U-NET FOR BIOMEDICAL IMAGE SEGMENTATION

U-Net and its variants

Introduction: So far Image Classification and Object Detection

This image is CC0 public domain

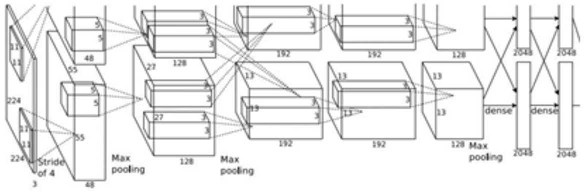
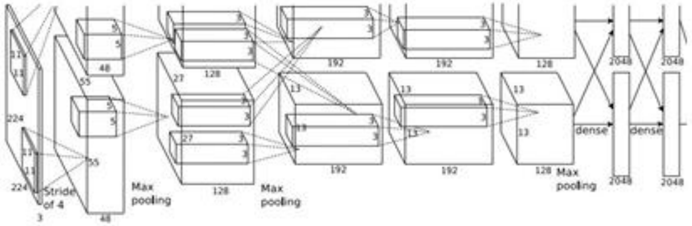


Figure copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

Vector: 4096

Fully-Connected: 4096 to 1000

Class Scores
Cat: 0.9
Dog: 0.05
Car: 0.01
...



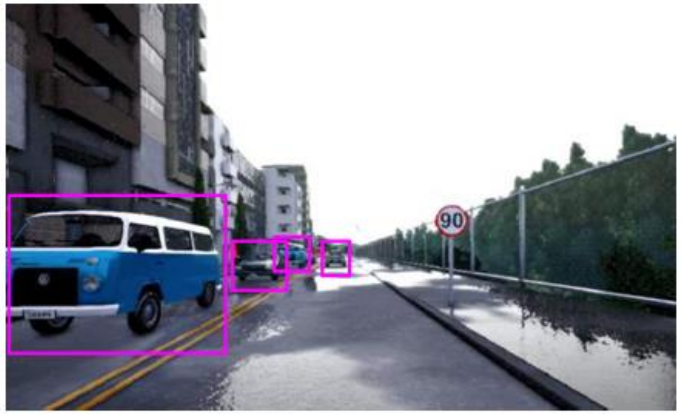
CAT: (x, y, w, h)

Introduction: Image Segmentation

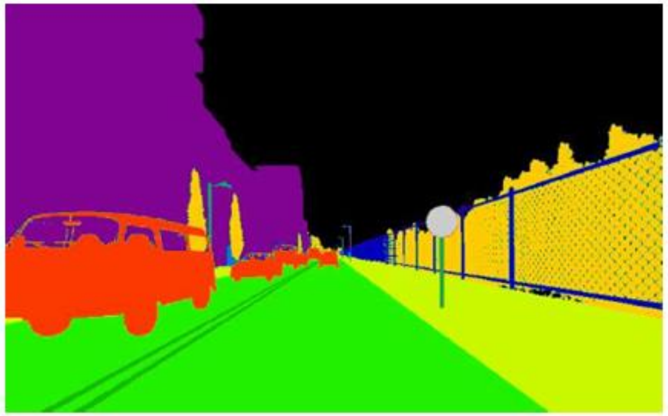
Definition: Image segmentation is a fundamental task in computer vision that involves dividing an image into multiple segments or regions, each of which corresponds to a meaningful object or part of the scene. The goal of image segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze.

