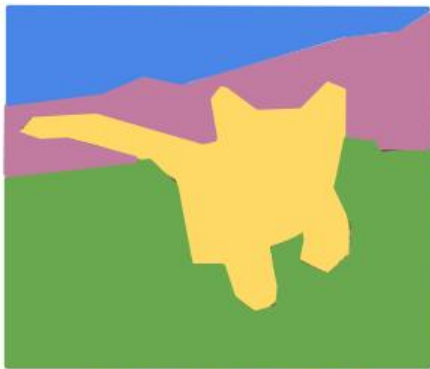
Nội dung

1. Giới thiệu bài toán phân đoạn ảnh
2. Lớp tăng độ phân giải upsampling
3. Hàm mục tiêu
4. Một số mạng phân đoạn ảnh tiêu biểu

Giới thiệu bài toán phân đoạn ảnh

Các bài toán thị giác máy

Semantic Segmentation



GRASS, CAT,
TREE, SKY

No objects, just pixels

Classification + Localization



CAT

Single Object

Object Detection



DOG, DOG, CAT

Multiple Object

Instance Segmentation

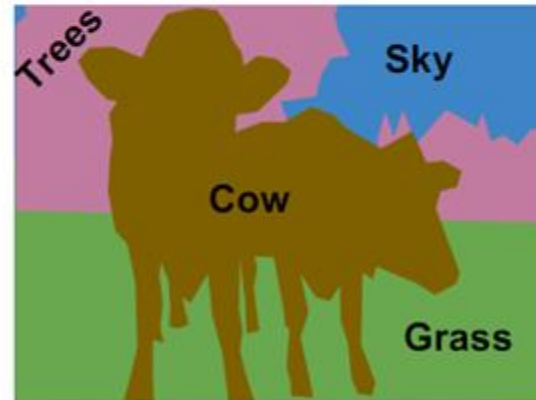
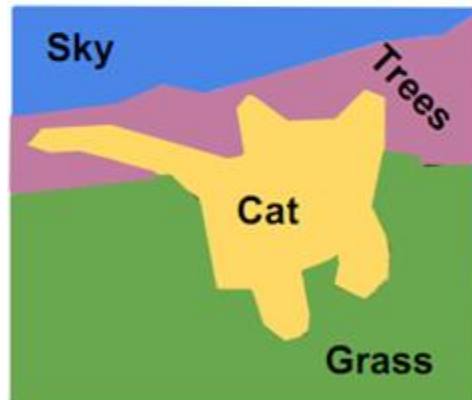


DOG, DOG, CAT

This image is CC0 public domain

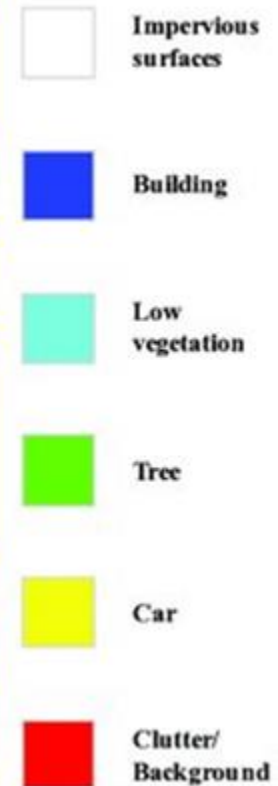
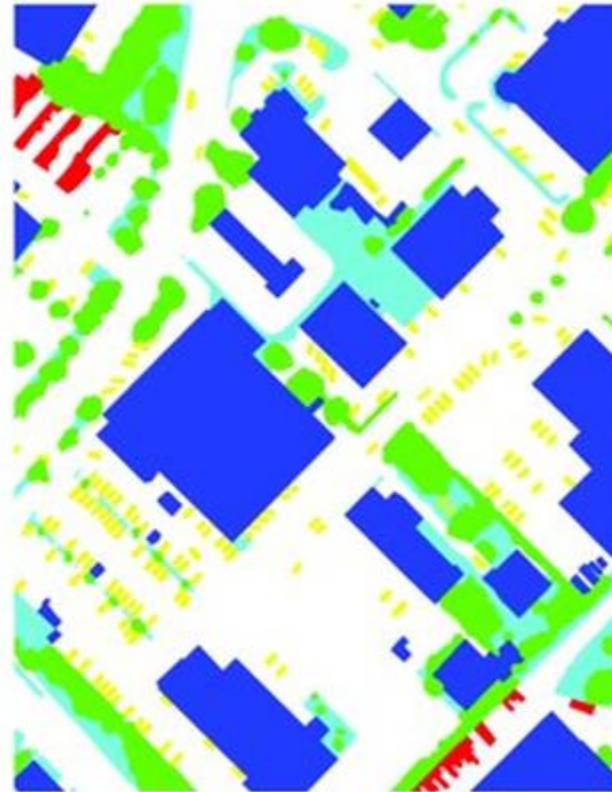
Phân vùng

- Phân lớp từng điểm ảnh trong ảnh
- Không phân biệt các đối tượng cùng lớp trong ảnh



Một số ứng dụng phân đoạn ảnh

- Phân đoạn ảnh vệ tinh và hàng không



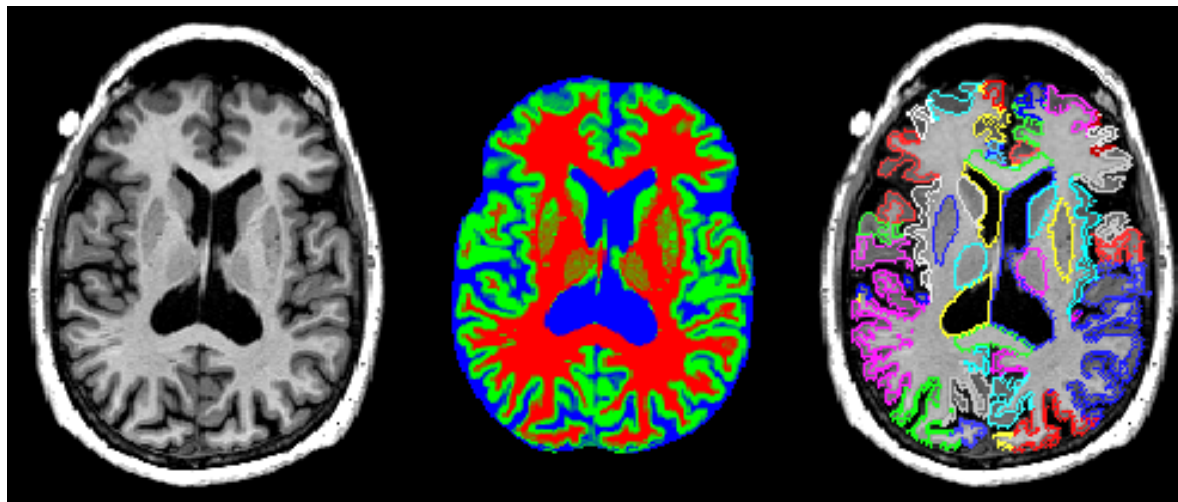
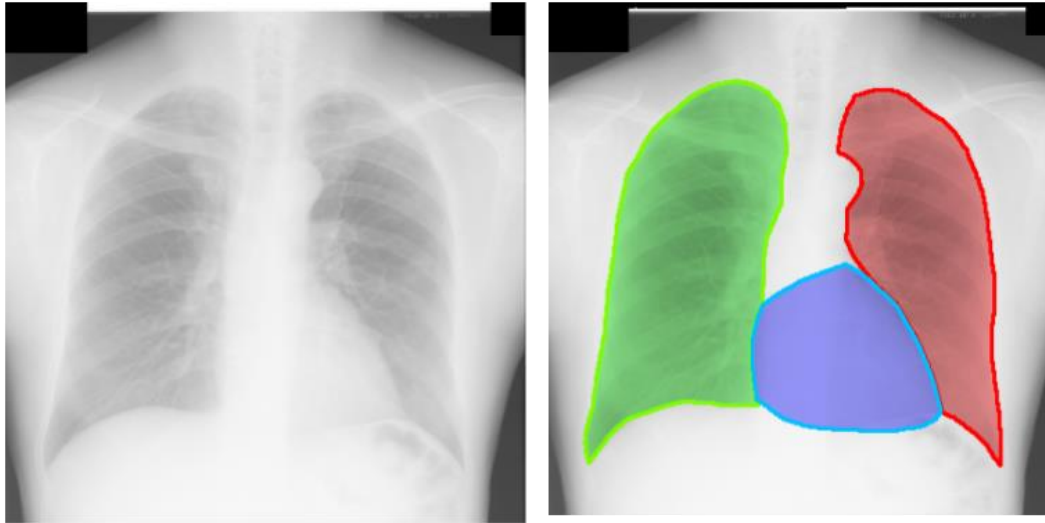
Một số ứng dụng phân đoạn ảnh

