Deep learning for medical imaging

Artificial Intelligence for Digital Health (AID)

M.Sc. in Digital Health – University of Pisa

Davide Bacciu (davide.bacciu@unipi.it)

Lecture(s) Outline

- Introduction to medical imaging
    - Image representation
    - Medical imaging as an inverse problem
    - Imaging modalities and their challenges
- Convolutional neural networks
    - Convolutional layers, filters/kernels, feature maps
    - Pooling layers and their role
    - Convolutional architectures and useful architectural tools
- Medical imaging tasks
    - classification, regression, segmentation, detection, registration, enhancement

Medical Imaging in Healthcare Practice

High Utilization in Healthcare

- Over 4.2 billion diagnostic medical imaging procedures performed globally each year

Impact on Diagnosis & Treatment

- Critical for cancer detection, surgical planning, chronic disease management
- Radiology drives ~80% of hospital diagnoses

Challenges

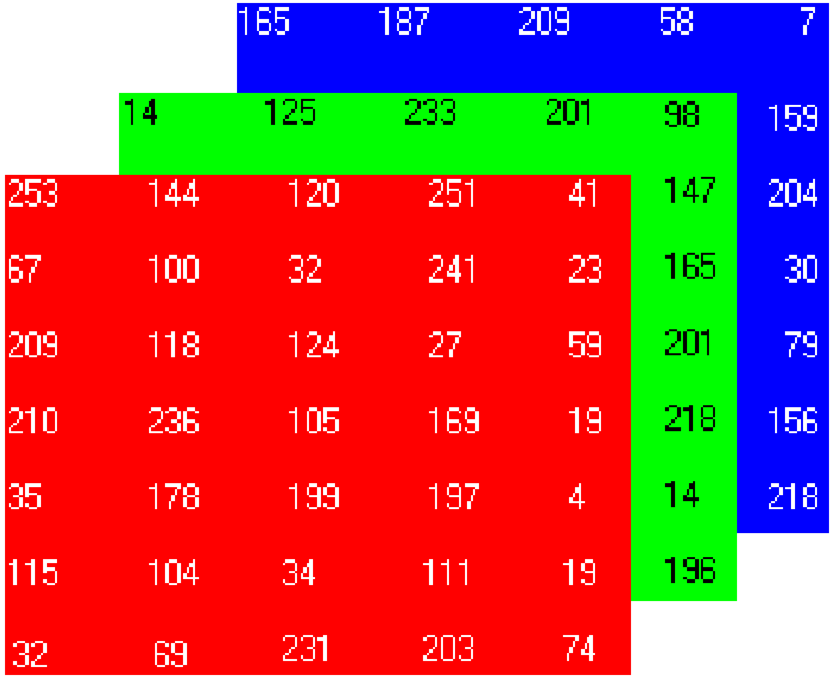
- Huge data volume → Necessitates automation and AI
- Variability in acquisition, reconstruction parameters

Image Data & Imaging Modalities

(General) Image Data

0	2	15	0	0	11	10	0	0	0	0	9	9	0	0	0
0	0	0	4	60	157	236	255	255	177	95	61	32	0	0	29
0	10	16	119	238	255	244	245	243	250	249	255	222	103	10	0
0	14	170	255	255	244	254	255	253	245	255	249	253	251	124	1
2	98	255	228	255	251	254	211	141	116	122	215	251	238	255	49
13	217	243	255	155	33	226	52	2	0	10	13	232	255	255	36
16	229	252	254	49	12	0	0	7	7	0	70	237	252	235	62
6	141	245	255	212	25	11	9	3	0	115	236	243	255	137	0
0	87	252	250	248	215	60	0	1	121	252	255	248	144	6	0
0	13	113	255	255	245	255	182	181	248	252	242	208	36	0	19
1	0	5	117	251	255	241	255	247	255	241	162	17	0	7	0
0	0	0	4	58	251	255	246	254	253	255	120	11	0	1	0
0	0	4	97	255	255	255	248	252	255	244	255	182	10	0	4
0	22	206	252	246	251	241	100	24	113	255	245	255	194	9	0
0	111	255	242	255	158	24	0	0	6	39	255	232	230	56	0
0	218	251	250	137	7	11	0	0	0	2	62	255	250	125	3
0	173	255	255	101	9	20	0	13	3	13	182	251	245	61	0
0	107	251	241	255	230	98	55	19	118	217	248	253	255	52	4
0	18	146	250	255	247	255	255	255	249	255	240	255	129	0	5
0	0	23	113	215	255	250	248	255	255	248	248	118	14	12	0
0	0	6	1	0	52	153	233	255	252	147	37	0	0	4	1
0	0	5	5	0	0	0	0	0	14	1	0	6	6	0	0

A matrix of NxM pixels with values in [0 - 255]



Three matrixes of NxM pixels with values in [0 - 255], one for red, green and blue channels

⇒ A NxMx3 tensor (not a mathematical tensor!)

Biomedical Images & Vision Tasks

Imaging methodologies use in healthcare generate very characteristic data (a.k.a. need dedicated solutions)

- Different modalities (spectral, temporal, spatiotemporal)
- Heterogeneity (multispectral, multimodal)
- Acquisition noise (ultrasound)

Main imaging modalities

- Radiology (X-ray, CT, MRI)
- Oncology imaging (PET scans, specialized MRI for tumor detection)
- Pathology (digital slides)
- Ultrasound, endoscopy, and other modalities

Fundamental tasks in medical imaging

- Classification (detection of disease)
- Regression (e.g., lesion size or tumor volume)
- Segmentation & detection (identifying and delineating tumors, organs, or structures)
- Registration (aligning structures between 2 different images)
- Enhancement (denoising, artifact removal, augmentation)

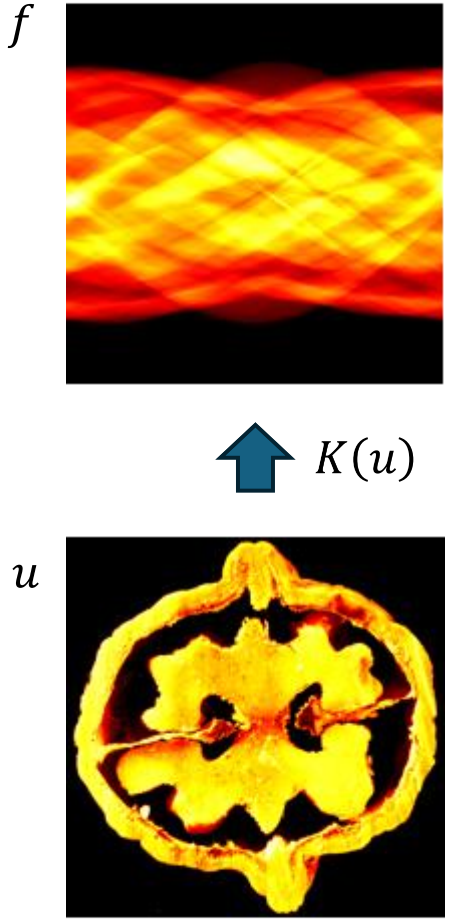
Physics of Biomedical Imaging – General Principle

Image formation (general principle)

- Emission or transmission of a wave (electromagnetic or acoustic)
- Detect and measure wave attenuation or reflection/scattering to reconstruct an image

Key mathematics/physics ideas

- Reconstructing an image u from measured signals f is an example of an inverse problem: i.e. extracting u from $K(u) = f$ where $K()$ is an operator (integral, linear system, …) modeling the physical process
- Modalities differ by type of wave (X-rays, radiofrequency for MRI, sound waves for ultrasound, positrons for PET)



Sinogram of CT Scan

CT scan image of walnut

Image source: here

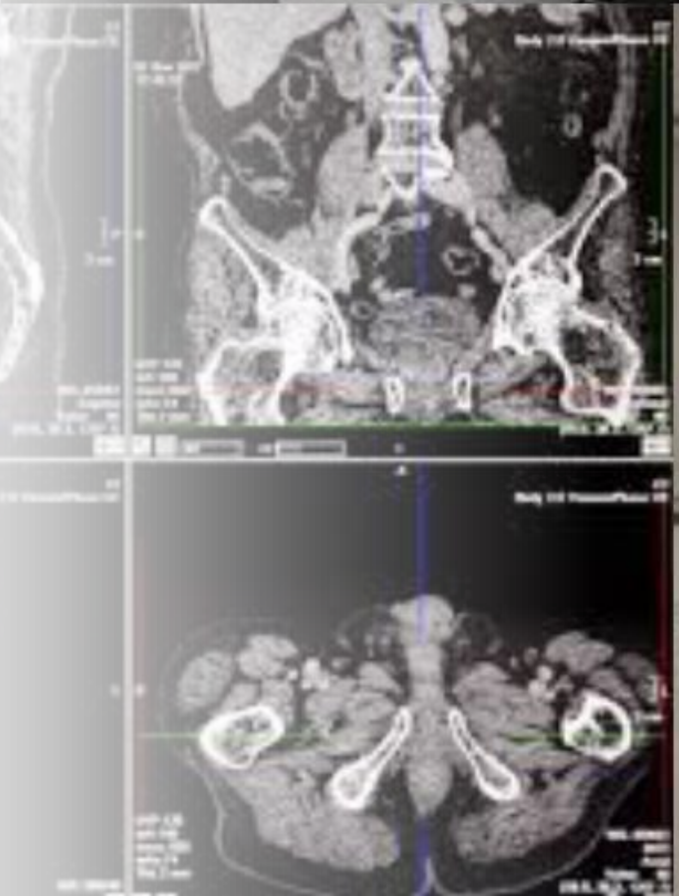
Radiology – X-ray, Computed Tomography

X-ray

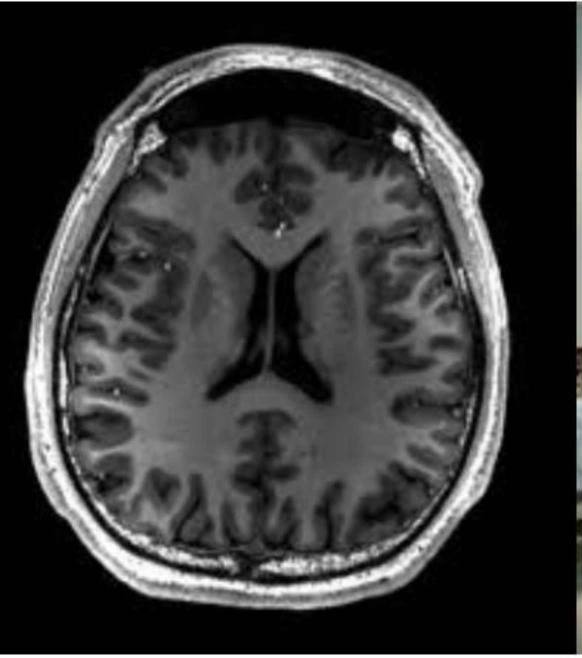
- 2D projection imaging using X-ray photons
- Attenuation depends on tissue density
- Applications: Chest radiographs, bone fractures

Computed Tomography (CT)

- Multiple X-ray projections from different angles
- Reconstructed via Radon transform or filtered back-projection
- Generates 3D volumetric data



Source: here



Source: here

Radiology –Magnetic Resonance Imaging, Ultrasound

Magnetic Resonance Imaging (MRI)