Object Detection vs. Image Segmentation

Input image



Object detection



Semantic segmentation

Image Segmentation: Semantic Segmentation vs. Instance Segmentation

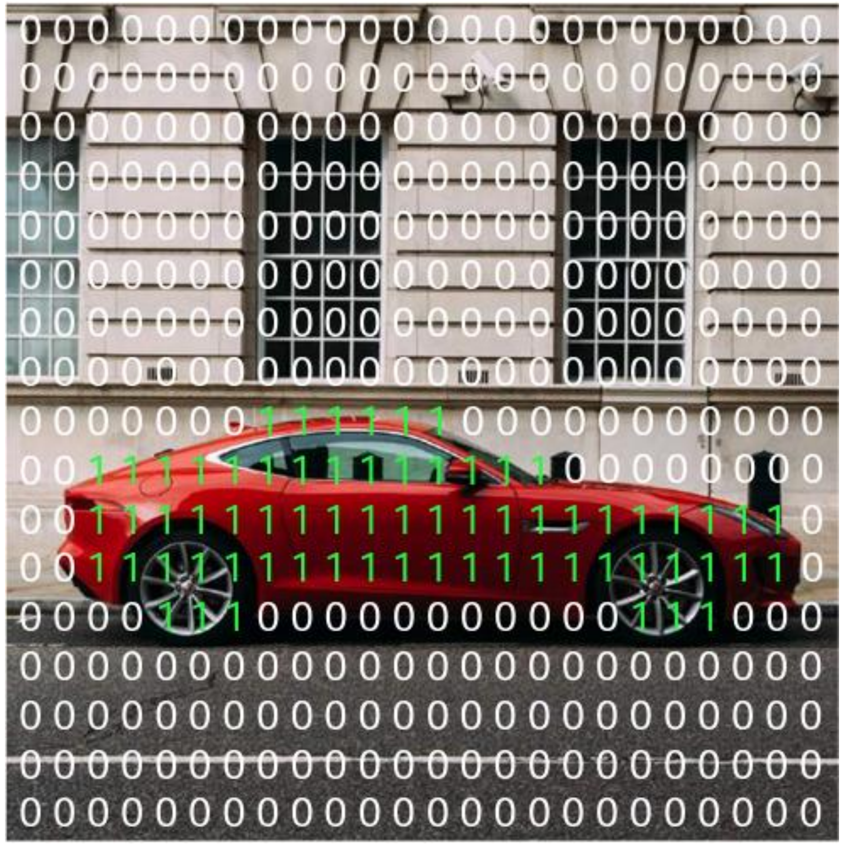
Semantic segmentation and **instance segmentation** are two different approaches to **segmenting objects** in an **image**, and they serve distinct purposes:

- **Semantic Segmentation:**
  - **Definition:** Semantic segmentation involves labeling each pixel in an image with a class label that corresponds to a specific category or object. It classifies each pixel into predefined categories, such as person, car, road, sky, etc.
  - **Output:** The output of semantic segmentation is a high-resolution map where each pixel is assigned a class label. Pixels belonging to the same class share similar visual characteristics.
  - **Example:** In an image containing a person, a car, and a road, semantic segmentation would label each pixel to indicate which class it belongs to (e.g., all pixels corresponding to the person would be labeled as "person").
  - **Use Case:** Semantic segmentation is widely used in tasks like scene understanding, object recognition, and image-to-text descriptions.