- Xe tự hành



Một số ứng dụng phân đoạn ảnh

- Y tế



- 7.2 David Goldblatt wrote in his world**
- British filmmaker and Canadian Foreign Minister Michael Ondaatje, former member of Parliament, Michael Ondaatje, and other Canadian officials, the highest Foreign Minister and 2000 to 2001 the ministerial and the minister. But I'm delighted. It's going to be rough, because they are very hard. And I can make some sense out of it.
- | Year | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 | 2026 | 2027 | 2028 | 2029 | 2030 | 2031 | 2032 | 2033 | 2034 | 2035 | 2036 | 2037 | 2038 | 2039 | 2040 | 2041 | 2042 | 2043 | 2044 | 2045 | 2046 | 2047 | 2048 | 2049 | 2050 | 2051 | 2052 | 2053 | 2054 | 2055 | 2056 | 2057 | 2058 | 2059 | 2060 | 2061 | 2062 | 2063 | 2064 | 2065 | 2066 | 2067 | 2068 | 2069 | 2070 | 2071 | 2072 | 2073 | 2074 | 2075 | 2076 | 2077 | 2078 | 2079 | 2080 | 2081 | 2082 | 2083 | 2084 | 2085 | 2086 | 2087 | 2088 | 2089 | 2090 | 2091 | 2092 | 2093 | 2094 | 2095 | 2096 | 2097 | 2098 | 2099 | 2100 | 2101 | 2102 | 2103 | 2104 | 2105 | 2106 | 2107 | 2108 | 2109 | 2110 | 2111 | 2112 | 2113 | 2114 | 2115 | 2116 | 2117 | 2118 | 2119 | 2120 | 2121 | 2122 | 2123 | 2124 | 2125 | 2126 | 2127 | 2128 | 2129 | 2130 | 2131 | 2132 | 2133 | 2134 | 2135 | 2136 | 2137 | 2138 | 2139 | 2140 | 2141 | 2142 | 2143 | 2144 | 2145 | 2146 | 2147 | 2148 | 2149 | 2150 | 2151 | 2152 | 2153 | 2154 | 2155 | 2156 | 2157 | 2158 | 2159 | 2160 | 2161 | 2162 | 2163 | 2164 | 2165 | 2166 | 2167 | 2168 | 2169 | 2170 | 2171 | 2172 | 2173 | 2174 | 2175 | 2176 | 2177 | 2178 | 2179 | 2180 | 2181 | 2182 | 2183 | 2184 | 2185 | 2186 | 2187 | 2188 | 2189 | 2190 | 2191 | 2192 | 2193 | 2194 | 2195 | 2196 | 2197 | 2198 | 2199 | 2200 | 2201 | 2202 | 2203 | 2204 | 2205 | 2206 | 2207 | 2208 | 2209 | 2210 | 2211 | 2212 | 2213 | 2214 | 2215 | 2216 | 2217 | 2218 | 2219 | 2220 | 2221 | 2222 | 2223 | 2224 | 2225 | 2226 | 2227 | 2228 | 2229 | 2230 | 2231 | 2232 | 2233 | 2234 | 2235 | 2236 | 2237 | 2238 | 2239 | 2240 | 2241 | 2242 | 2243 | 2244 | 2245 | 2246 | 2247 | 2248 | 2249 | 2250 | 2251 | 2252 | 2253 | 2254 | 2255 | 2256 | 2257 | 2258 | 2259 | 2260 | 2261 | 2262 | 2263 | 2264 | 2265 | 2266 | 2267 | 2268 | 2269 | 2270 | 2271 | 2272 | 2273 | 2274 | 2275 | 2276 | 2277 | 2278 | 2279 | 2280 | 2281 | 2282 | 2283 | 2284 | 2285 | 2286 | 2287 | 2288 | 2289 | 2290 | 2291 | 2292 | 2293 | 2294 | 2295 | 2296 | 2297 | 2298 | 2299 | 2300 | 2301 | 2302 | 2303 | 2304 | 2305 | 2306 | 2307 | 2308 | 2309 | 2310 | 2311 | 2312 | 2313 | 2314 | 2315 | 2316 | 2317 | 2318 | 2319 | 2320 | 2321 | 2322 | 2323 | 2324 | 2325 | 2326 | 2327 | 2328 | 2329 | 2330 | 2331 | 2332 | 2333 | 2334 | 2335 | 2336 | 2337 | 2338 | 2339 | 2340 | 2341 | 2342 | 2343 | 2344 | 2345 | 2346 | 2347 | 2348 | 2349 | 2350 | 2351 | 2352 | 2353 | 2354 | 2355 | 2356 | 2357 | 2358 | 2359 | 2360 | 2361 | 2362 | 2363 | 2364 | 2365 | 2366 | 2367 | 2368 | 2369 | 2370 | 2371 | 2372 | 2373 | 2374 | 2375 | 2376 | 2377 | 2378 | 2379 | 2380 | 2381 | 2382 | 2383 | 2384 | 2385 | 2386 | 2387 | 2388 | 2389 | 2390 | 2391 | 2392 | 2393 | 2394 | 2395 | 2396 | 2397 | 2398 | 2399 | 2400 | 2401 | 2402 | 2403 | 2404 | 2405 | 2406 | 2407 | 2408 | 2409 | 2410 | 2411 | 2412 | 2413 | 2414 | 2415 | 2416 | 2417 | 2418 | 2419 | 2420 | 2421 | 2422 | 2423 | 2424 | 2425 | 2426 | 2427 | 2428 | 2429 | 2430 | 2431 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|

9

Trượt cửa sổ

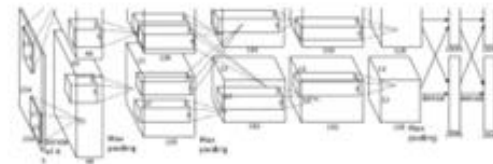
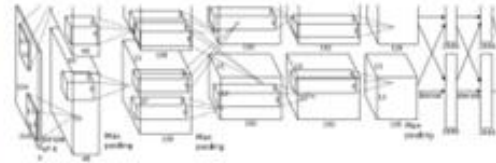
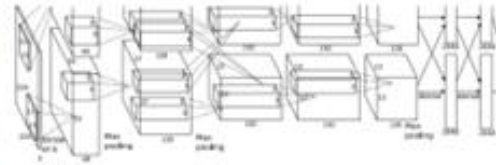
Ảnh đầu vào



Trích từng vùng ảnh



Phân lớp điểm ảnh trung tâm bằng mạng CNN



Cow

Cow

Grass

Trượt cửa sổ

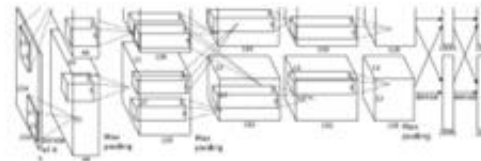
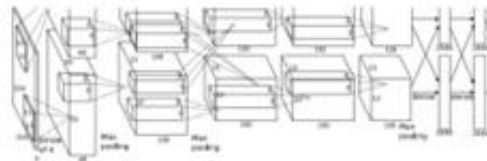
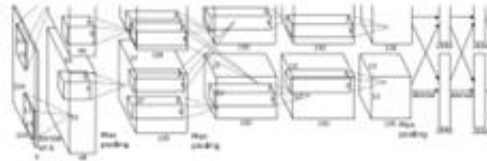
Ảnh đầu vào



Trích từ vùng ảnh



Phân lớp điểm ảnh trung tâm bằng mạng CNN



Cow

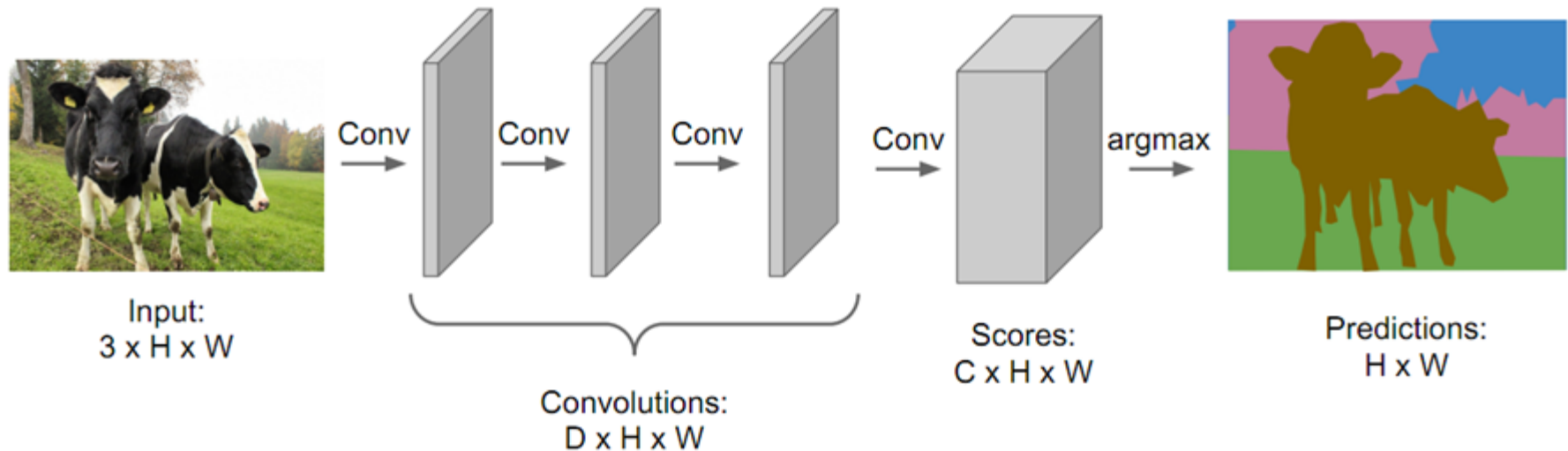
Cow

Grass

Không hiệu quả! Không tái sử dụng được đặc trưng các vùng trùng nhau

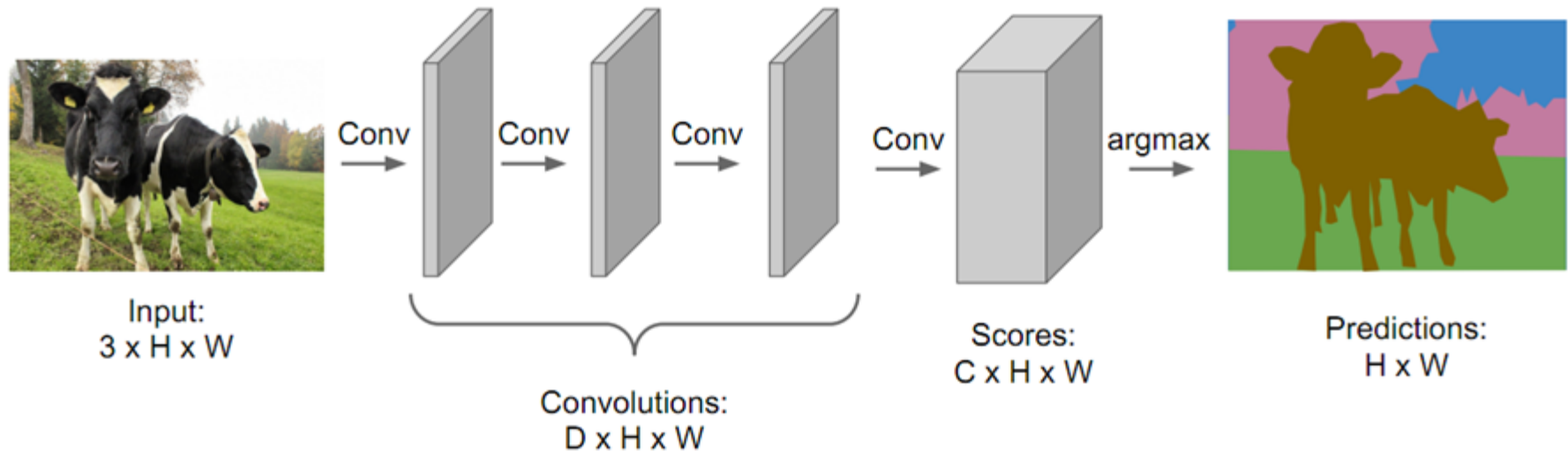
Tích chập hoàn toàn (Fully Convolutional)

- Thiết kế mạng CNN gồm nhiều lớp tích chập để phân lớp đồng thời tất cả các điểm ảnh.