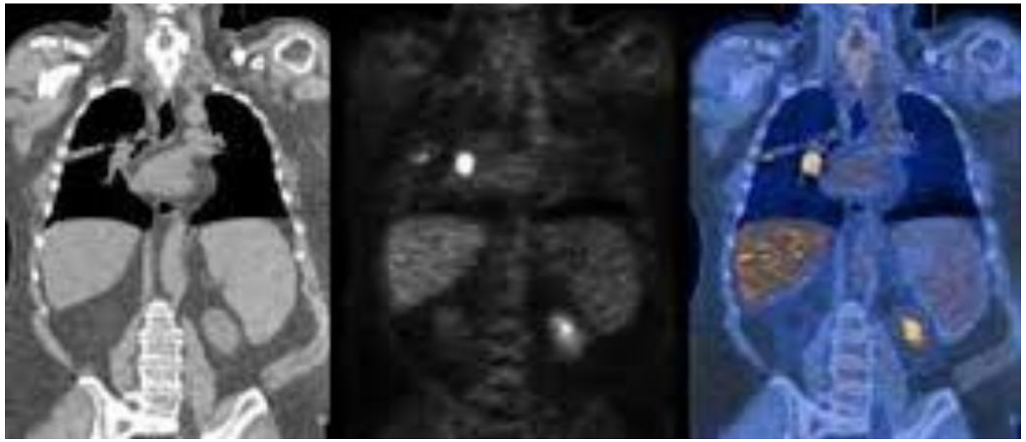
- Manipulates proton spin alignments via strong magnetic fields & RF pulses
- Signal measured in k-space, reconstructed via inverse Fourier transform
- Good soft-tissue contrast

Ultrasound Imaging (USI)

- Uses high-frequency sound waves, reflection captured by a transducer
- Real-time imaging, widely used for obstetrics, cardiac echo
- Safe (no ionizing radiation), but operator-dependent
- Short video (temporal data)

Oncology Imaging – PET & Specialized MRI

Source: here

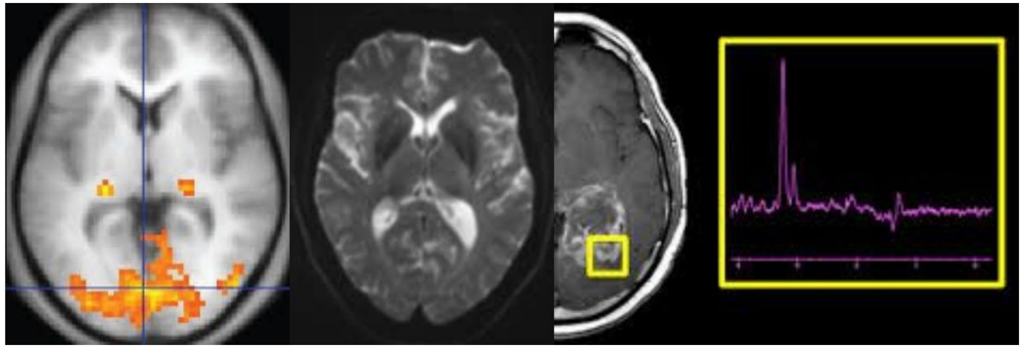


Positron Emission Tomography (PET)

- Inject radioactive tracer that emits positrons
- Detect annihilation photons, reconstruct distribution of tracer uptake
- Highlights metabolic activity, commonly used for tumour detection and staging

Specialized MRI

- fMRI for brain function mapping
- DWI/ADC for tumour characterization and cellularity
- MRS (Magnetic Resonance Spectroscopy) for metabolic profiling



fMRI DWI MRS

Pathology Imaging - Digital Slides & Advanced Stains

Digital Pathology

- High-resolution scanning of tissue slides (large magnification)
- Resulting images can be gigapixel-level

Types of Microscopy & Staining

- H&E (Hematoxylin & Eosin): Standard stain for tissue morphology
- Histochemical stains: Highlight specific chemical components
- Immunohistochemistry (IHC): Antibody-based staining for specific proteins
- In situ hybridization: Detect specific nucleic acid sequences

Source: here

Other Modalities & 4D Imaging

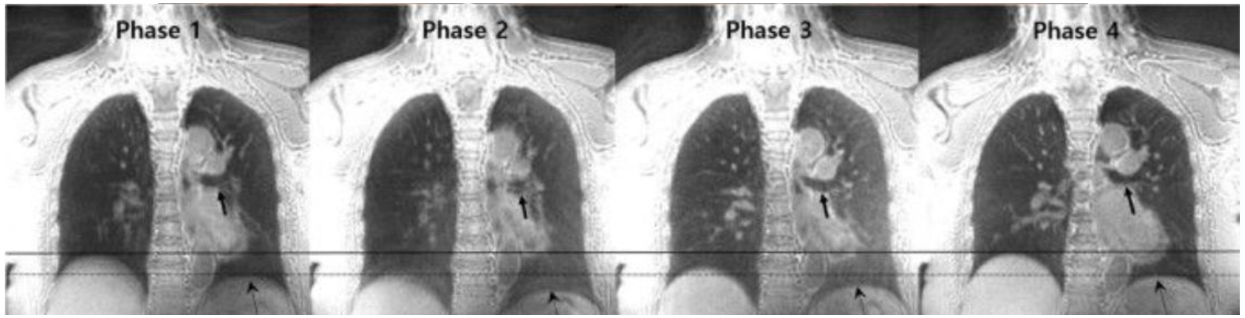
Endoscopy

- Direct visualization using cameras inserted into body cavities
- Often recorded as video (temporal dimension)

4D Imaging

- 3D + time: 4D CT in radiotherapy planning for moving organs (lungs)
- Real-time MRI sequences

Source: here



Convolutional Neural Networks

Dense Vector Multiplication

Processing images: the dense (MLP) way

32x32x3 image

Reshape it into a vector

$\boldsymbol{x}$

3072

An input-sized weight vector for each hidden neuron

$\boldsymbol{W}$

100x3072

$\boldsymbol{W}\boldsymbol{x}^{T}$

Each element contains the activation of 1 neuron

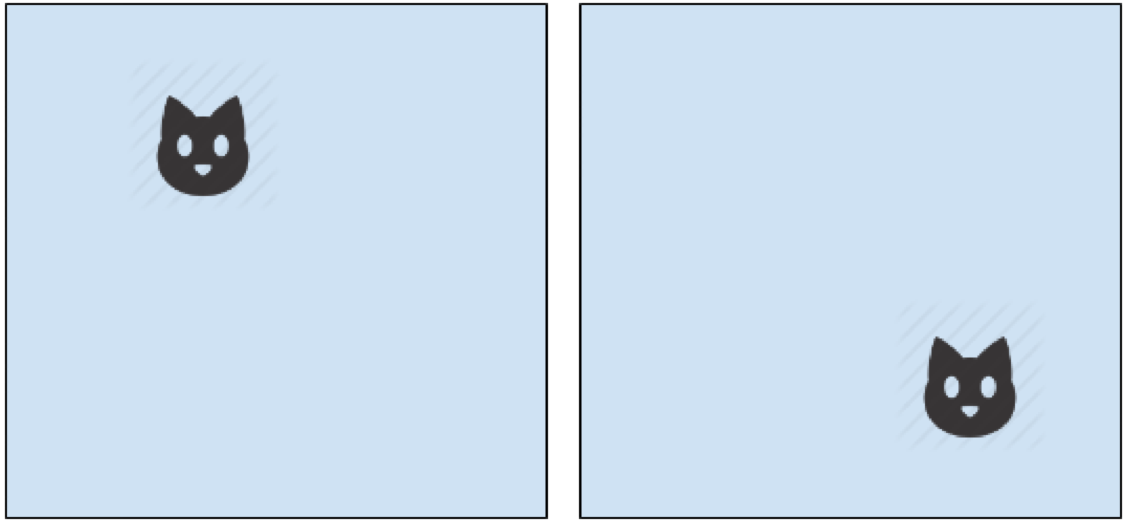
100

Not a good solution considering efficiency aspects!

About invariances

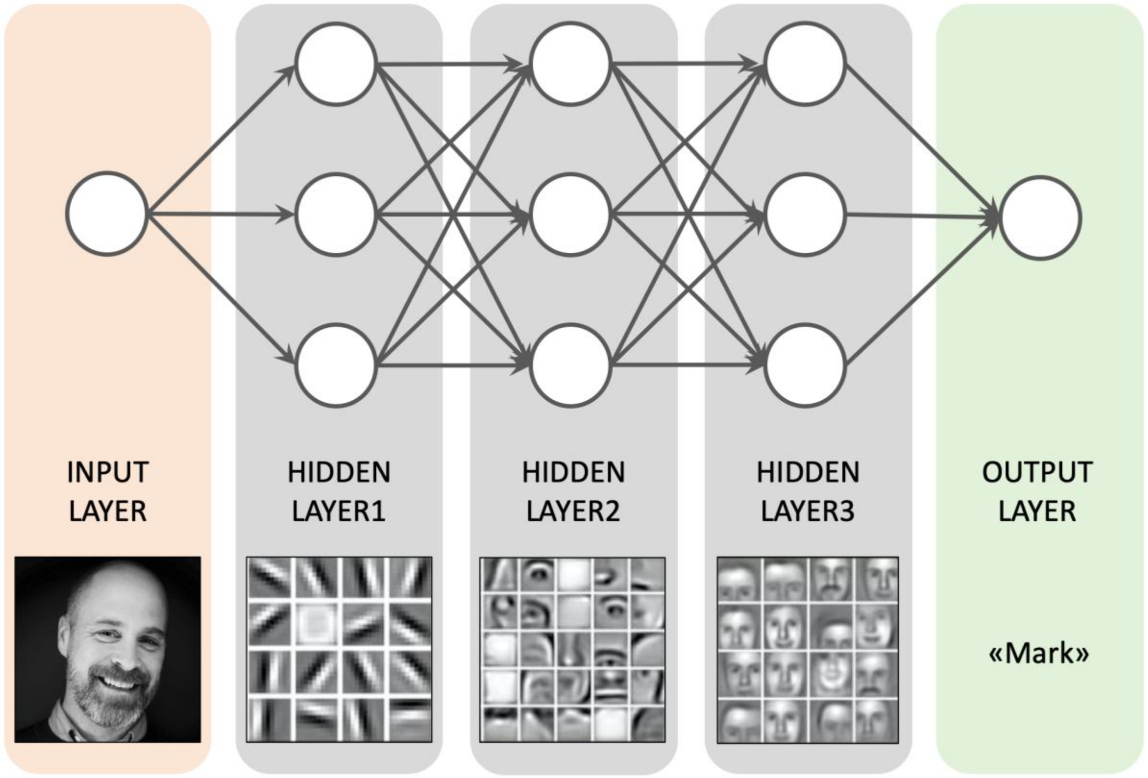
MLPs are positional

We (most likely) need translation invariance!