Image Segmentation: Semantic Segmentation vs. Instance Segmentation

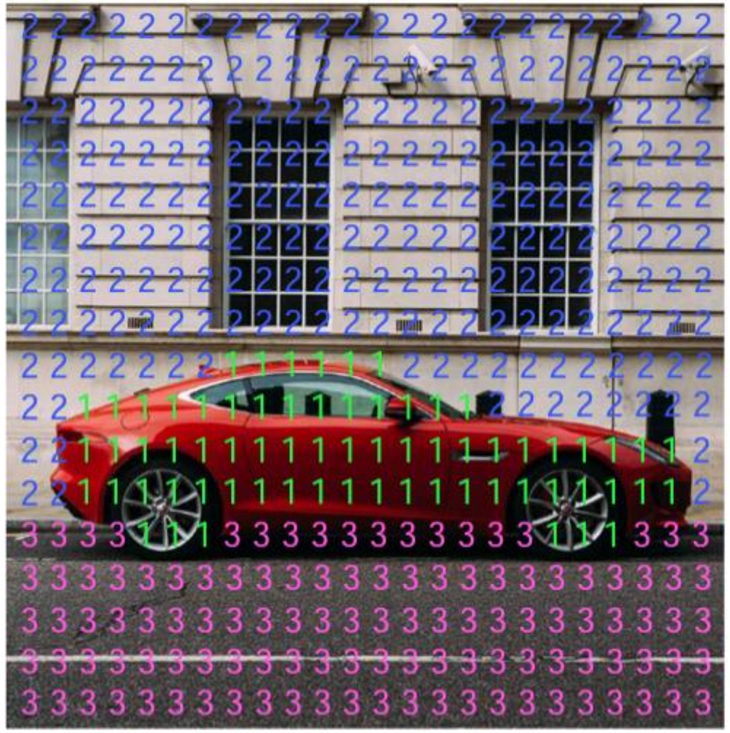
- **Instance Segmentation:**
  - Definition: Instance segmentation goes a step further than semantic segmentation. In addition to labeling each pixel with a class, it also distinguishes between individual instances of the same class. This means that if there are multiple objects of the same class (e.g., multiple cars or multiple persons) in the image, they will each be assigned a unique label.
  - Output: The output of instance segmentation provides a unique label for each object instance. This allows for precise delineation of individual objects.
  - Example: In an image with multiple cars, instance segmentation would label each car separately, assigning a unique identifier to each one.
  - Use Case: Instance segmentation is particularly important in scenarios where it is necessary to differentiate between individual objects of the same class, such as in robotics, autonomous driving, and object counting.

Per-pixel Class Labels



Class 0 = Not a car
Class 1 = Car

Per-pixel Class Labels



1. Car
2. Building
3. Road

2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2
2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
3	3	3	3	1	1	1	3	3	3	3	3	3	3	3	3	3	3	1	1	1	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	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	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
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

Image Segmentation: Fully Convolutional Network (FCN)

A Fully Convolutional Network (FCN) is a type of neural network architecture designed for semantic segmentation tasks in computer vision. Semantic segmentation involves classifying each pixel in an image into specific categories, enabling the understanding of the scene at a pixel level. FCNs are particularly well-suited for tasks where spatial information is crucial, such as object detection and image segmentation.

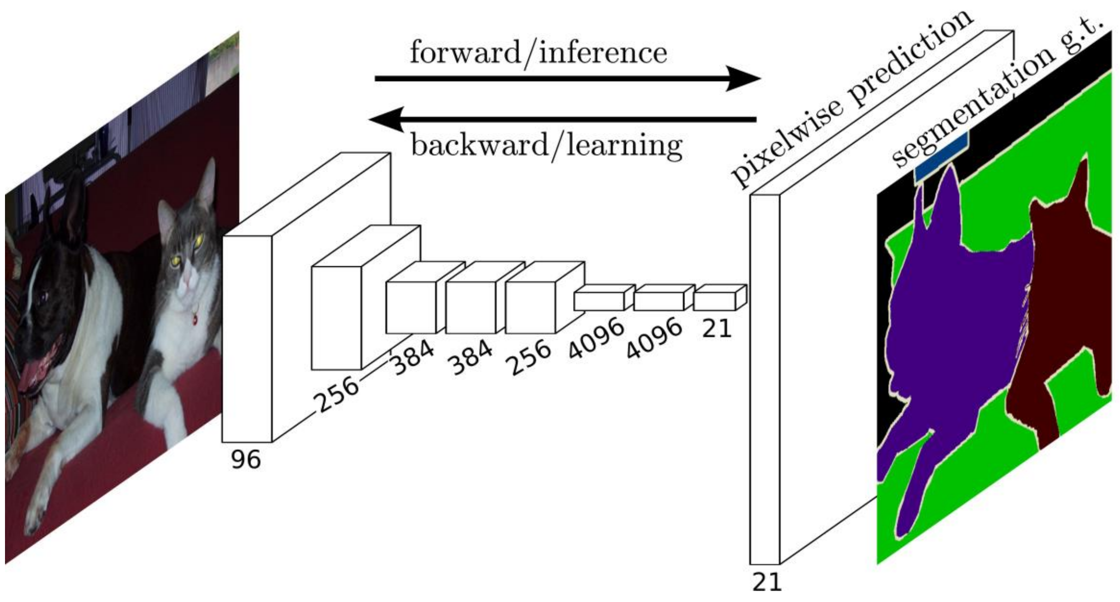
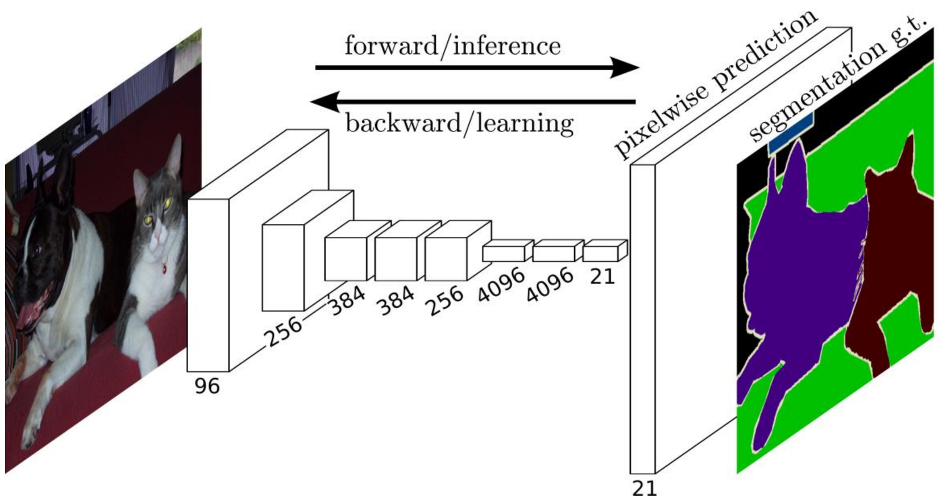