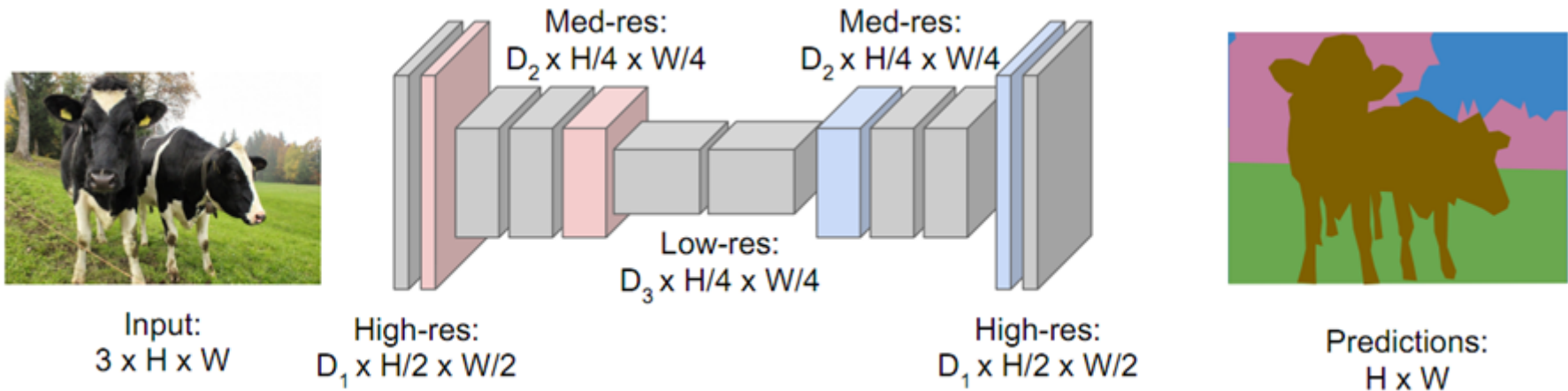
Tích chập hoàn toàn (Fully Convolutional)

- Thiết kế mạng CNN gồm nhiều lớp tích chập để phân lớp đồng thời tất cả các điểm ảnh.
- Vấn đề: Tích chập với các lớp đầu vào có độ phân giải cao đòi hỏi nhiều chi phí tính toán



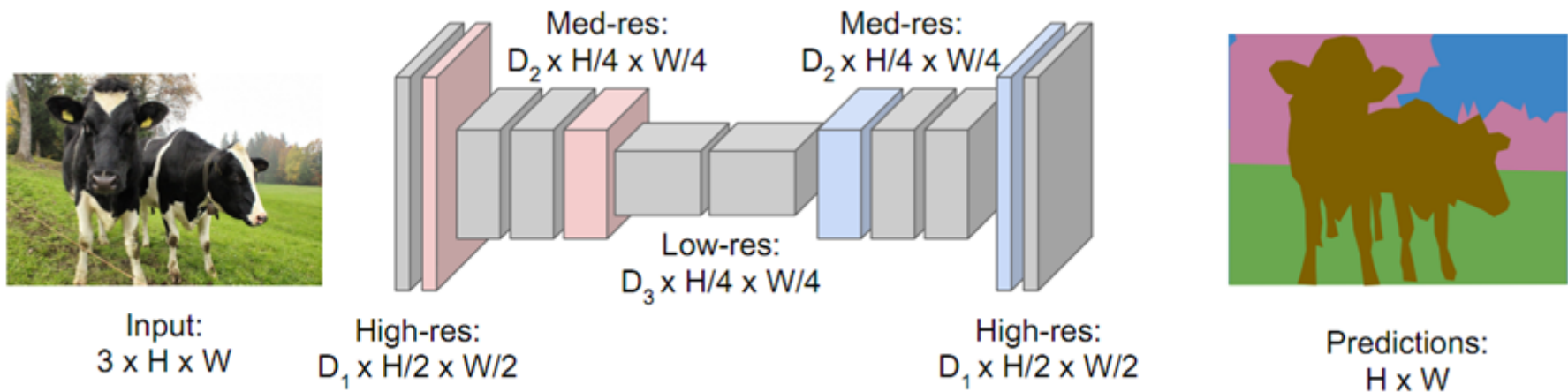
Tích chập hoàn toàn (Fully Convolutional)

- Thiết kế mạng CNN với các lớp giảm độ phân giải (downsampling) và tăng độ phân giải (upsampling)



Tích chập hoàn toàn (Fully Convolutional)

- Thiết kế mạng CNN với các lớp giảm độ phân giải (downsampling) và tăng độ phân giải (upsampling)
- Giảm độ phân giải: max pooling hay strided conv
- Tăng độ phân giải?



Lớp tăng độ phân giải upsampling

Lớp Unpooling

- Các lớp này không có tham số

Nearest Neighbor

| | |
|---|---|
| 1 | 2 |
| 3 | 4 |



| | | | |
|---|---|---|---|
| 1 | 1 | 2 | 2 |
| 1 | 1 | 2 | 2 |
| 3 | 3 | 4 | 4 |
| 3 | 3 | 4 | 4 |

Input: 2 x 2

Output: 4 x 4

“Bed of Nails”

| | |
|---|---|
| 1 | 2 |
| 3 | 4 |



| | | | |
|---|---|---|---|
| 1 | 0 | 2 | 0 |
| 0 | 0 | 0 | 0 |
| 3 | 0 | 4 | 0 |
| 0 | 0 | 0 | 0 |

Input: 2 x 2

Output: 4 x 4

Lớp Max Unpooling

Max Pooling

Ghi nhớ vị trí phần tử lớn nhất

| | | | |
|---|---|---|---|
| 1 | 2 | 6 | 3 |
| 3 | 5 | 2 | 1 |
| 1 | 2 | 2 | 1 |
| 7 | 3 | 4 | 8 |

Input: 4 x 4



| | |
|---|---|
| 5 | 6 |
| 7 | 8 |

Output: 2 x 2



Rest of the network

Max Unpooling

Sử dụng vị trí đã ghi nhớ khi pooling

| | |
|---|---|
| 1 | 2 |
| 3 | 4 |

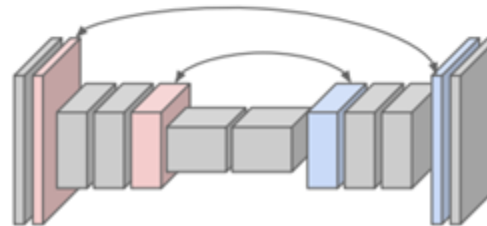
Input: 2 x 2



| | | | |
|---|---|---|---|
| 0 | 0 | 2 | 0 |
| 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 4 |

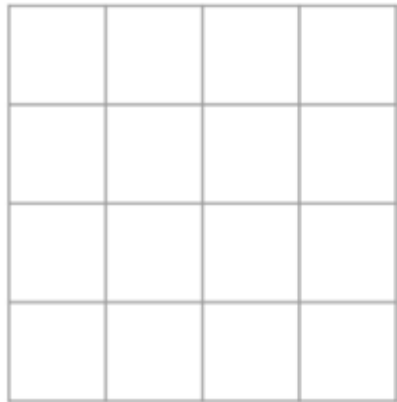
Output: 4 x 4

Các cặp max pooling và max unpooling được dùng đối xứng nhau trong mạng

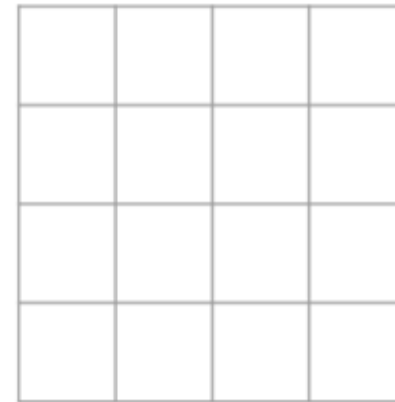


Tích chập chuyển vị

- Là phép tăng độ phân giải (upsampling) có chứa các tham số có thể huấn luyện được



Input: 4 x 4

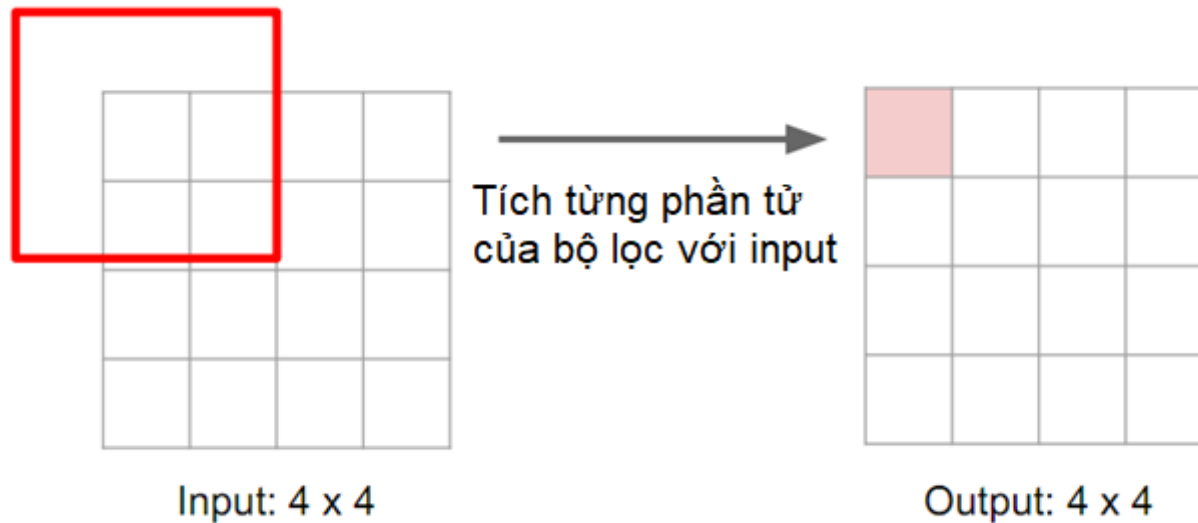


Output: 4 x 4

- Xem lại ví dụ tích chập conv 3x3, bước nhảy stride 1 và thêm viền padding 1

Tích chập chuyển vị

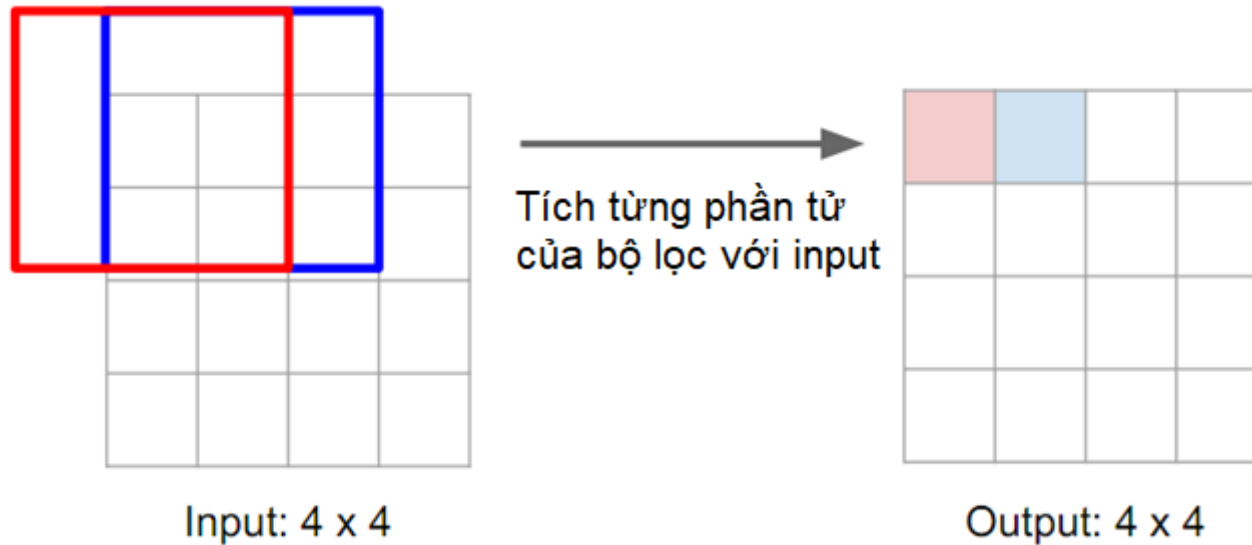
- Là phép tăng độ phân giải (upsampling) có chứa các tham số có thể huấn luyện được