- If we unfold the two images into two vectors, the features identifying the cat will be in different positions
- But this still remains a picture of a cat, which we would like to classify as such irrespectively of its position in the image

About compositionality



Images have a compositional nature: simple features are composed into increasingly more complex ones

An inductive bias to keep in mind

Nearby pixels are more correlated than far away ones

The input representation should not destroy pixel relationships (like vectorization does)

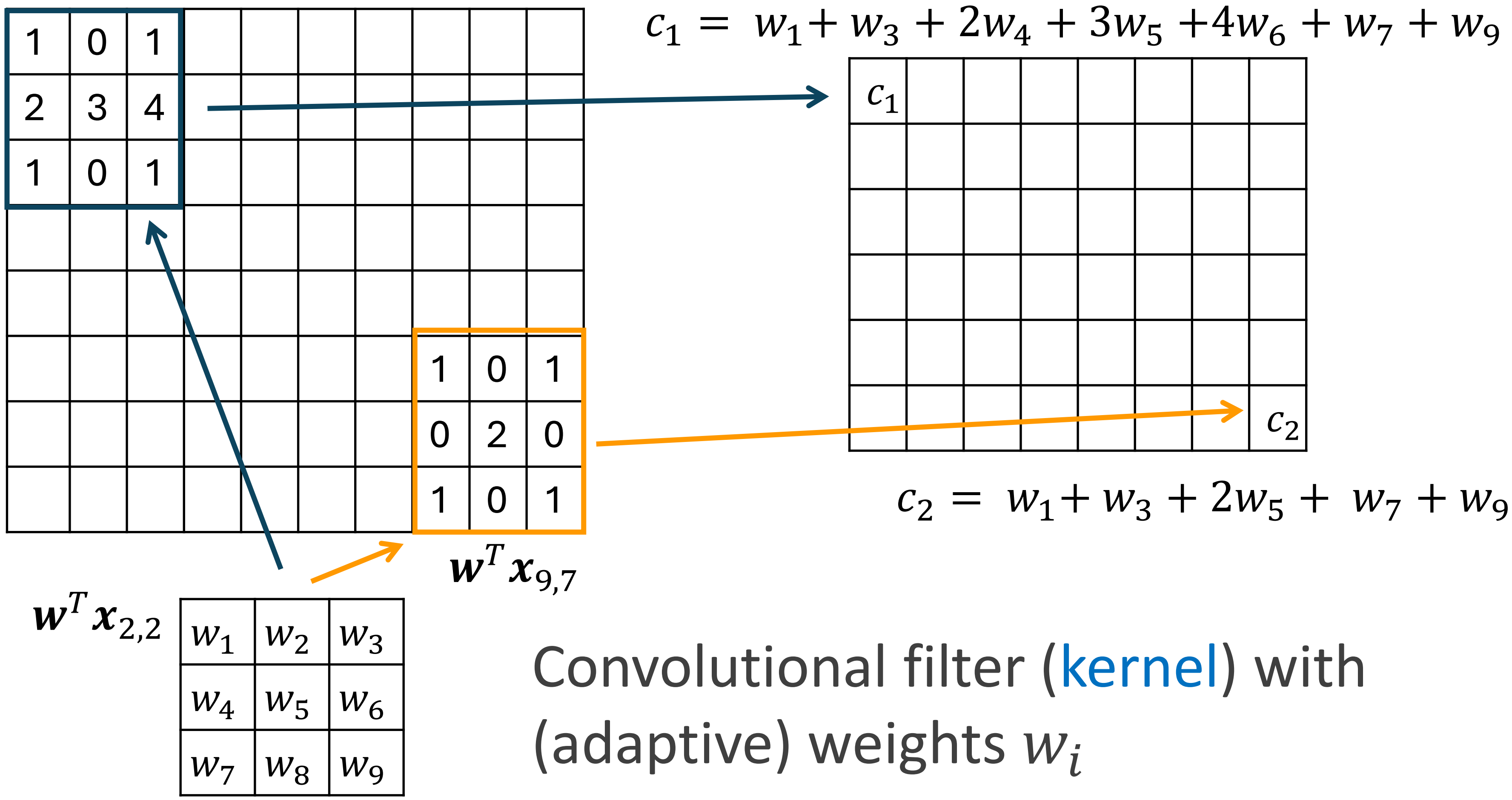
Wrapping-up

We need a neural architecture that is

- Efficient
- Translation invariant
- Compositional in nature
- Capable of exploiting pixel relationships

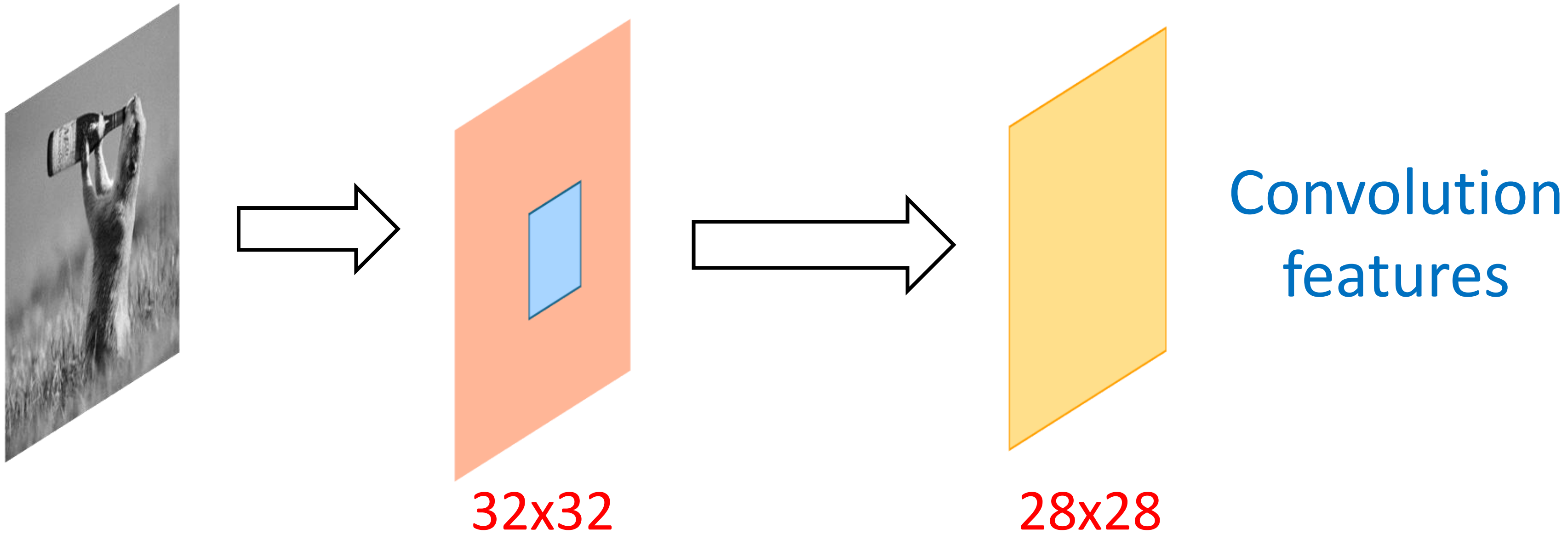
The answer we give today is: Convolutional Neural Networks

Convolution – A Linear Combination Operator



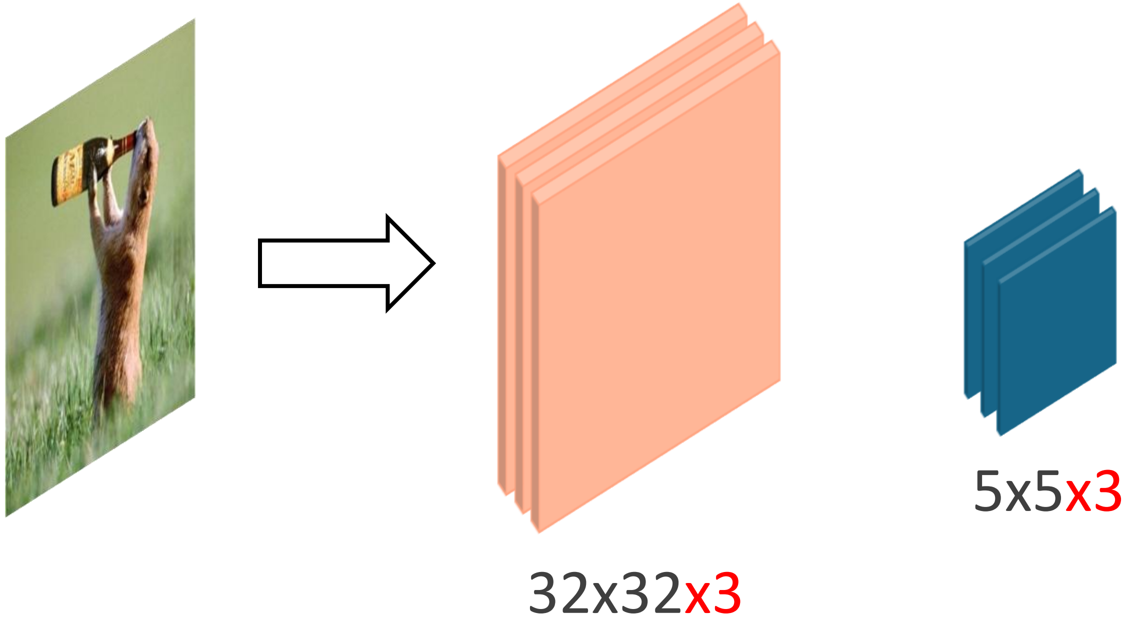
Convolutional filter (kernel) with (adaptive) weights w_i

Convolutional Features



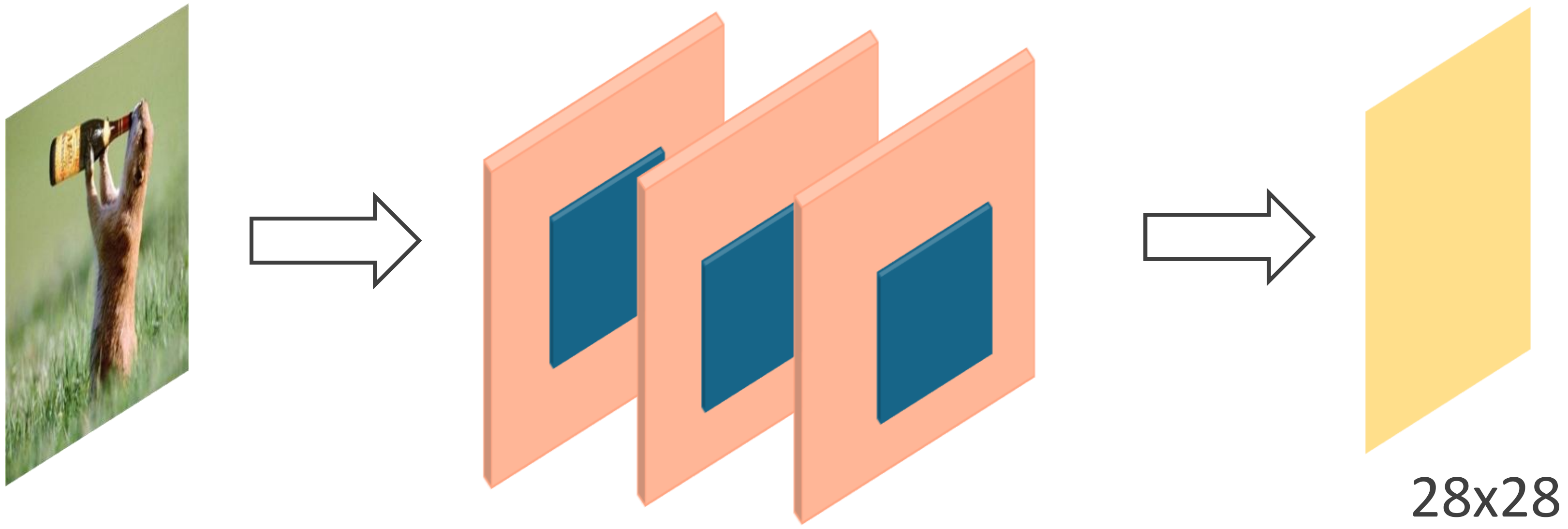
Slide the filter on the image computing elementwise products and summing up