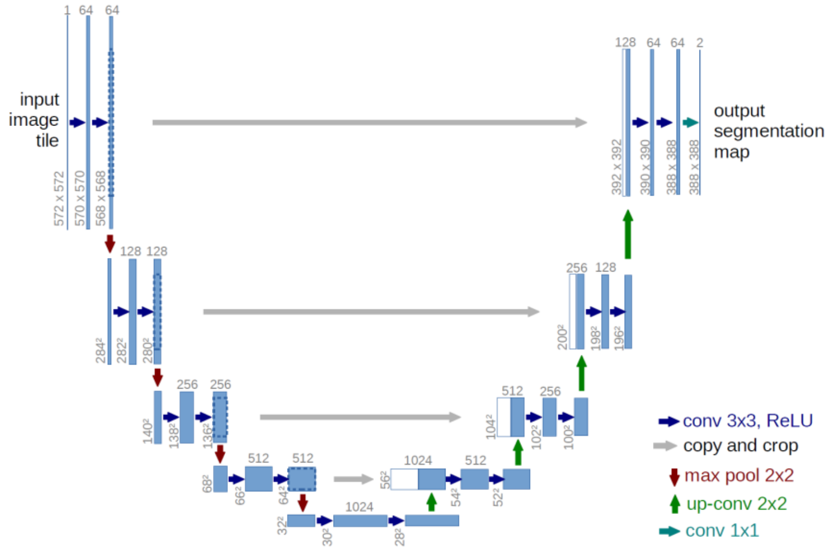
Image Segmentation: Fully Convolutional Network (FCN)

The original Fully Convolutional Network (FCN) architecture was introduced by Jonathan Long, Evan Shelhamer, and Trevor Darrell in the paper titled "Fully Convolutional Networks for Semantic Segmentation" in 2015. FCNs have since become a fundamental building block for many state-of-the-art models in semantic segmentation and related computer vision tasks. **Variants** of **FCNs**, such as **U-Net**, **SegNet**, and **Deeplab**, have been developed to address specific challenges and improve performance in different scenarios.