- Xem lại ví dụ tích chập conv 3x3, bước nhảy stride 1 và thêm viền padding 1

Tích chập chuyển vị

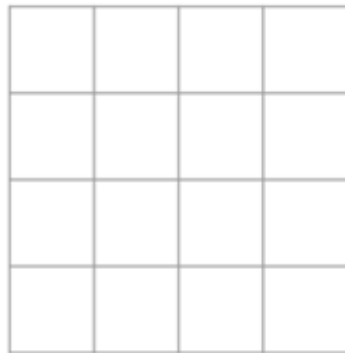
- Là phép tăng độ phân giải (upsampling) có chứa các tham số có thể huấn luyện được



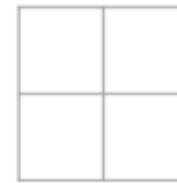
- Xem lại ví dụ tích chập conv 3x3, bước nhảy stride 1 và thêm viền padding 1

Tích chập chuyển vị

- Là phép tăng độ phân giải (upsampling) có chứa các tham số có thể huấn luyện được



Input: 4 x 4

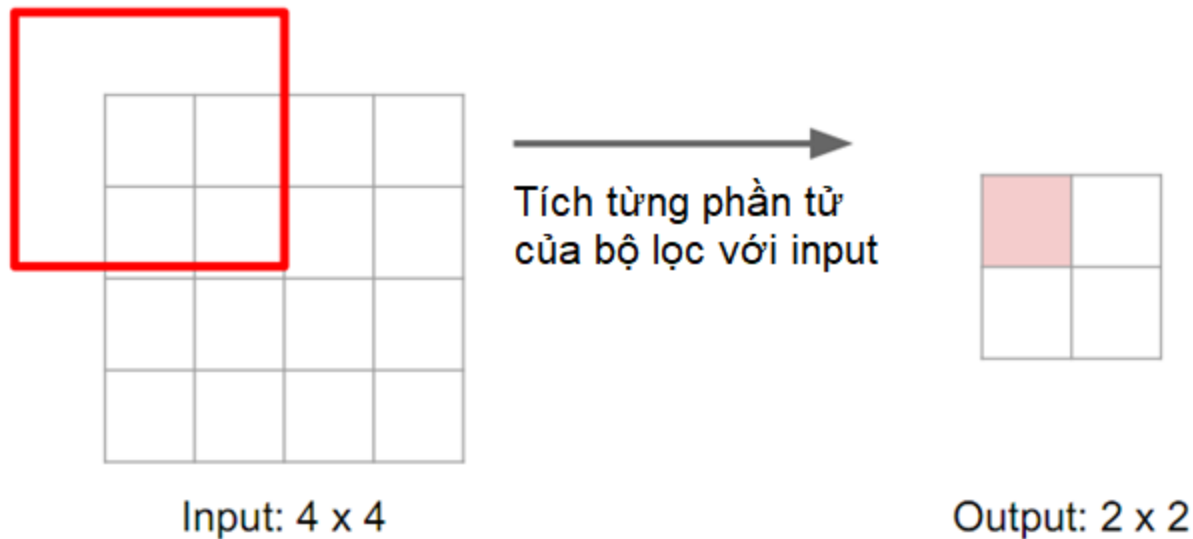


Output: 2 x 2

- Xem lại ví dụ tích chập conv 3x3, bước nhảy stride 2 và thêm viền padding 1

Tích chập chuyển vị

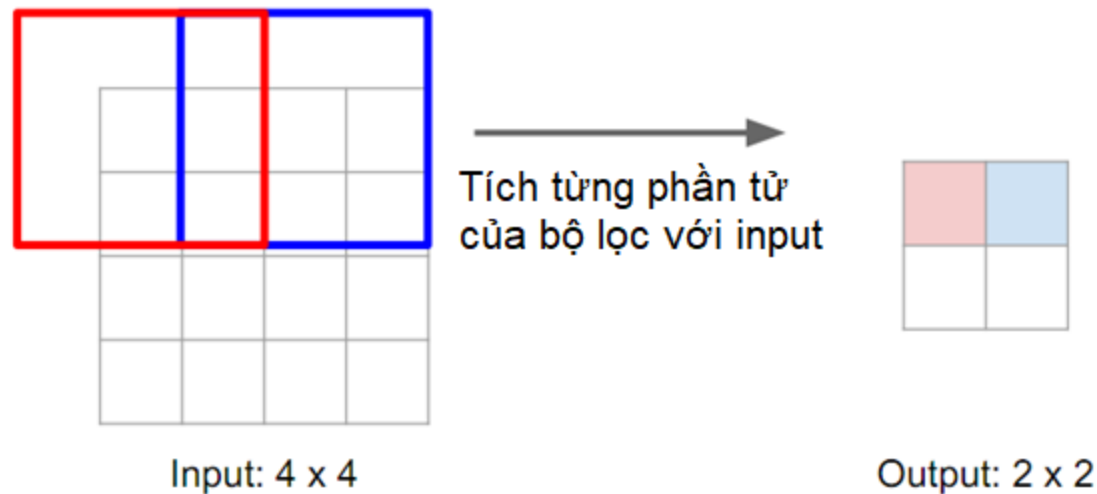
- Là phép tăng độ phân giải (upsampling) có chứa các tham số có thể huấn luyện được



- Xem lại ví dụ tích chập conv 3x3, bước nhảy stride 2 và thêm viền padding 1

Tích chập chuyển vị

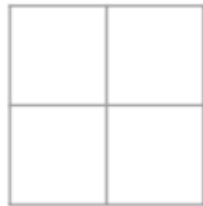
- Là phép tăng độ phân giải (upsampling) có chứa các tham số có thể huấn luyện được



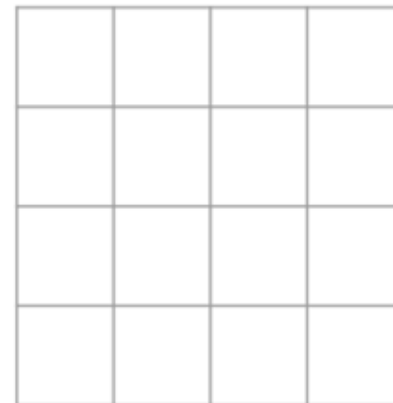
- Xem lại ví dụ tích chập conv 3x3, bước nhảy stride 2 và thêm viền padding 1

Tích chập chuyển vị

- Là phép tăng độ phân giải (upsampling) có chứa các tham số có thể huấn luyện được



Input: 2 x 2

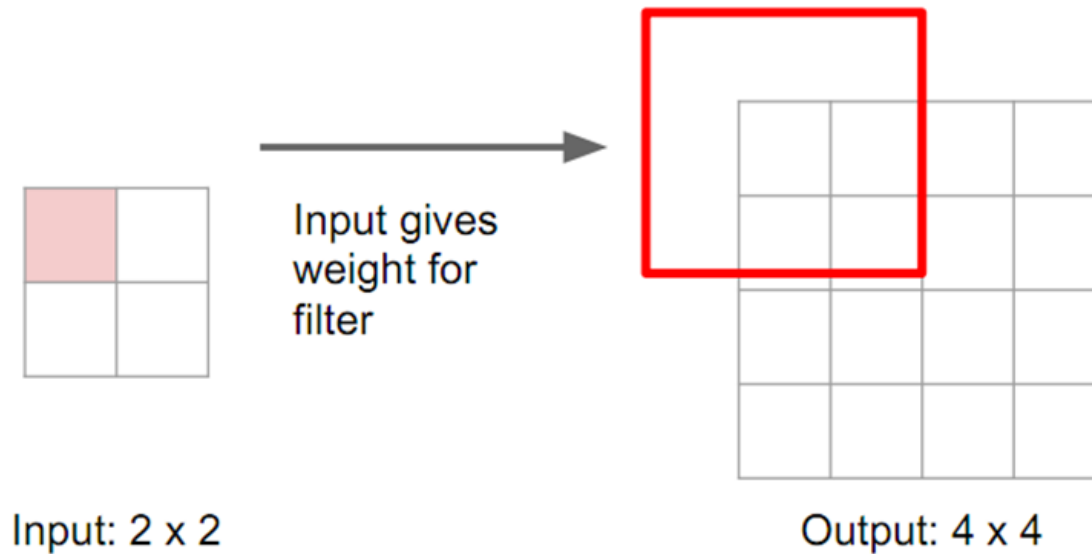


Output: 4 x 4

- Tích chập chuyển vị conv 3x3, bước nhảy stride 2 và thêm viền padding 1

Tích chập chuyển vị

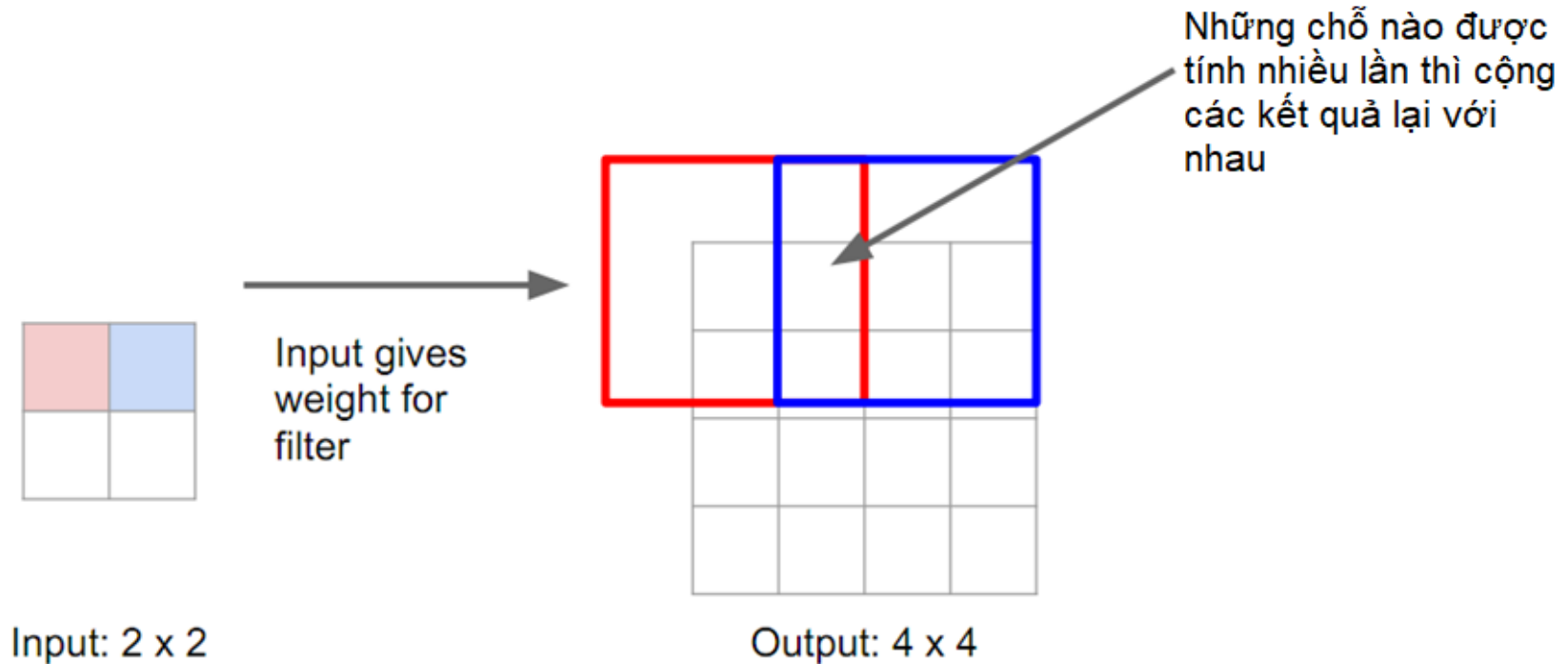
- Là phép tăng độ phân giải (upsampling) có chứa các tham số có thể huấn luyện được