Multi-Channel Convolution



Convolution filter has a number of slices equal to the number of image channels

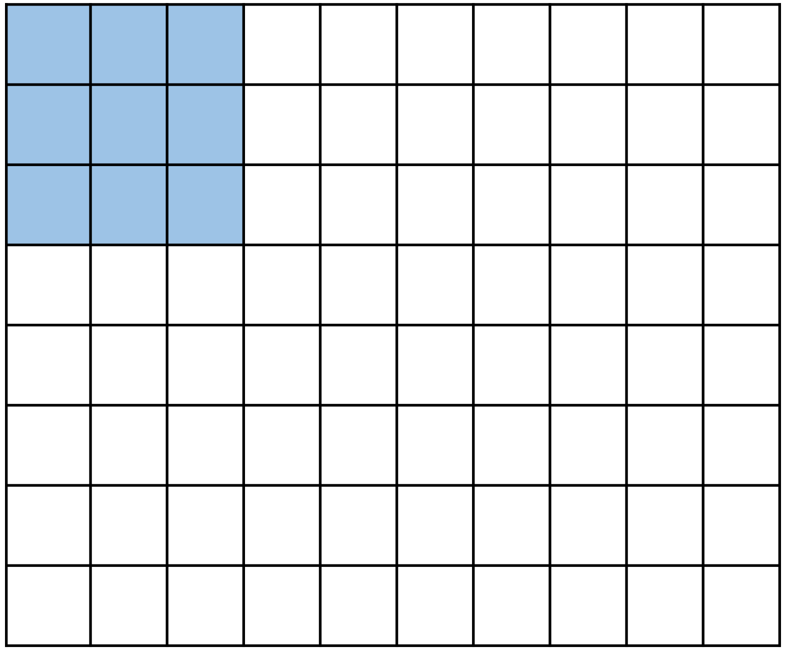
Multi-Channel Convolution



All channels are typically convolved together

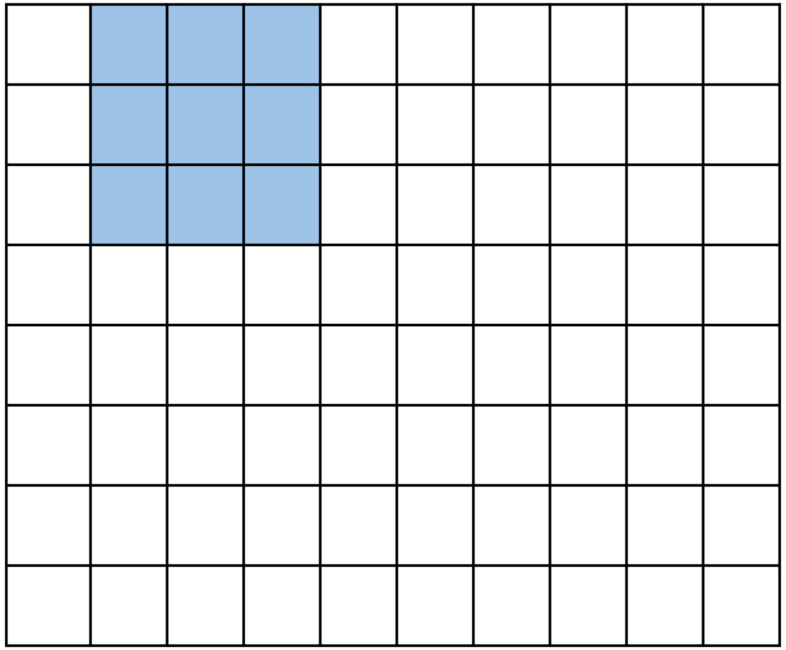
- They are summed-up in the convolution
- The convolution map stays bi-dimensional

Stride



- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1

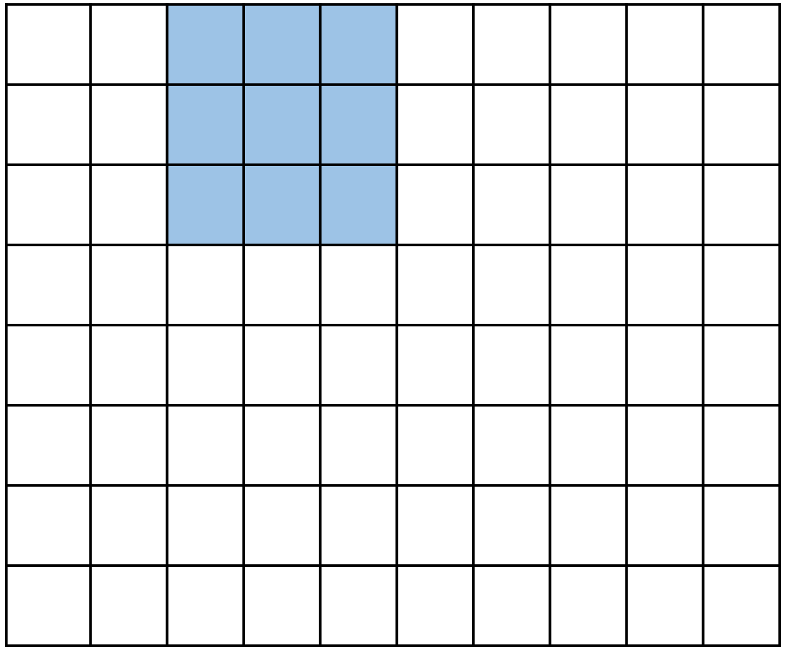
Stride



stride = 1

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1

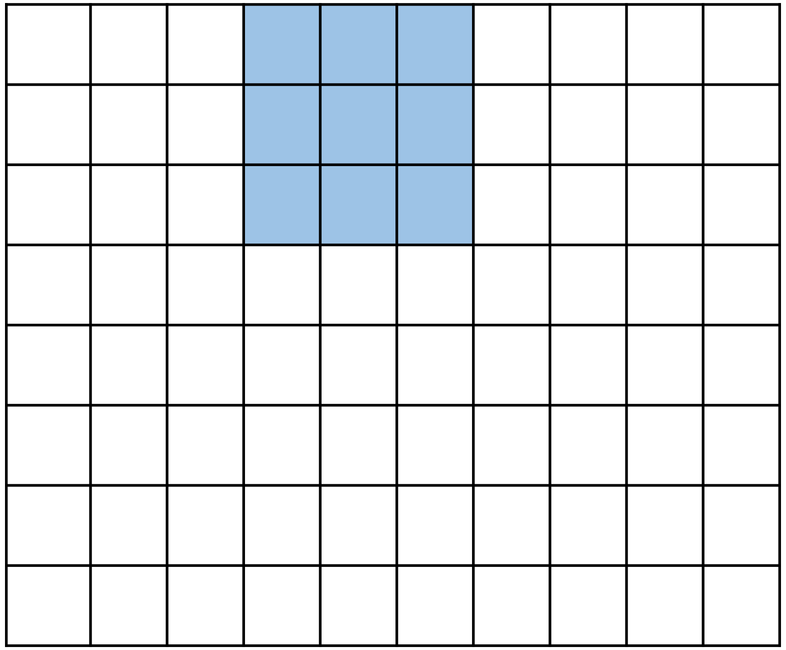
Stride



stride = 1

- Basic convolution slides the filter on the image one pixel at a time
    - Stride = 1

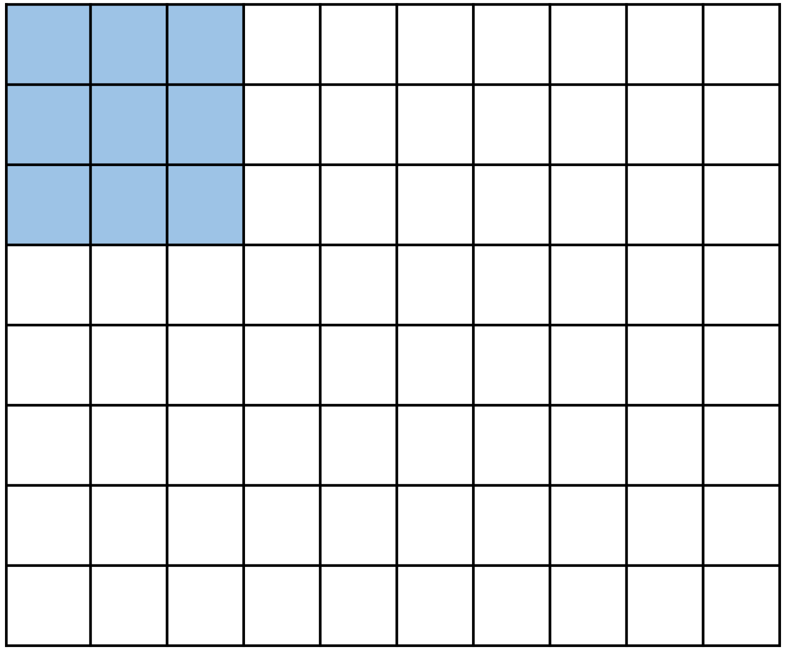
Stride



stride = 1

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1

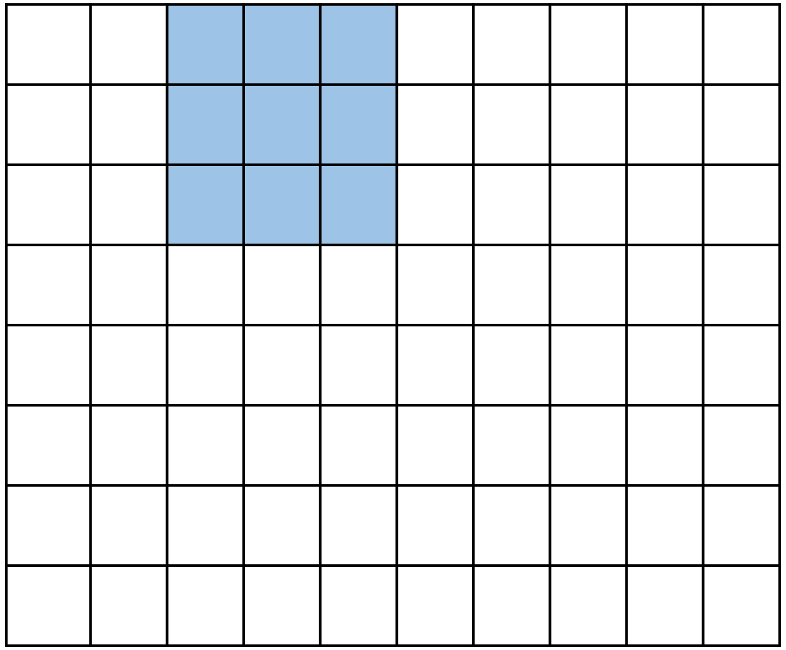
Stride



stride = 2

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1
- Can define a different stride
  - Hyperparameter

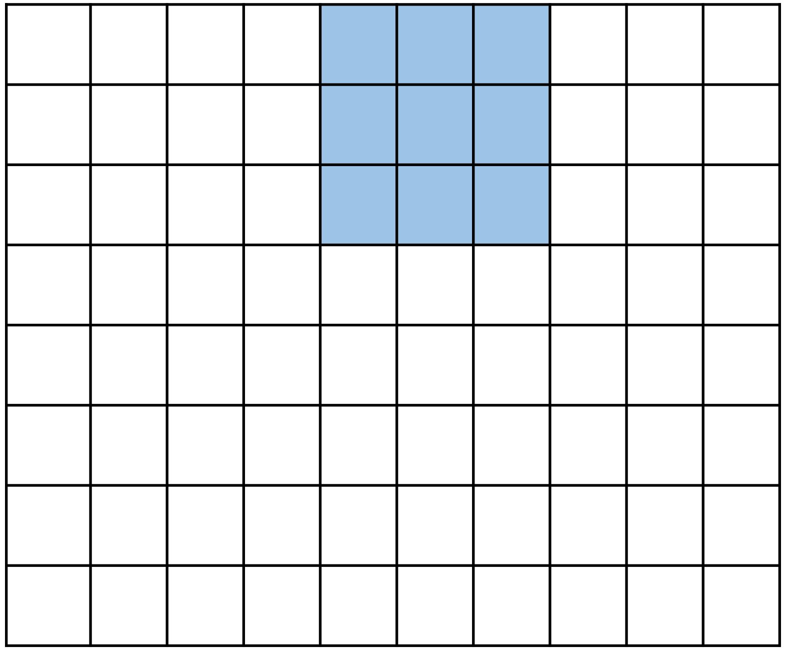
Stride



stride = 2

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1
- Can define a different stride
  - Hyperparameter