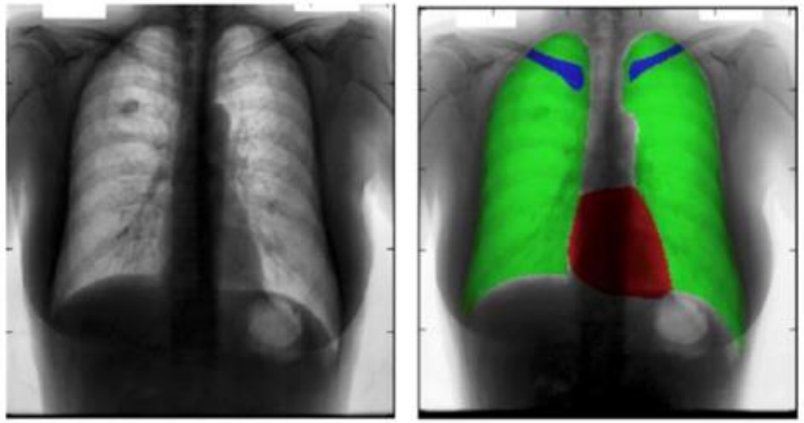


U-NET: Introduction

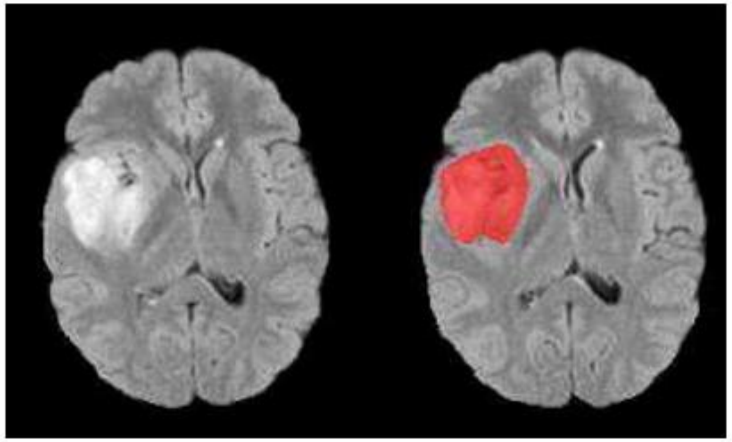
- **U-Net** is a popular **convolutional neural network architecture** designed for image segmentation tasks. It was introduced by **Olaf Ronneberger**, **Philipp Fischer**, and **Thomas Brox** in their 2015 paper titled "U-Net: Convolutional Networks for Biomedical Image Segmentation."
- The architecture of **U-Net** is characterized by its **U-shaped structure (Figure on the right)**, which consists of a **contracting path (left side)** and an **expansive path (right side).** This unique design allows the network to capture both local and global information, making it **highly effective for tasks like biomedical image segmentation.**
- **Fun Fact:** When the researchers wrote the original U-NET paper, they were thinking of the application of biomedical image segmentation. But these ideas turned out to be useful for many other computer vision semantic segmentation applications as well.