- Tích chập chuyển vị conv 3x3, bước nhảy stride 2 và thêm viền padding 1

Tích chập chuyển vị

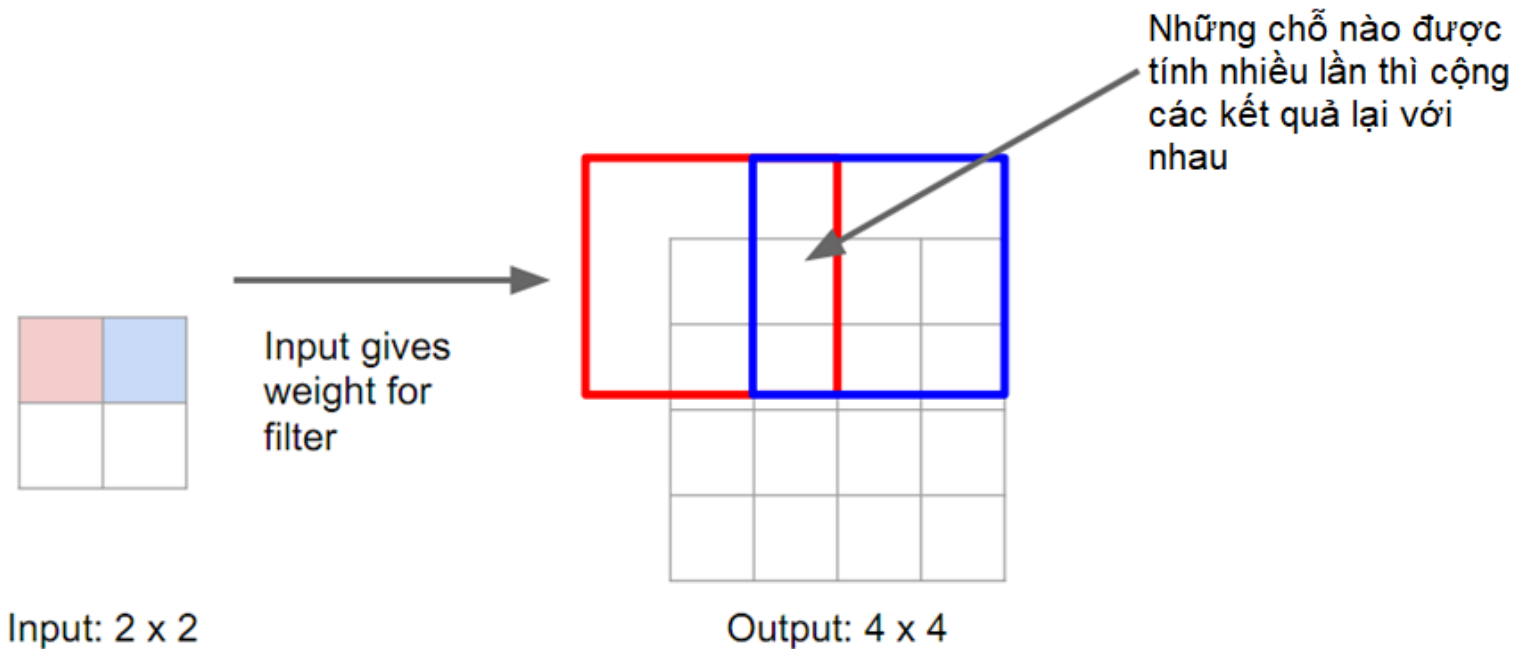
- Là phép tăng độ phân giải (upsampling) có chứa các tham số có thể huấn luyện được



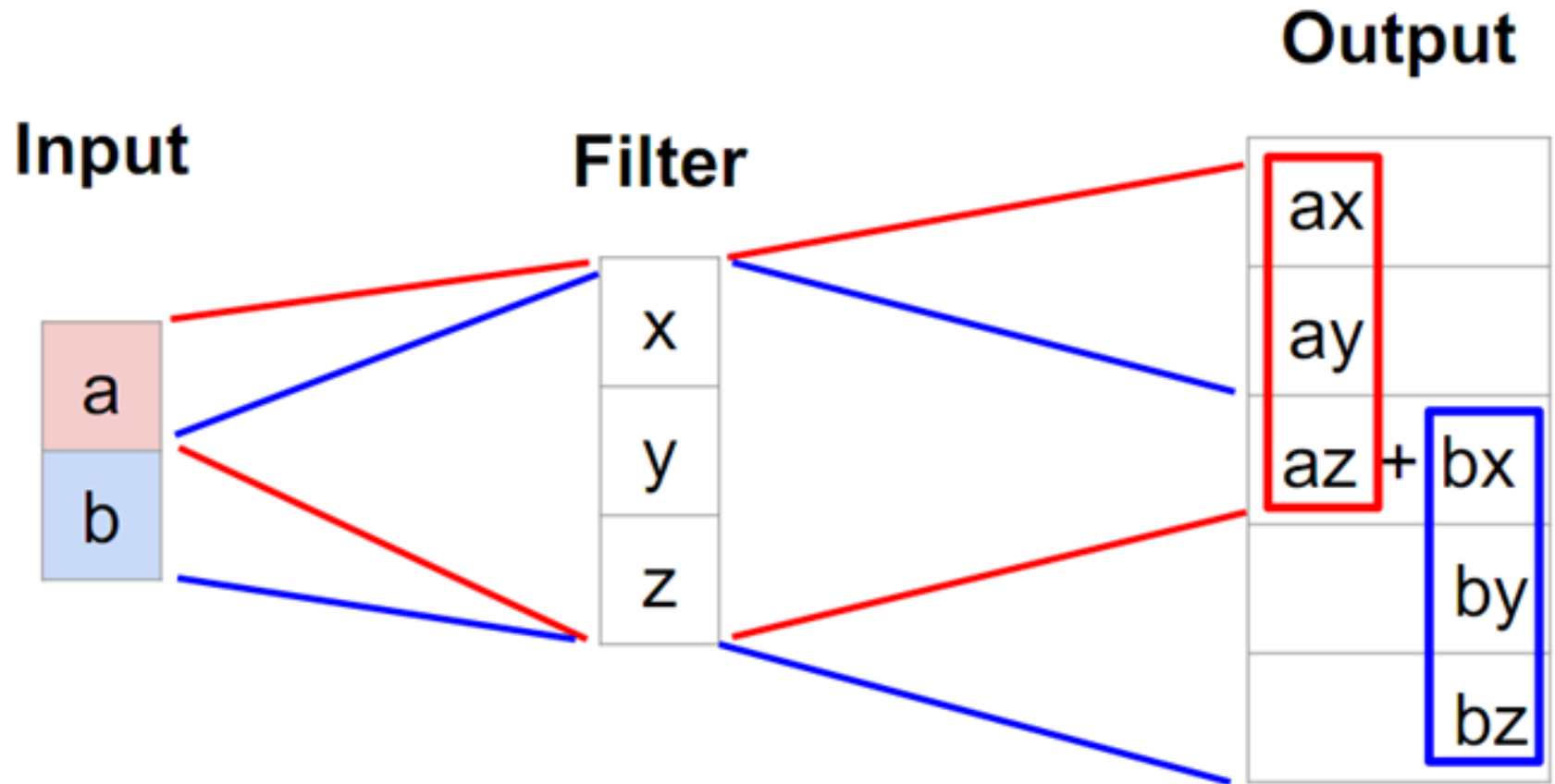
- Tích chập chuyển vị conv 3x3, bước nhảy stride 2 và thêm viền padding 1

Tích chập chuyển vị

- Tên gọi khác:
 - Deconvolution (không nên, dễ gây hiểu nhầm)
 - Upconvolution
 - Fractionally strided convolution
 - Backward strided convolution