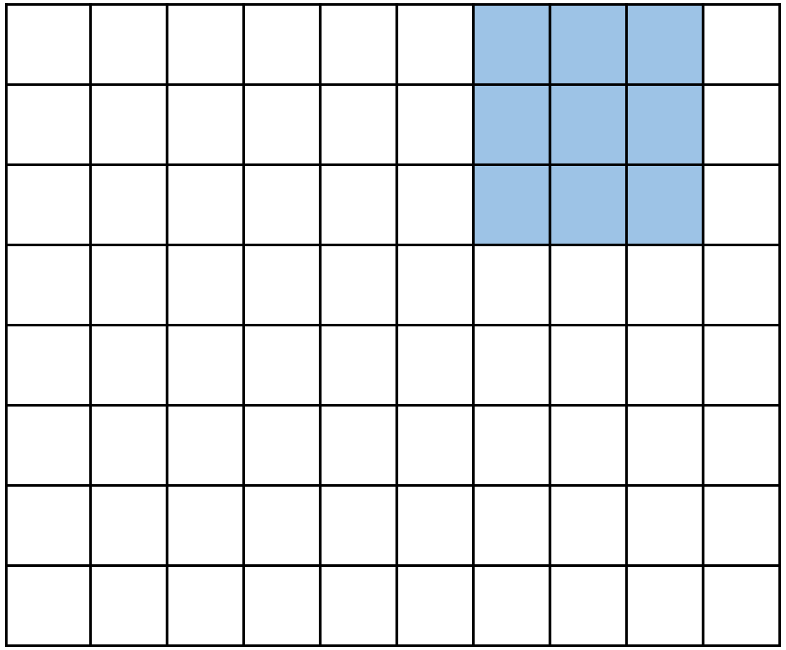
Stride



stride = 2

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1
- Can define a different stride
  - Hyperparameter

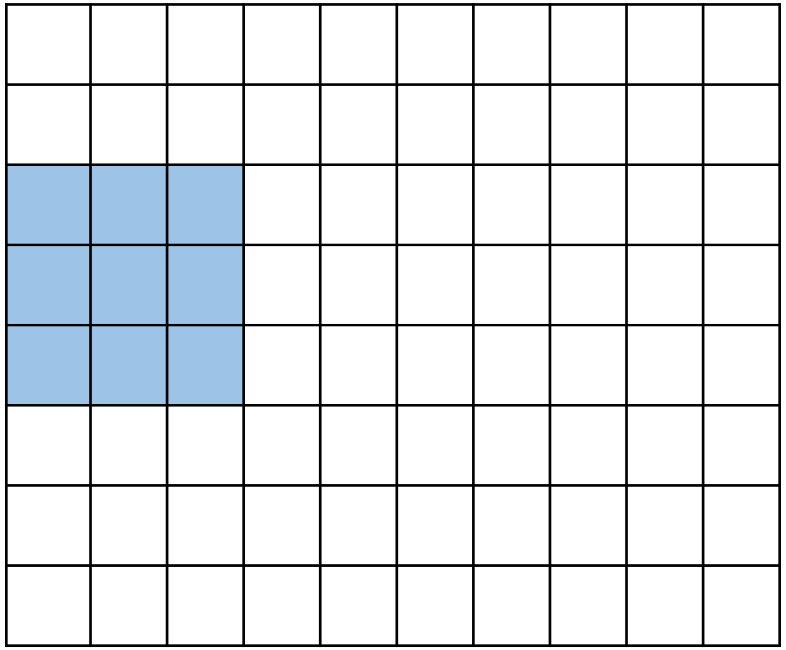
Stride



stride = 2

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1
- Can define a different stride
  - Hyperparameter

Stride

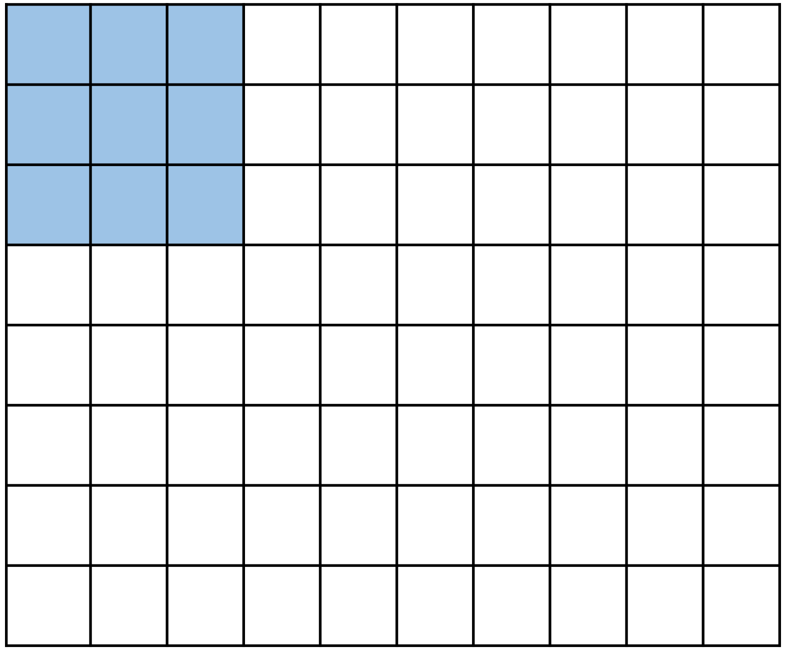


stride = 2

Works in both directions!

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1
- Can define a different stride
  - Hyperparameter

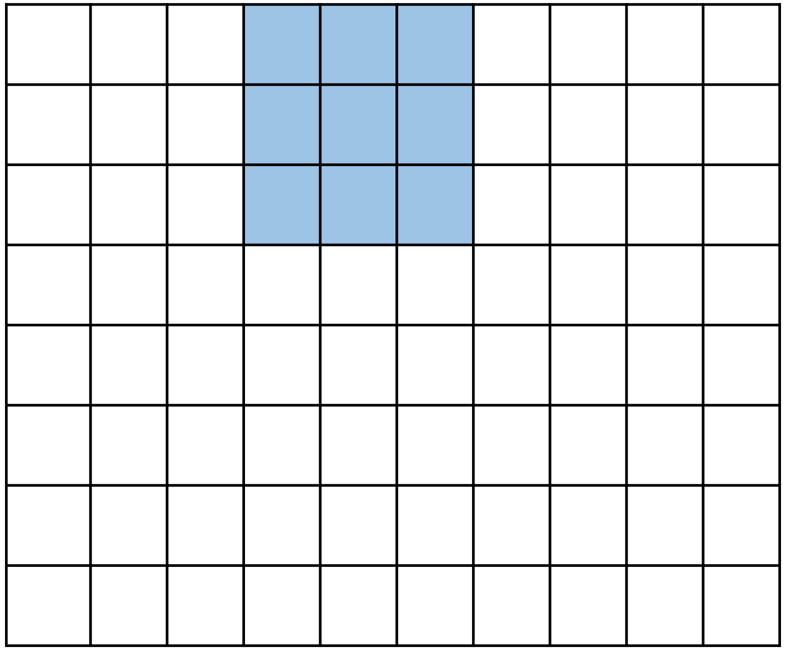
Stride



stride = 3

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1
- Can define a different stride
  - Hyperparameter
- Stride reduces the number of multiplications
  - Subsamples the image

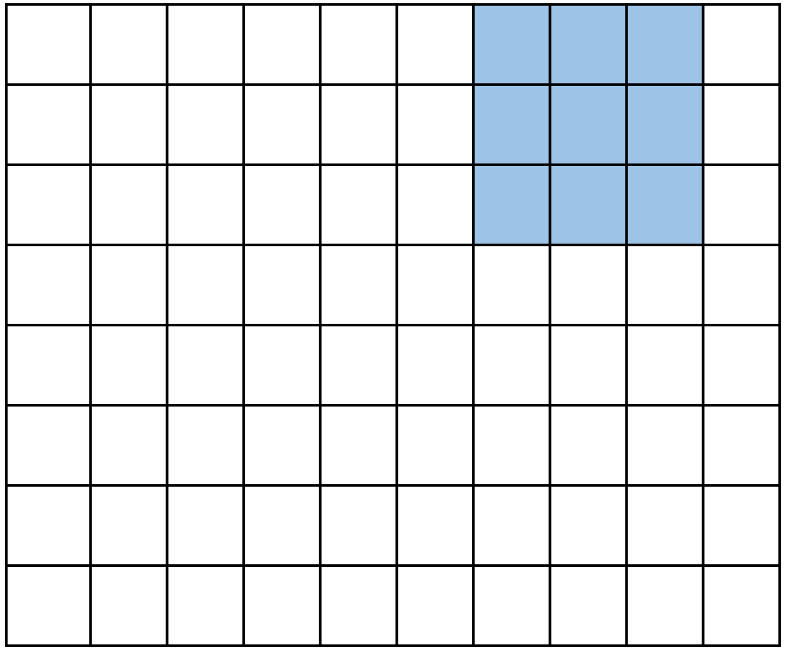
Stride



stride = 3

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1
- Can define a different stride
  - Hyperparameter
- Stride reduces the number of multiplications
  - Subsamples the image

Stride

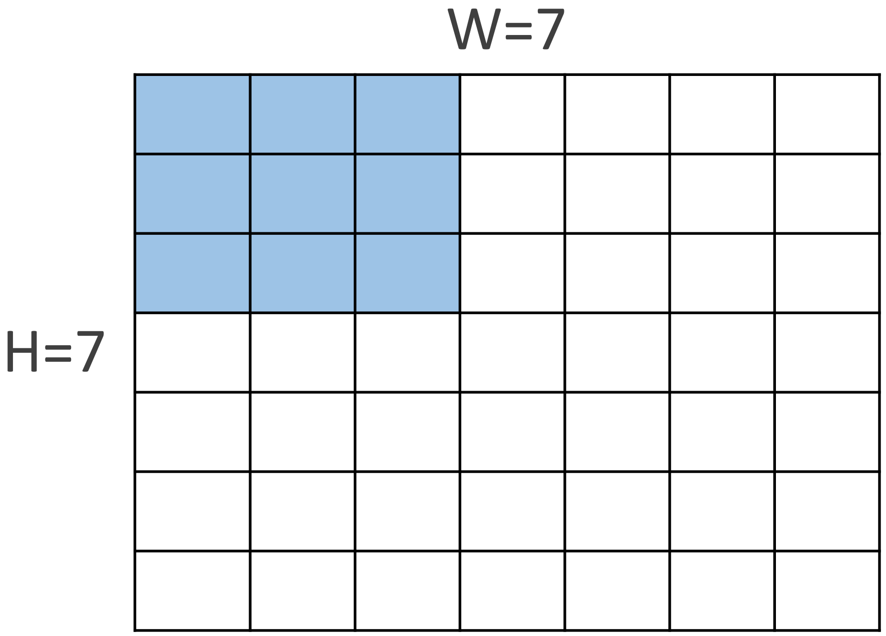


stride = 3

- Basic convolution slides the filter on the image one pixel at a time
  - Stride = 1
- Can define a different stride
  - Hyperparameter
- Stride reduces the number of multiplications
  - Subsamples the image

Activation Map Size

What is the size of the image after application of a filter with a given size and stride?