Motivations for U-NET



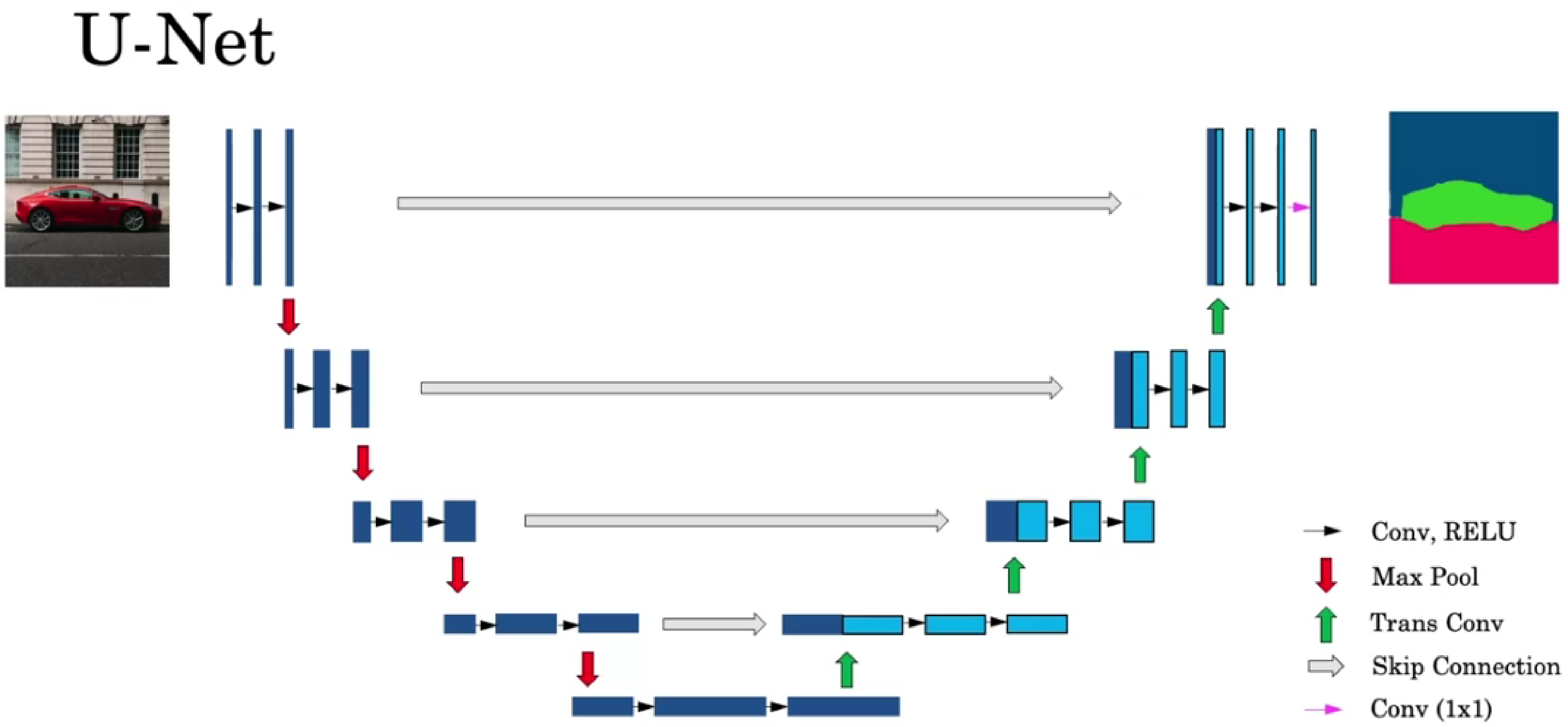
Chest X-Ray



Brain MRI

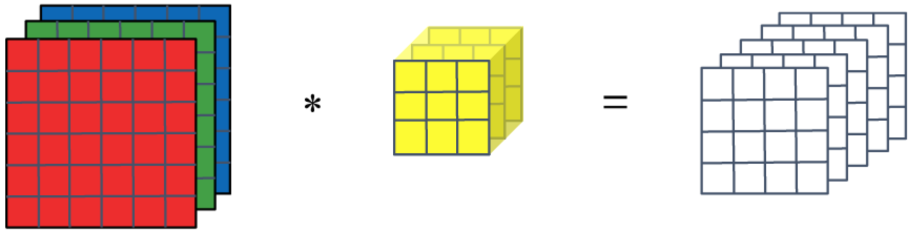
- [Novikov et al., 2017, Fully Convolutional Architectures for Multi-Class Segmentation in Chest Radiographs]
- [Dong et al., 2017, Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks]