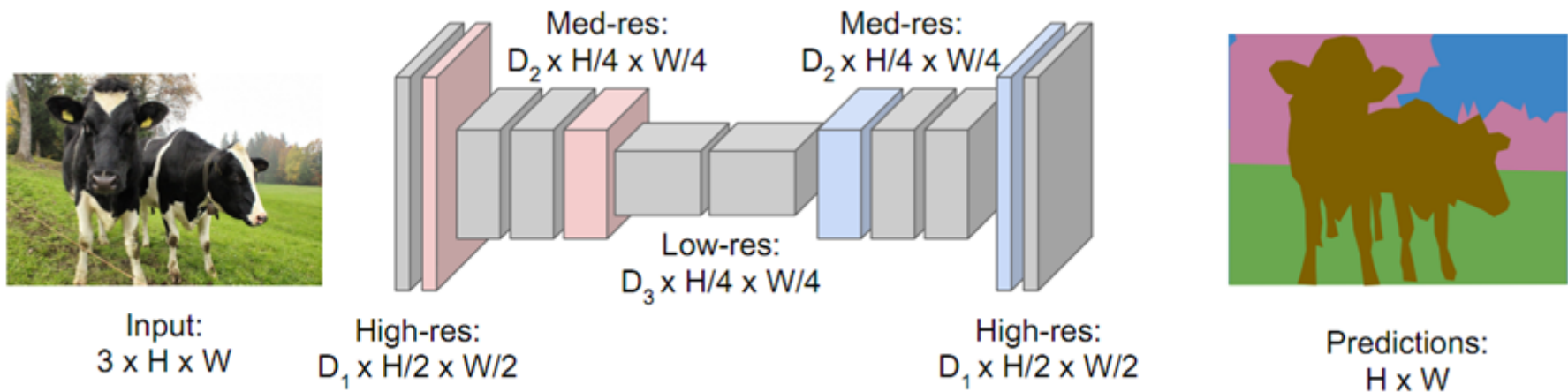


Ví dụ tích chập chuyển vị trong 1D



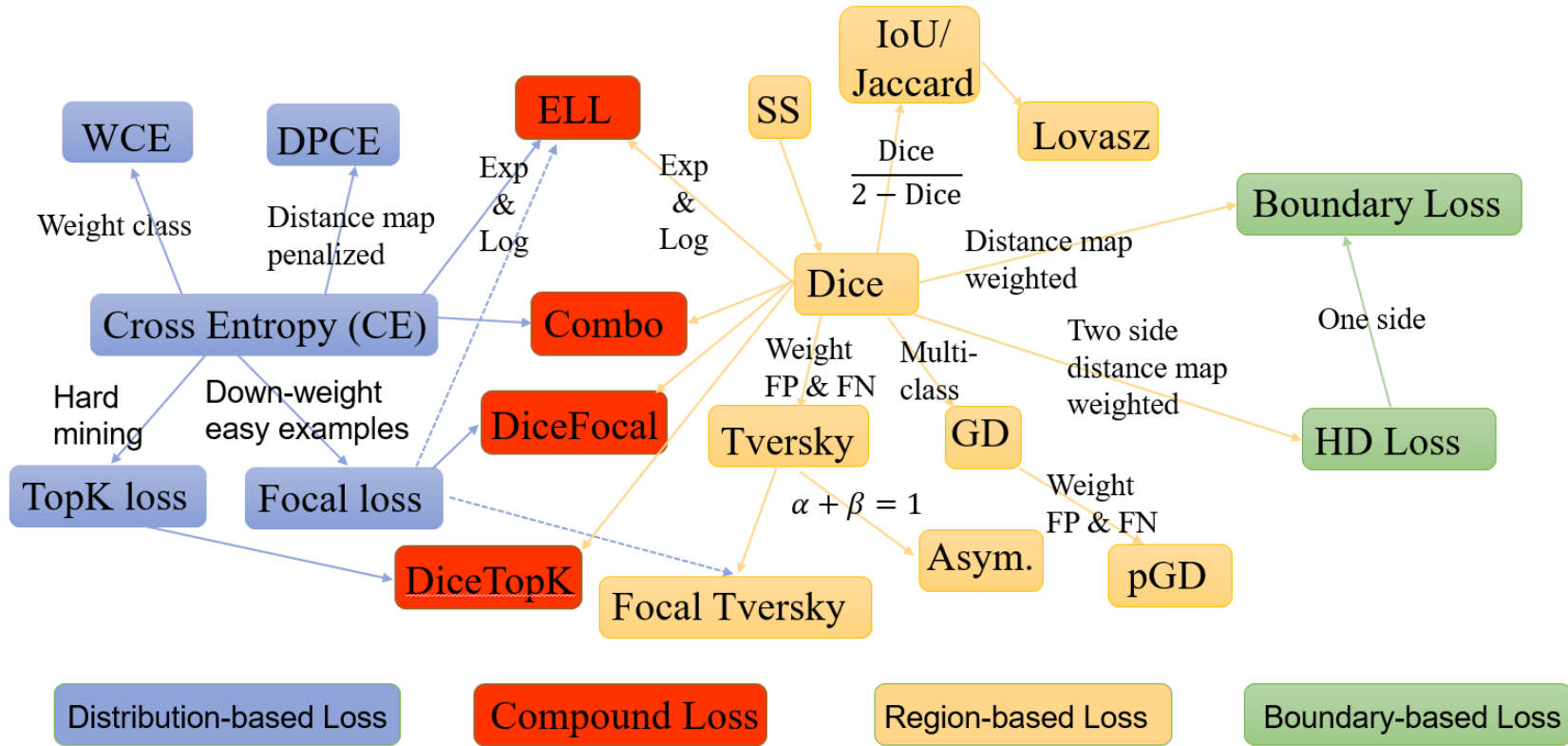
Tích chập hoàn toàn (Fully Convolutional)

- Thiết kế mạng CNN với các lớp giảm độ phân giải (downsampling) và tăng độ phân giải (upsampling)
- Giảm độ phân giải: max pooling hay strided conv
- Tăng độ phân giải: unpooling hoặc transpose conv



Hàm mục tiêu cho bài toán phân đoạn ảnh

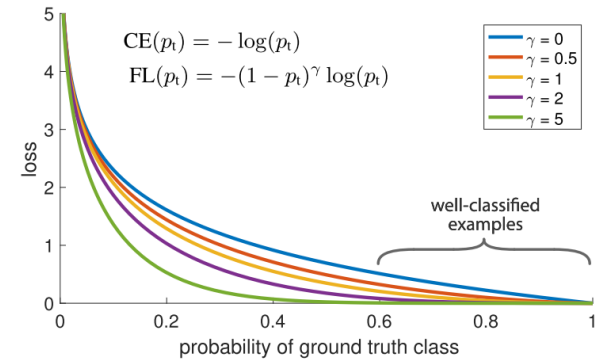
Hàm mục tiêu



Hàm mục tiêu dựa trên phân phối

- Cross Entropy (CE):

$$CE(p, \hat{p}) = -(p \log(\hat{p}) + (1 - p) \log(1 - \hat{p}))$$



- Weighted CE: mỗi lớp có trọng số khác nhau

$$WCE(p, \hat{p}) = -(\beta p \log(\hat{p}) + (1 - p) \log(1 - \hat{p}))$$

- Focal loss: giải quyết vấn đề mất cân bằng lớn giữa lớp nền và lớp đối tượng quan tâm. Giá trị hàm mục tiêu đối với những mẫu dễ phân loại được giảm xuống thấp để mạng tập trung hơn vào mẫu khó.

$$FL(p, \hat{p}) = -(\alpha(1 - \hat{p})^\gamma p \log(\hat{p}) + (1 - \alpha)\hat{p}^\gamma (1 - p) \log(1 - \hat{p}))$$

Hàm mục tiêu dựa trên vùng

- Dice coefficient và IoU:

$$DC = \frac{2TP}{2TP + FP + FN} = \frac{2|X \cap Y|}{|X| + |Y|}$$

$$IoU = \frac{TP}{TP + FP + FN} = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|}$$

- Dice loss:

$$DL(p, \hat{p}) = 1 - \frac{2p\hat{p} + 1}{p + \hat{p} + 1}$$

- Tversky loss:

$$TI(p, \hat{p}) = \frac{p\hat{p}}{p\hat{p} + \beta(1 - p)\hat{p} + (1 - \beta)p(1 - \hat{p})}$$

Hàm mục tiêu kết hợp

- Dice loss + CE:

$$\text{CE}(p, \hat{p}) + \text{DL}(p, \hat{p})$$

- Dice loss + Focal loss

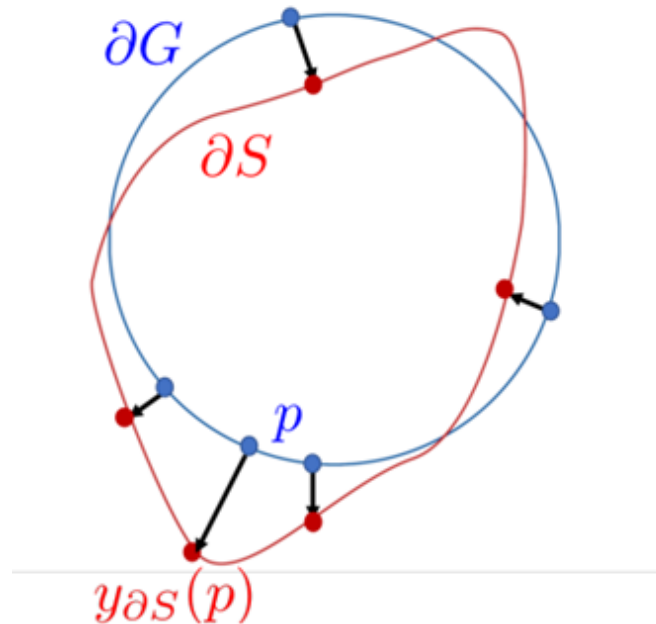
$$\text{CE}(p, \hat{p}) + \text{FL}(p, \hat{p})$$

- ...

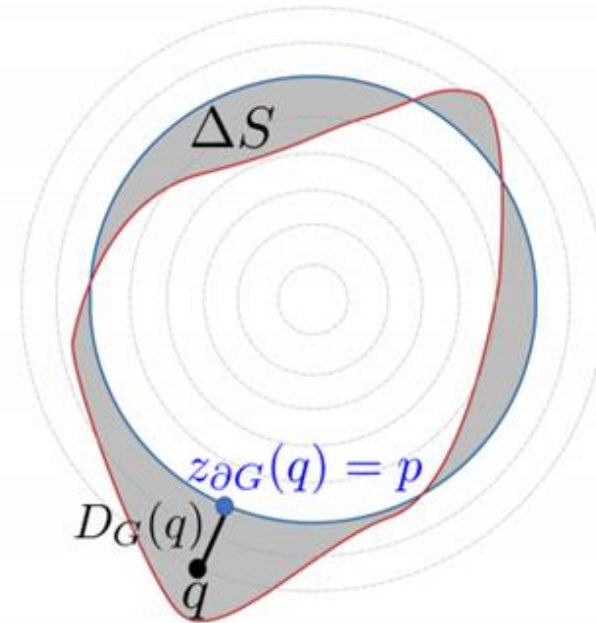
Hàm mục tiêu boundary loss

$$\text{Dist}(\partial G, \partial S) = \int_{\partial G} \|y_{\partial S}(p) - p\|^2 dp$$

$$\text{Dist}(\partial G, \partial S) = 2 \int_{\Delta S} D_G(q) dq$$



(a) Differential

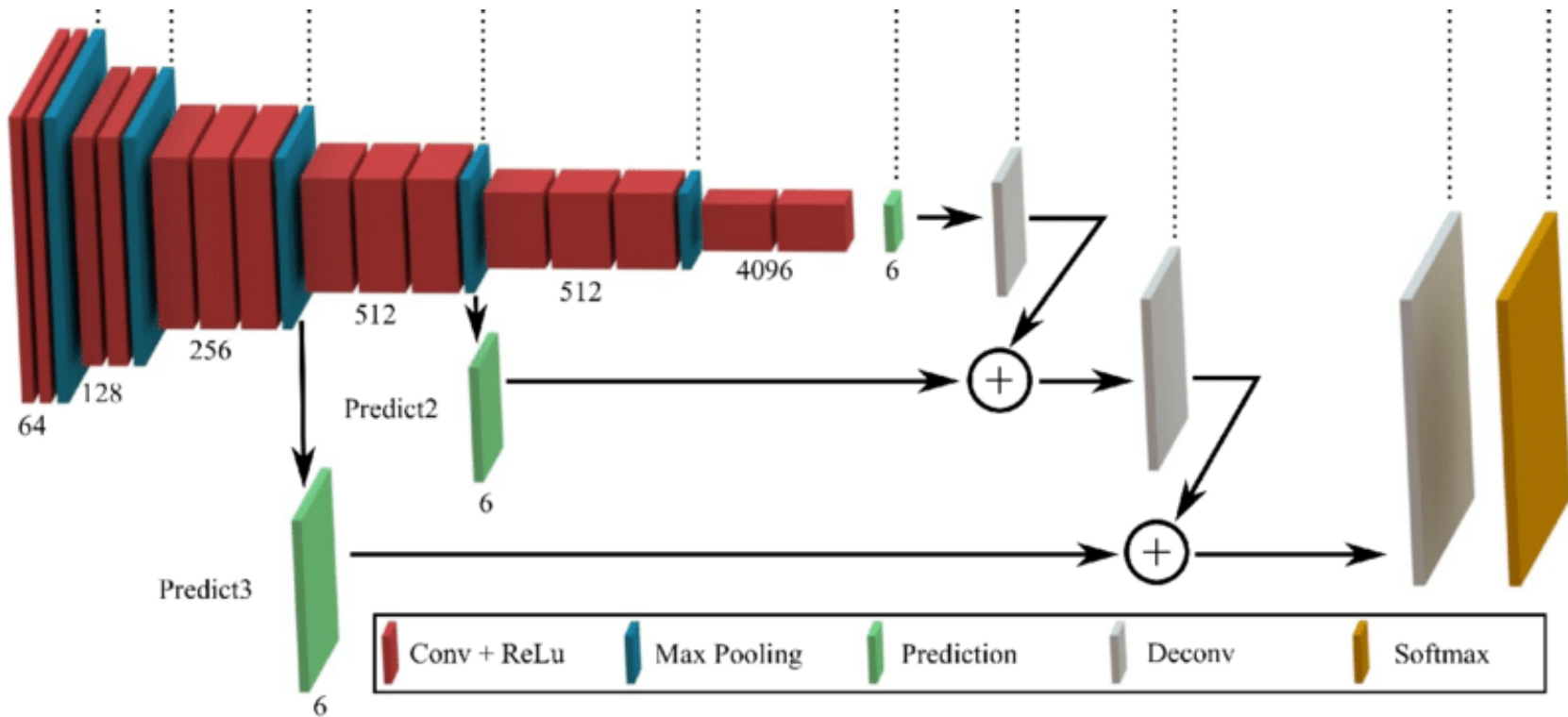


(b) Integral

$$\frac{1}{2} \text{Dist}(\partial G, \partial S) = \int_{\Omega} \phi_G(q) s(q) dq - \int_{\Omega} \phi_G(q) g(q) dq$$