Take a 3x3 filter with stride 1

K=3, S=1

Output image is: 5x5

Zero Padding

Add columns and rows of zeros to the border of the image

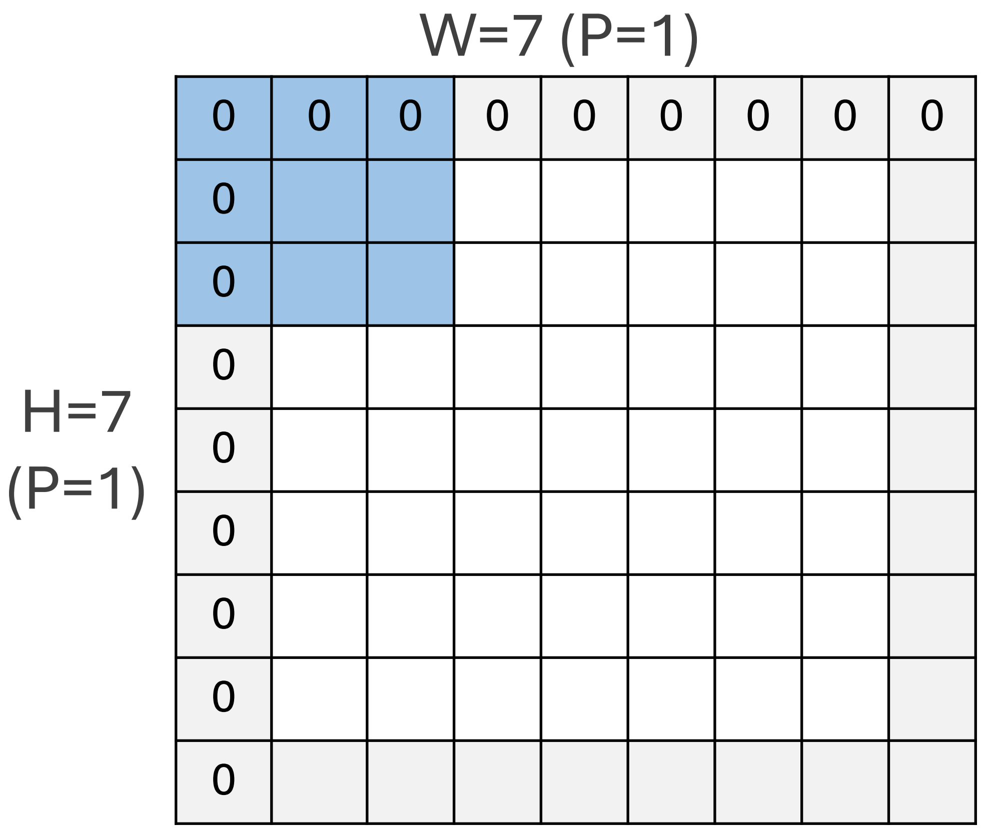
W=7

H=7

0	0	0	0	0	0	0	0	0
0								
0								
0								
0								
0								
0								
0								
0								

Zero Padding

Add columns and rows of zeros to the border of the image

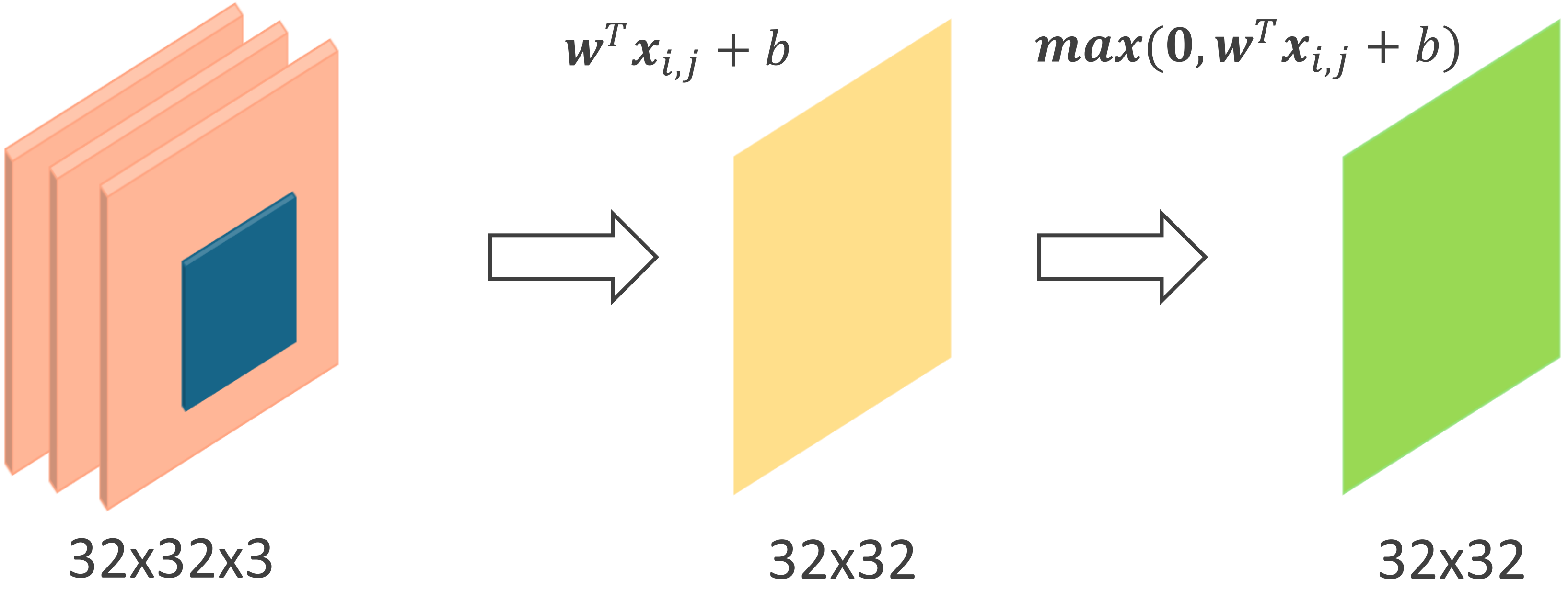


Zero padding serves to retain the original size of image

$$P = \frac{K-1}{2}$$

Pad as necessary to perform convolutions with a given stride S

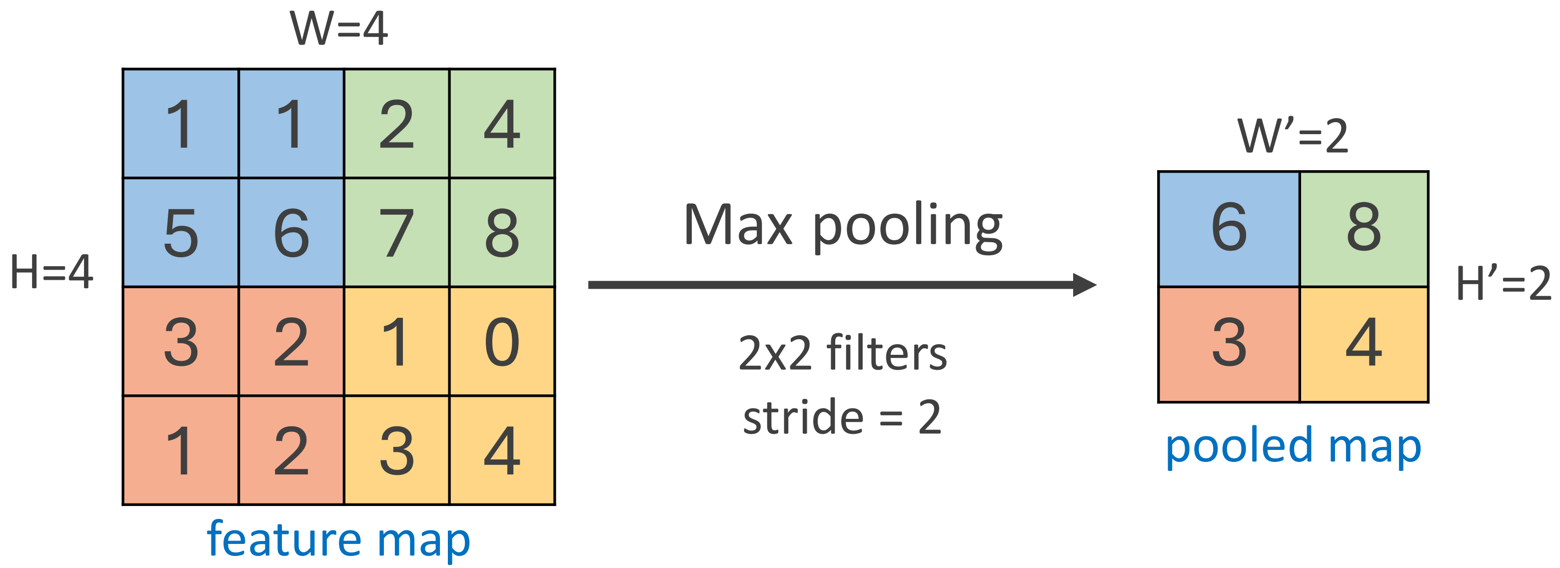
Feature Map Transformation



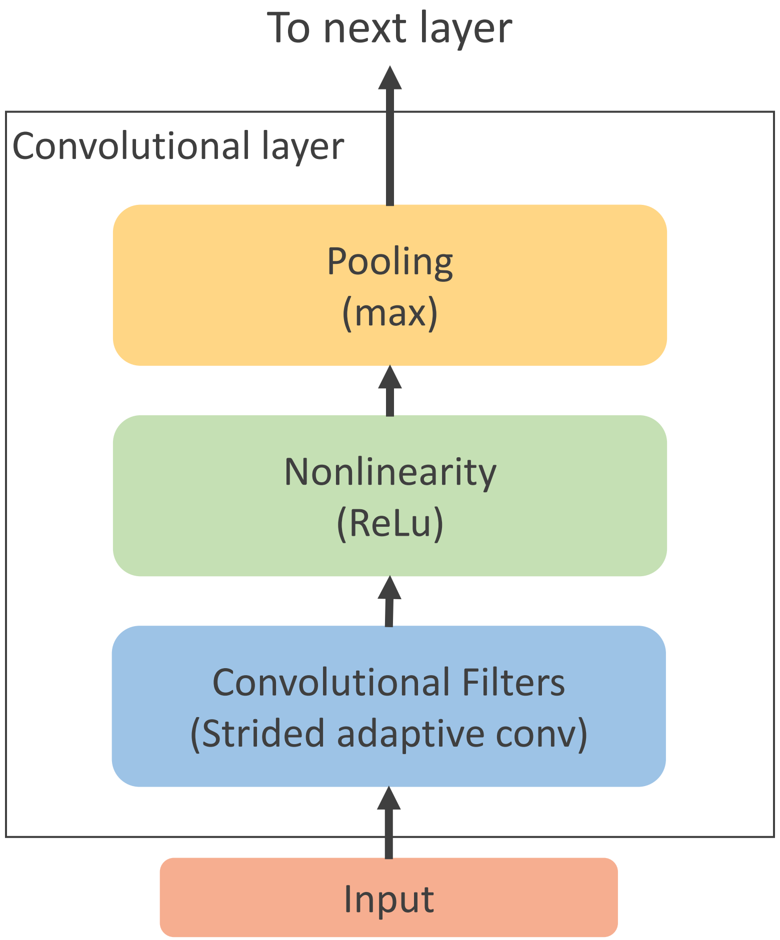
- Convolution is a linear operator
- Apply an element-wise nonlinearity to obtain a transformed feature map

Pooling

- Operates on the feature map to make the representation
  - Smaller (subsampling)
  - Robust to (some) transformations
- Max pooling is the one used more frequently, but other forms are possible



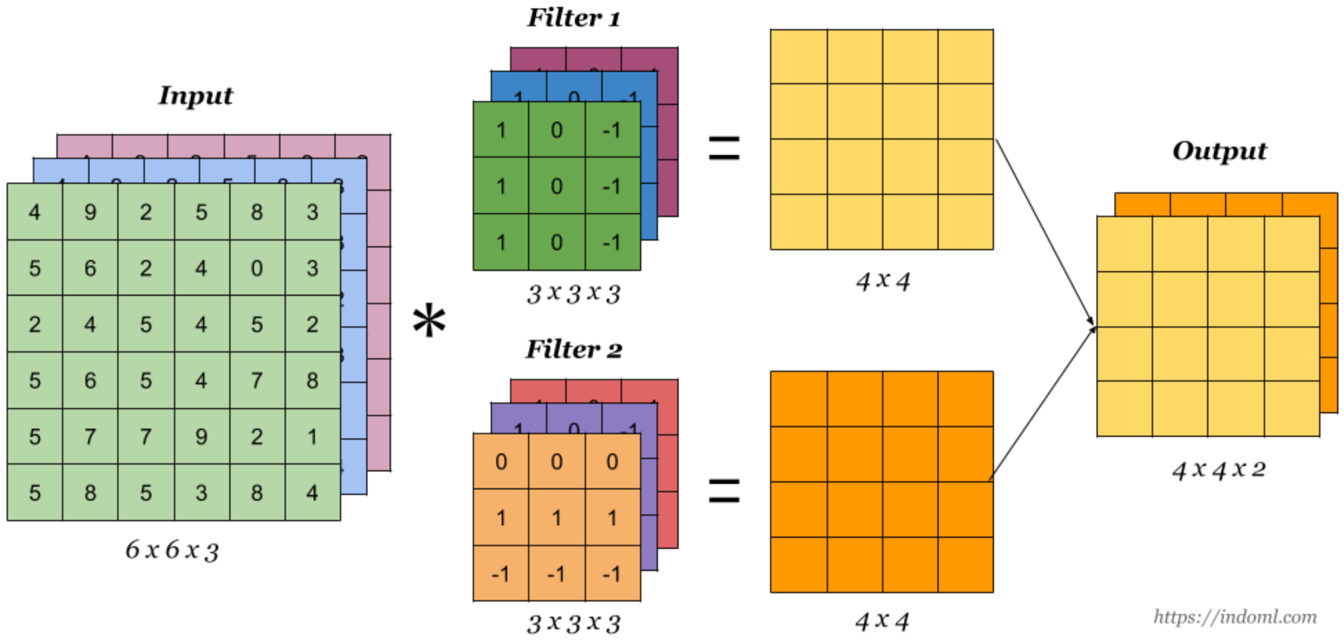
Combining the single ingredients



- An architecture made by a hierarchical composition of the basic elements
- Pooling is typically not applied after every convolution
- Network parameters are in the convolutional component