U-NET: Architecture



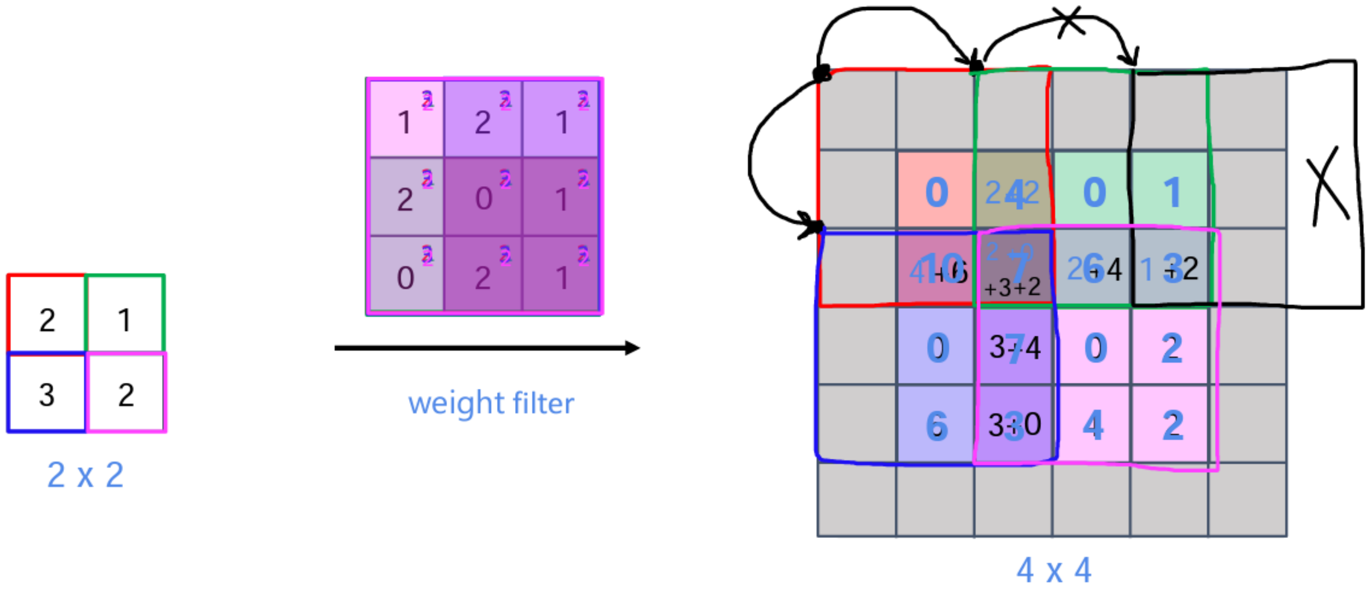
U-NET: Transpose Convolution

Normal Convolution



Transpose Convolution

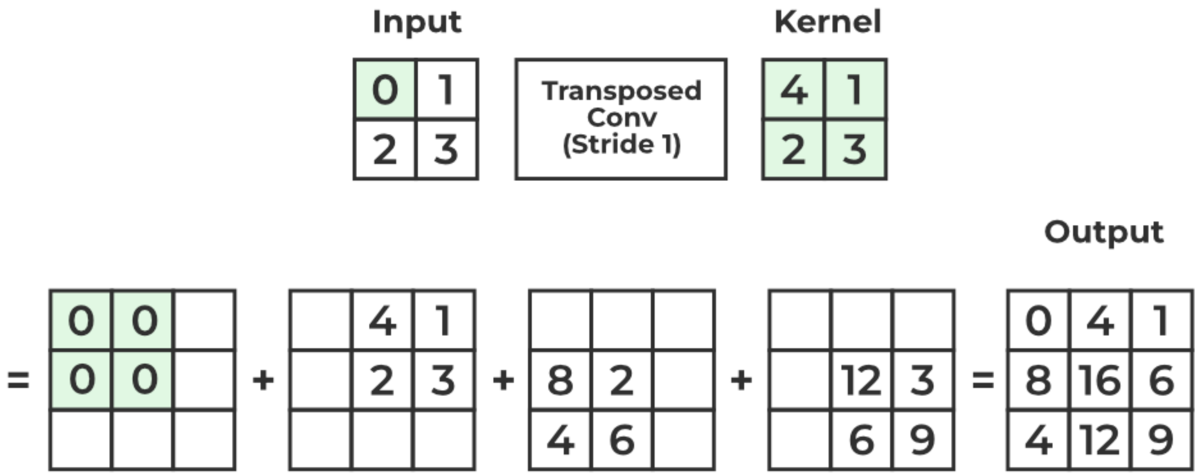
U-NET: Transpose Convolution



filter f x f = 3 x 3 padding p = 1 stride s = 2

Source: [Coursera]

U-NET: Transpose Convolution



Deep Learning in Medical Images