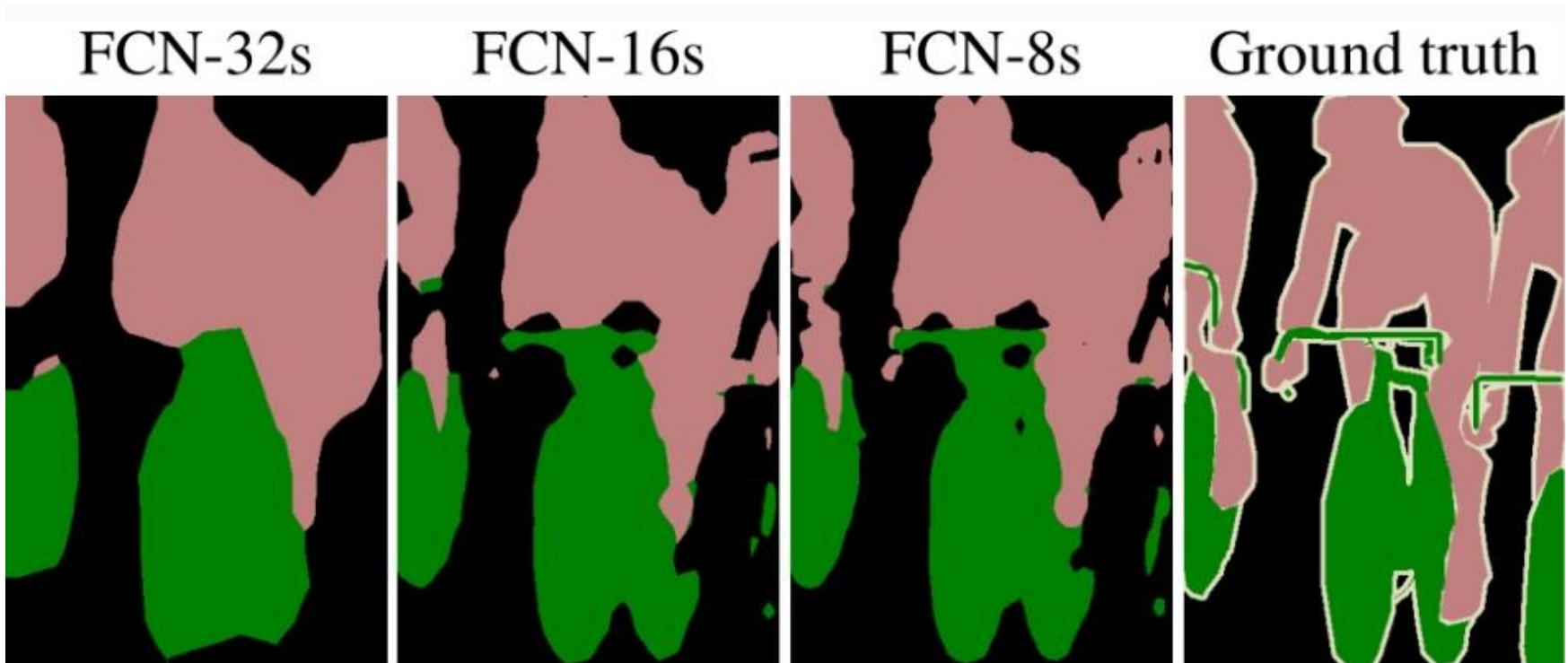
Một số mạng phân đoạn ảnh tiêu biểu

FCN với 2 kết nối tắt



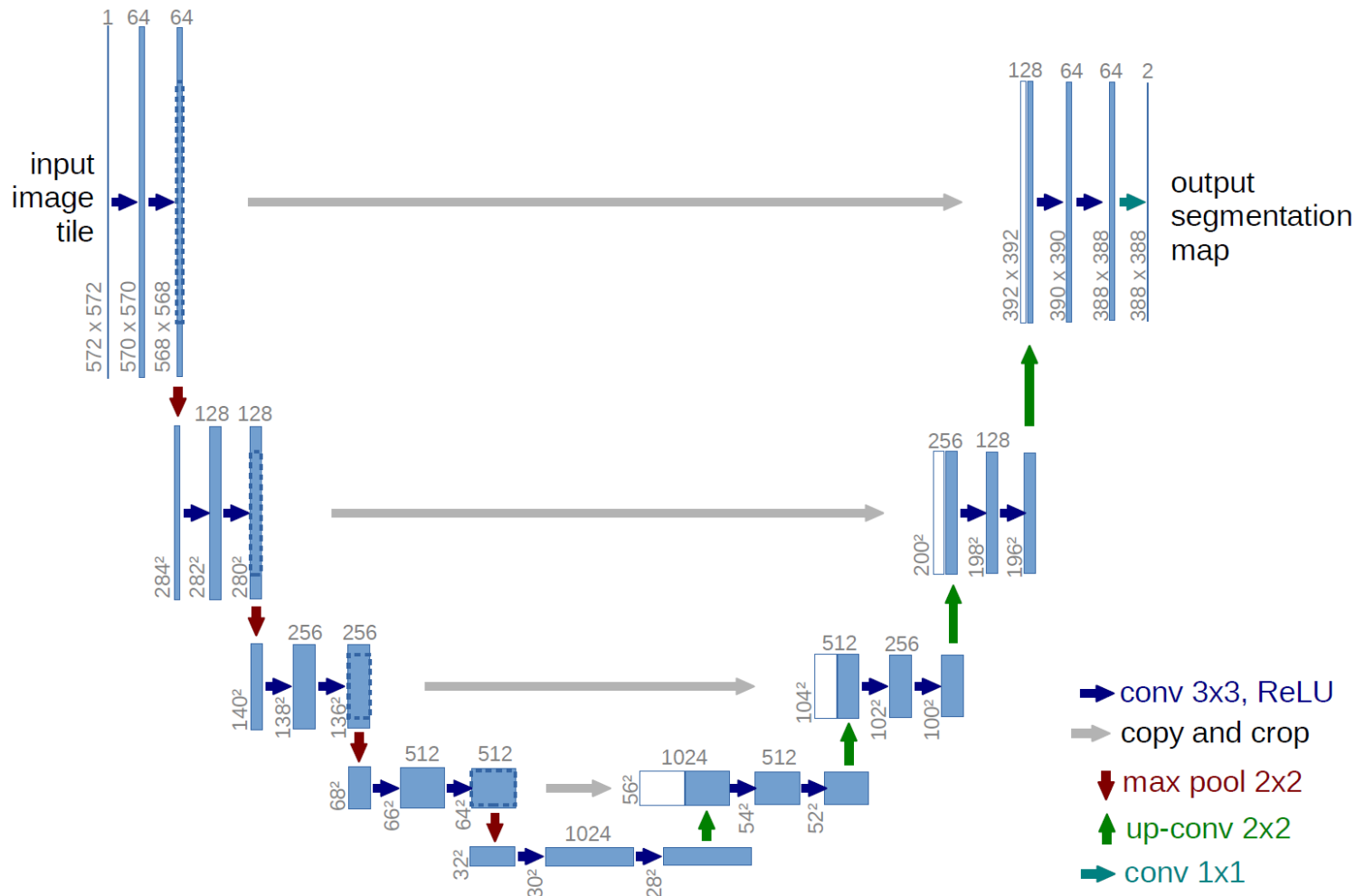
FCN với 2 kết nối tắt

- Minh họa kết quả FCN với các mức độ phân giải khác nhau

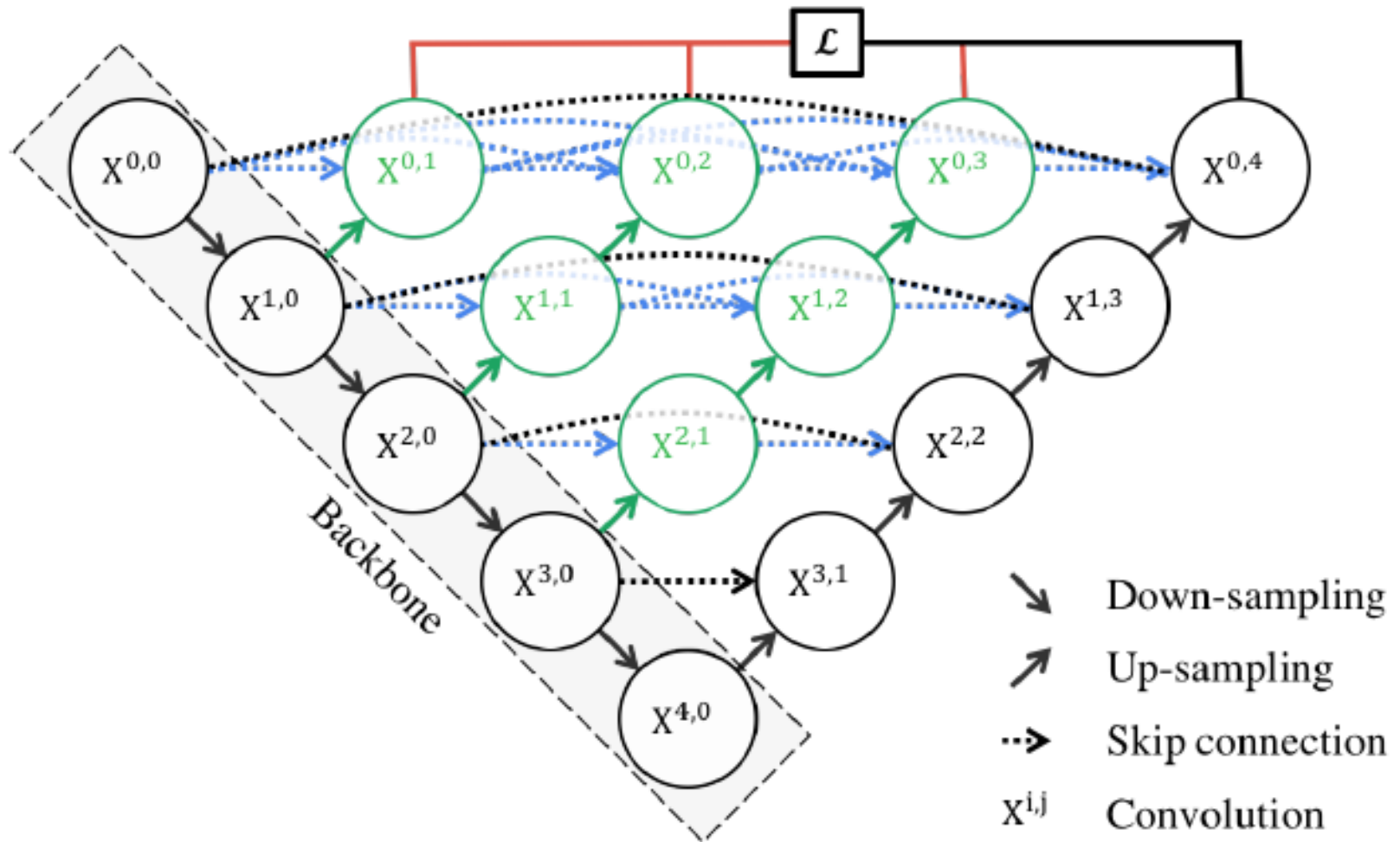


U-Net

- Được sử dụng rộng rãi trong y tế

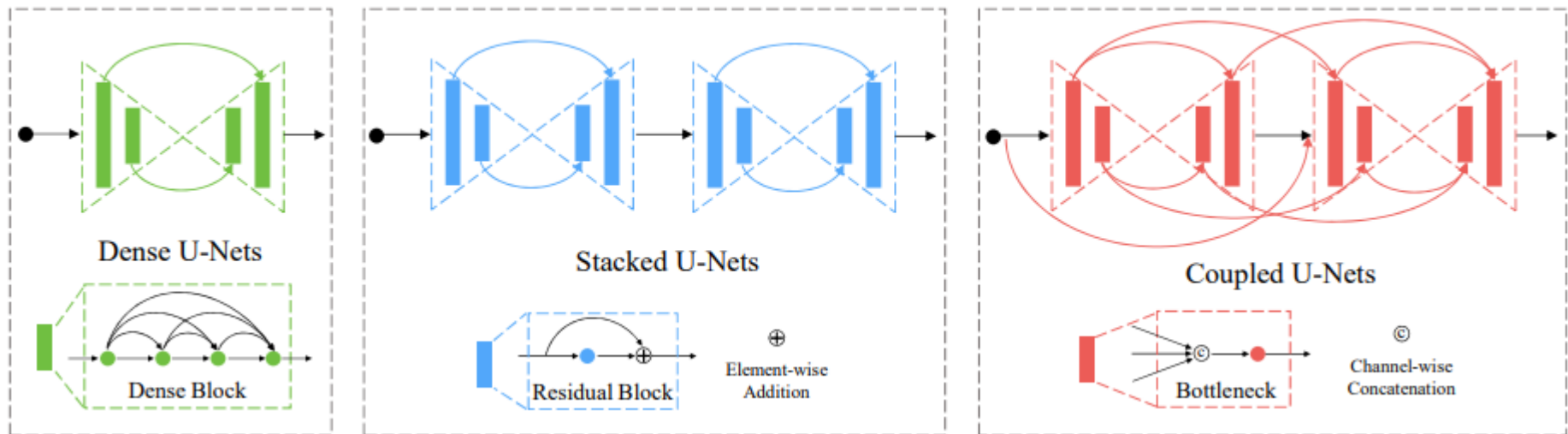


U-Net++

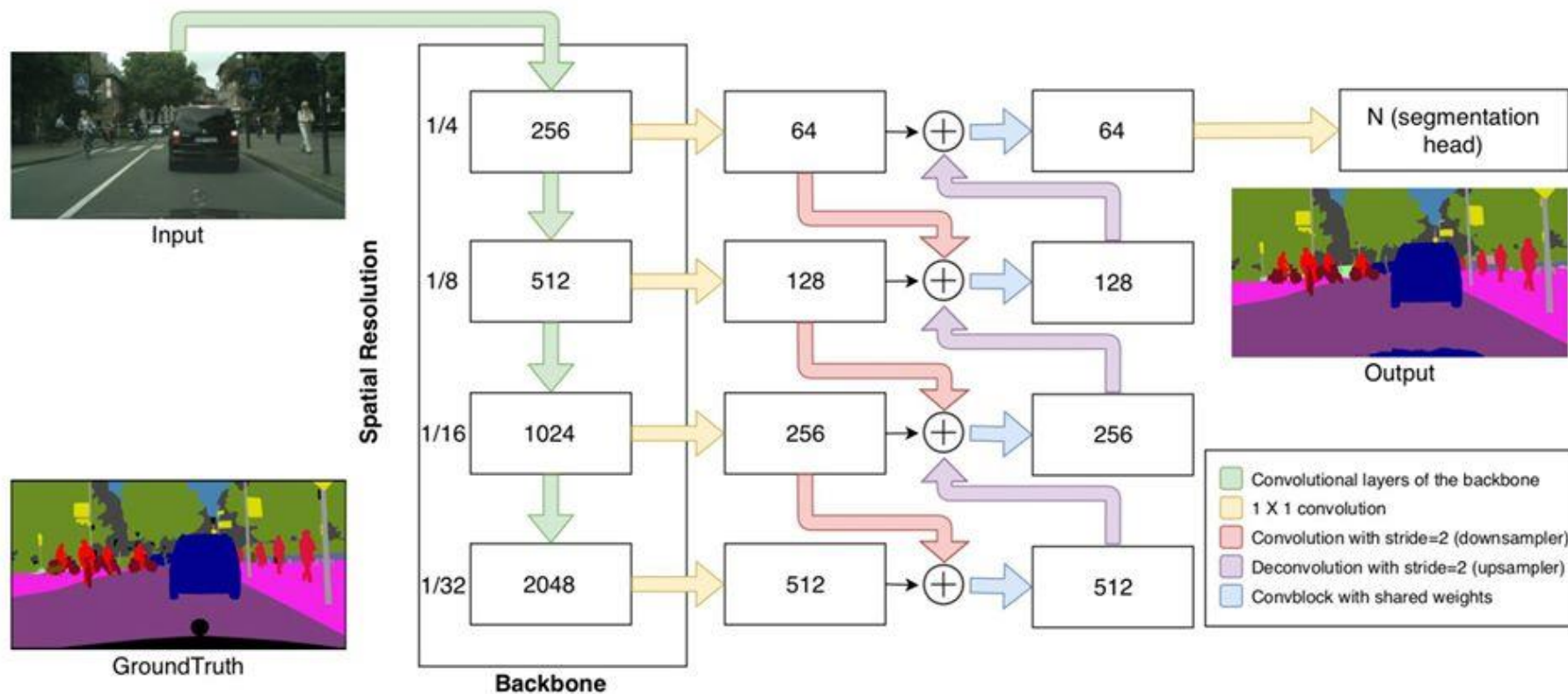


Stacked UNets và CUNets