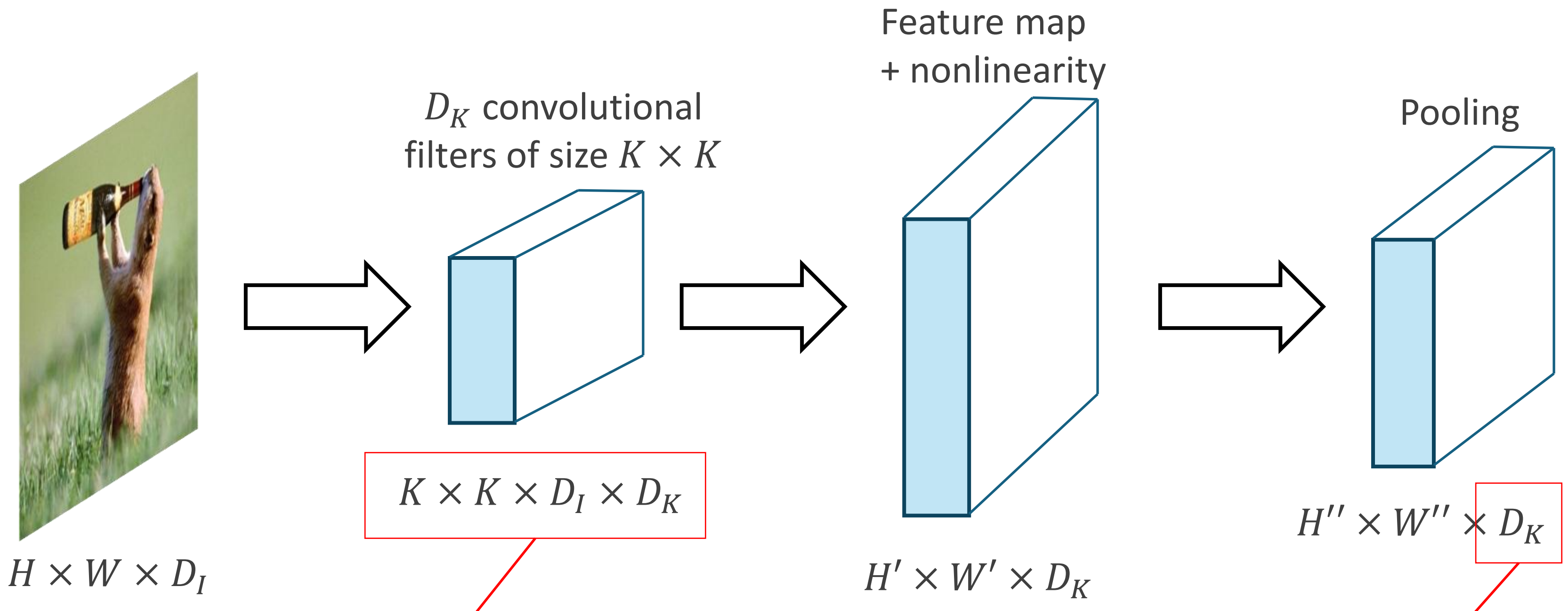
Combining Multiple Filters

We typically work with more than one convolutional filter



- Convolutions across multiple channels are summed
- Feature maps from multiple convolutions are concatenated

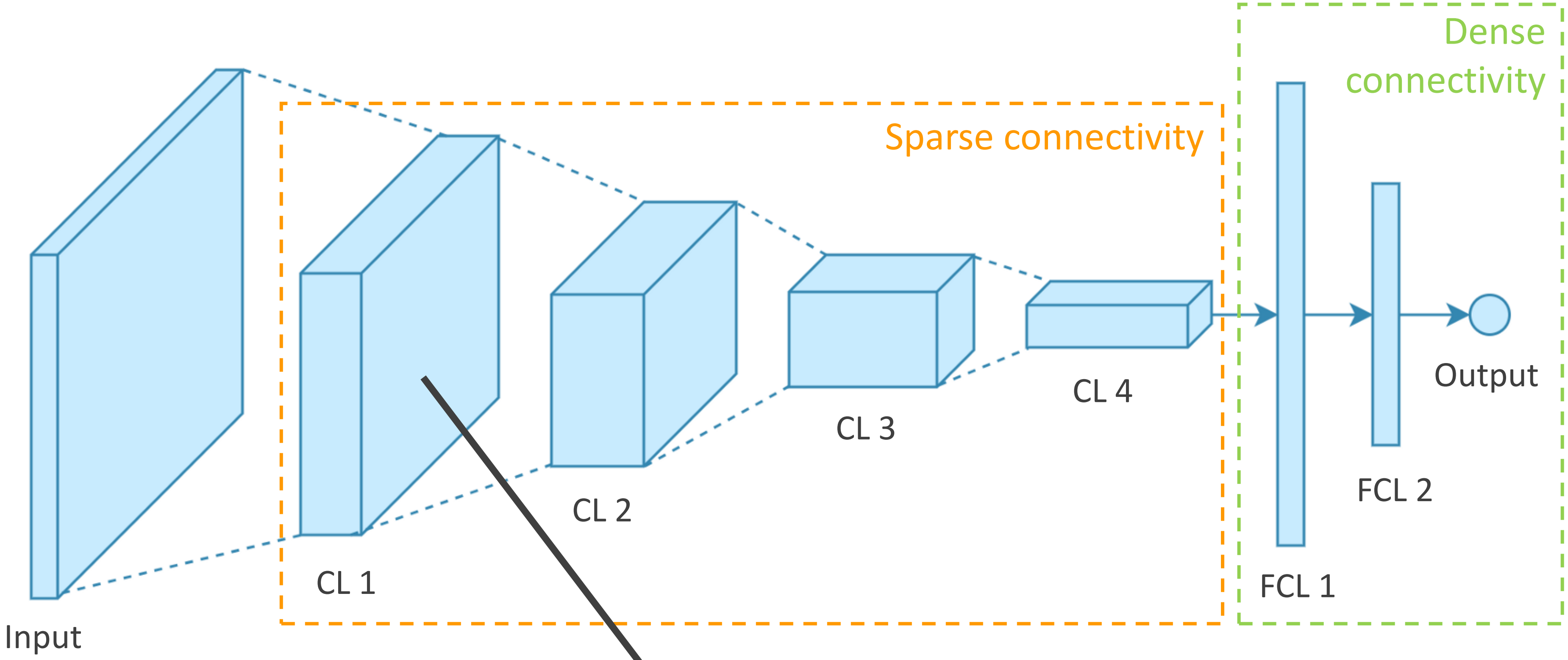
Convolutional Filter Banks



Number of model parameters due to this convolution element (add D_K bias terms)

Pooling is often (not always) applied independently on the D_K convolutions

A (Prototypical) Convolutional Architecture

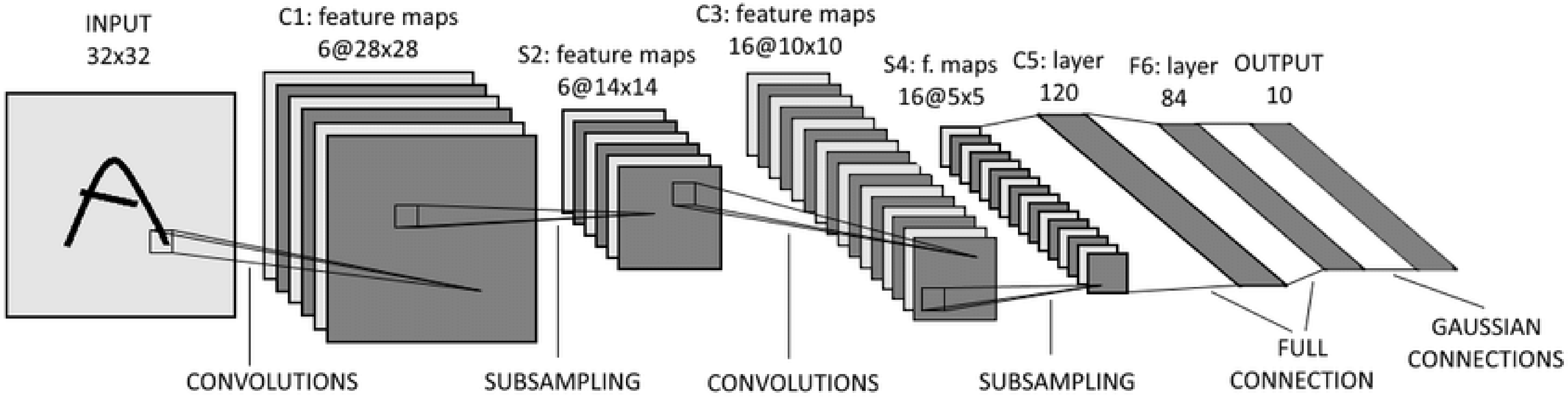


CL -> Convolutional Layer
FCL -> Fully Connected Layer

Contains several convolutional filters with different size and stride

Architectures and useful tricks

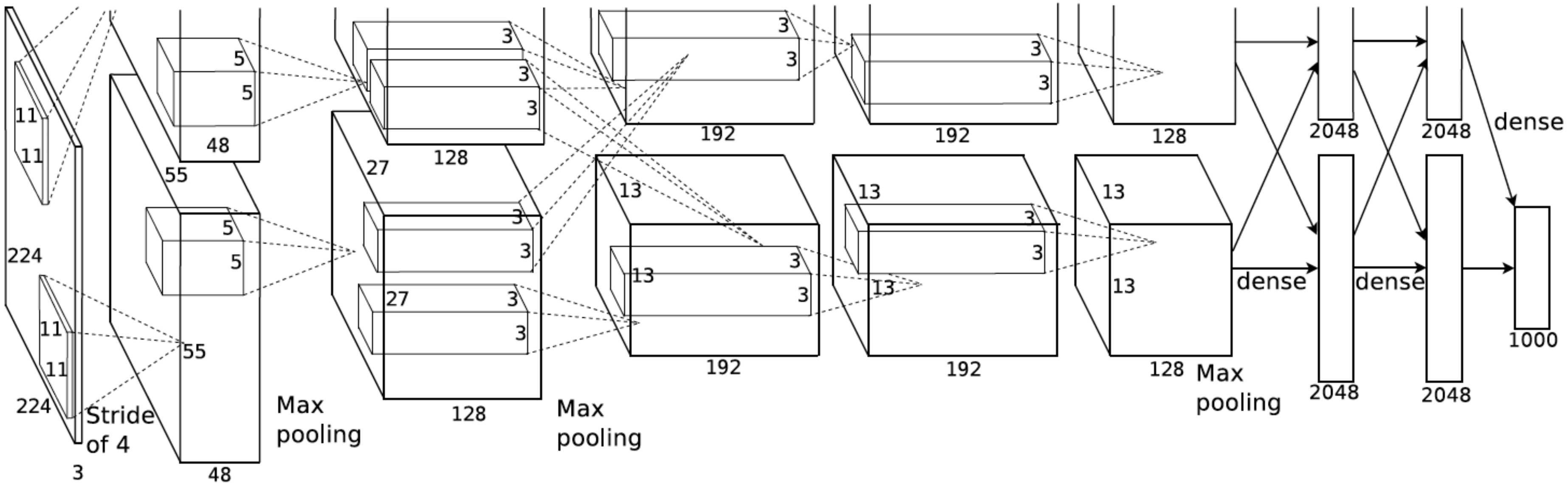
LeNet-5 (1989)