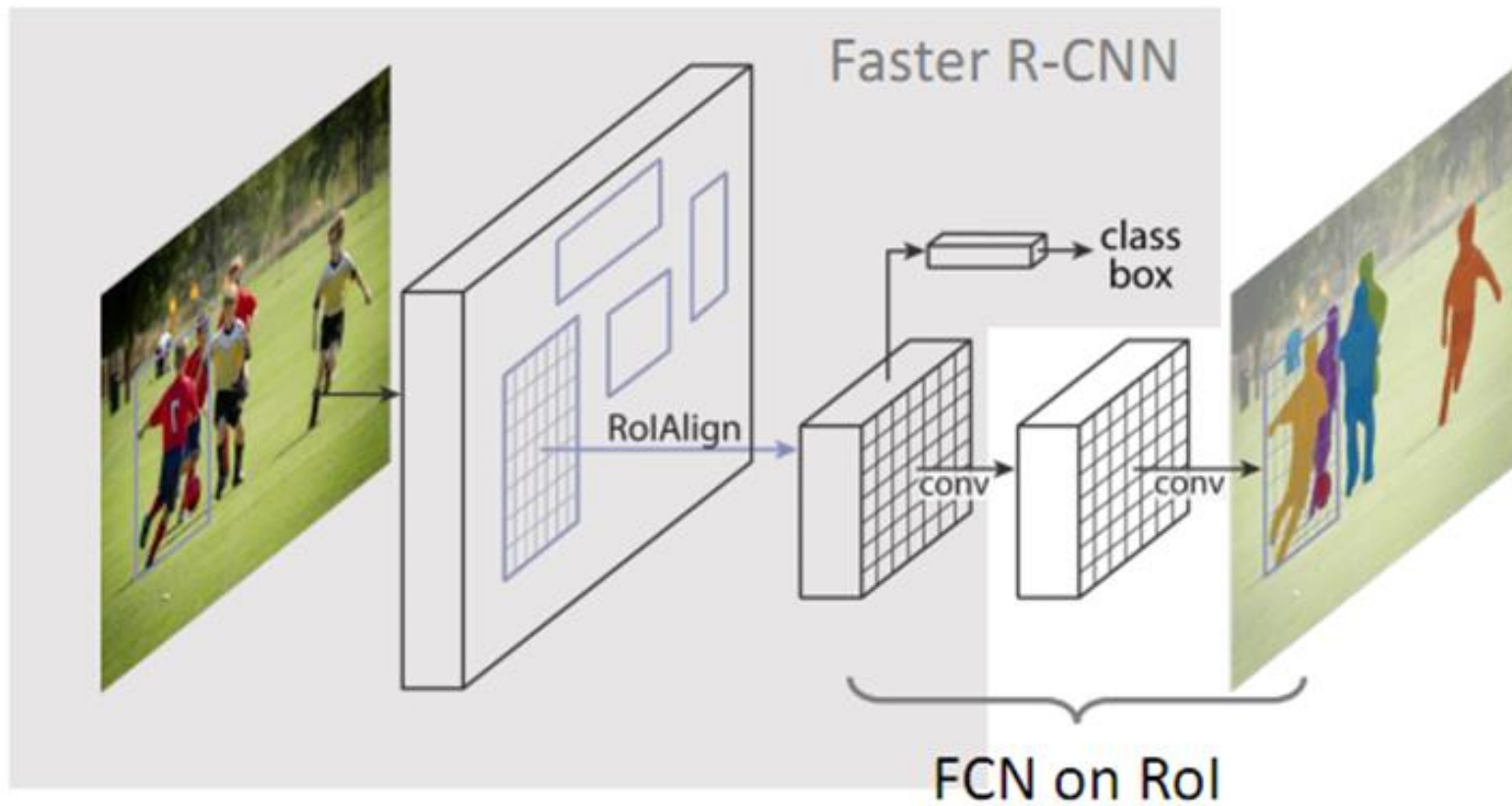
- Stacked UNets: ghép nhiều UNet nối tiếp nhau
- CUNets: cũng ghép nhiều UNet nối tiếp nhau nhưng có thêm các kết nối tắt giữa các UNet với nhau



RGPNet



Mask R-CNN



Tài liệu tham khảo

1. Khóa cs231n của Stanford:

<http://cs231n.stanford.edu>

2. Hàm mục tiêu cho bài toán phân đoạn ảnh:

<https://lars76.github.io/neural-networks/object-detection/losses-for-segmentation/>