The first CNN on images

- Grayscale images
- Filters are 5x5 with stride 1 (sigmoid nonlinearity)
- Pooling is 2x2 with stride 2

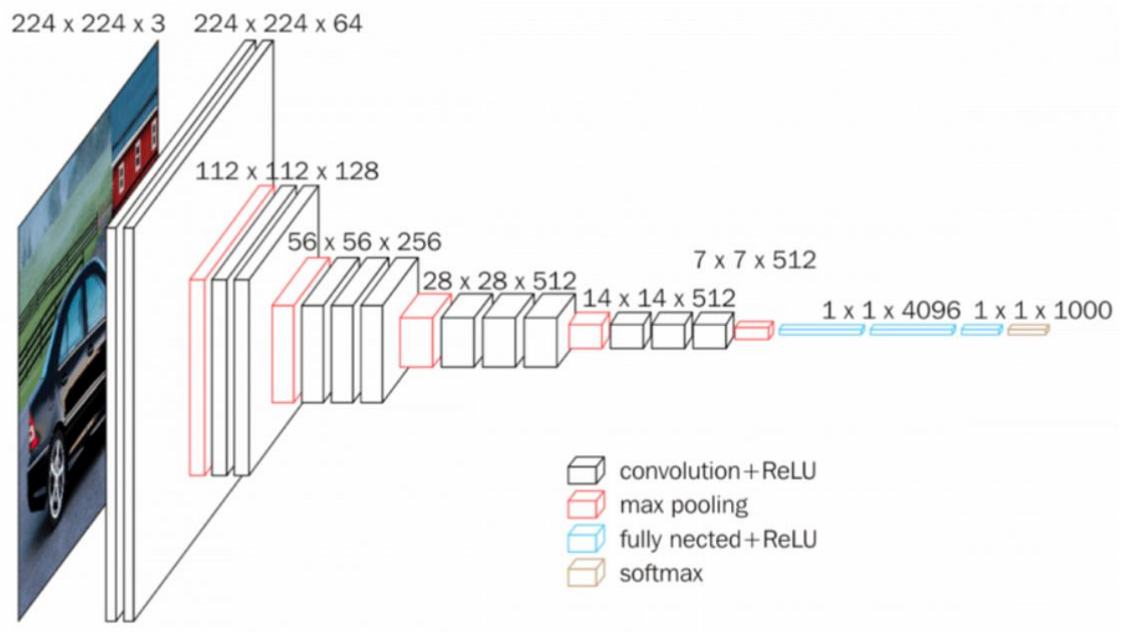
AlexNet (2012) - Architecture



Began the deep learning revolution

- RGB images 227x227x3
- 5 convolutional layers + 3 fully connected layers

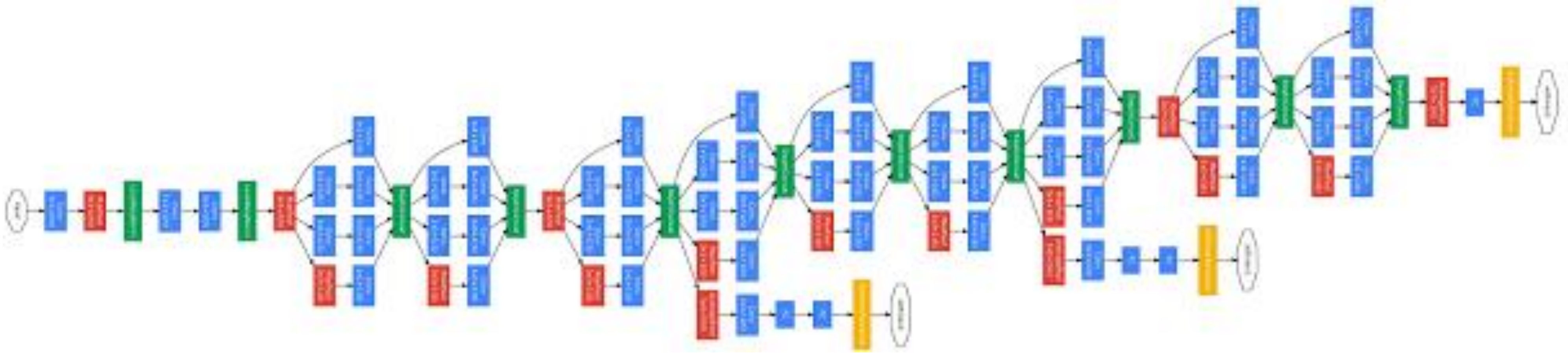
VGG-16 (2014)



Scaled computer vision tasks to thousands of tasks

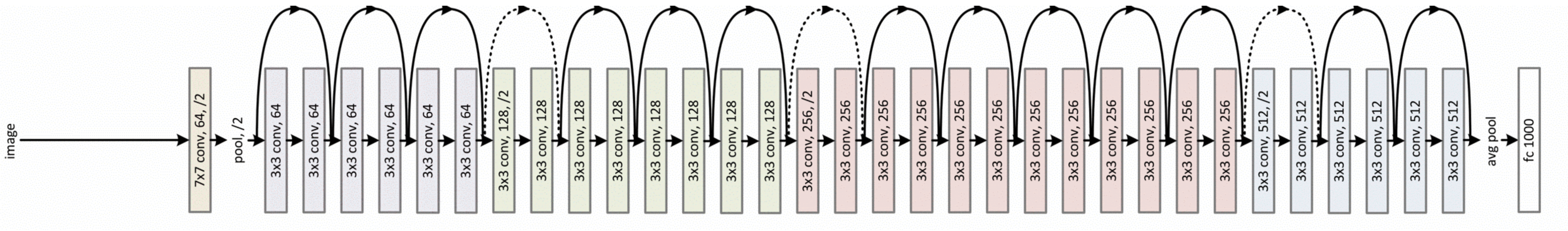
Parameter heavy and inefficient, but still somehow in use

GoogLeNet/InceptionNet (2015)



- Introduced rich multi-resolution kernels and heavy use of 1x1 convolutions
- A first parameter-efficient CNN

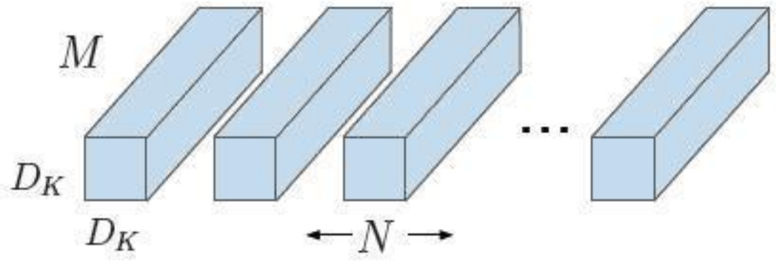
ResNet (2015)



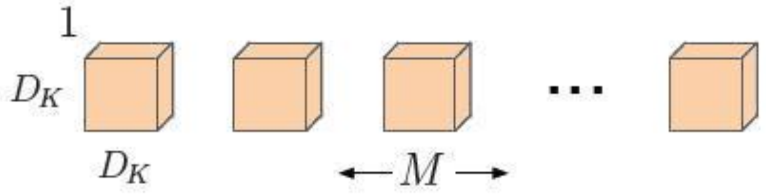
- Introduced residual connections
- Began the era of very-deep networks

MobileNets

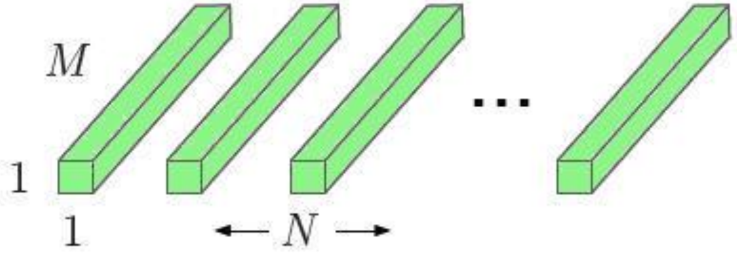
Making CNNs efficient to run on mobile devices by depth-wise separable convolutions



(a) Standard Convolution Filters



(b) Depthwise Convolutional Filters



(c) 1×1 Convolutional Filters called Pointwise Convolution in the context of Depthwise Separable Convolution

arxiv.org/pdf/1704.04861.pdf

CNN Training

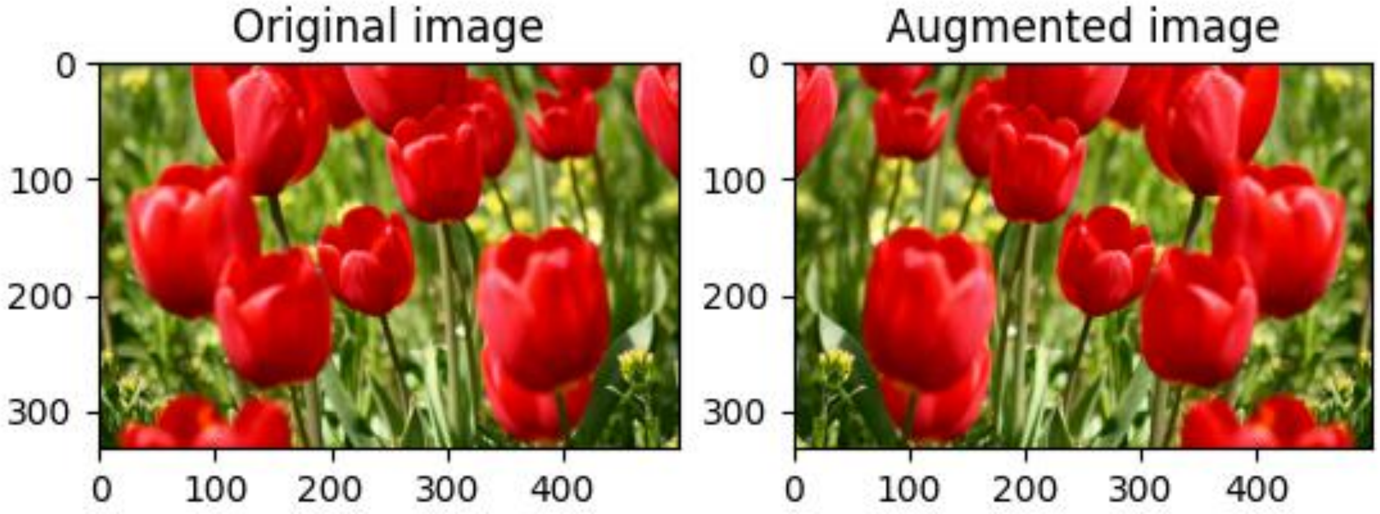
As usual by backpropagation and most recommendations for DNNs also apply to CNNs

- **Many hyperparameters** → follow proven best practices
- **Regularization is key** → CNNs overfit less than MLPs, but still need dropout, weight decay, batch/layer normalization, etc.
- **Use ReLU or similar activations** → avoid sigmoid/tanh
- **Optimizer & learning rate** → in model selection!
- **Deep networks?** Watch out for **vanishing gradients** (aka check the gradient norms)

Bonus (for images): You can use **data augmentation** to massively expand your training set

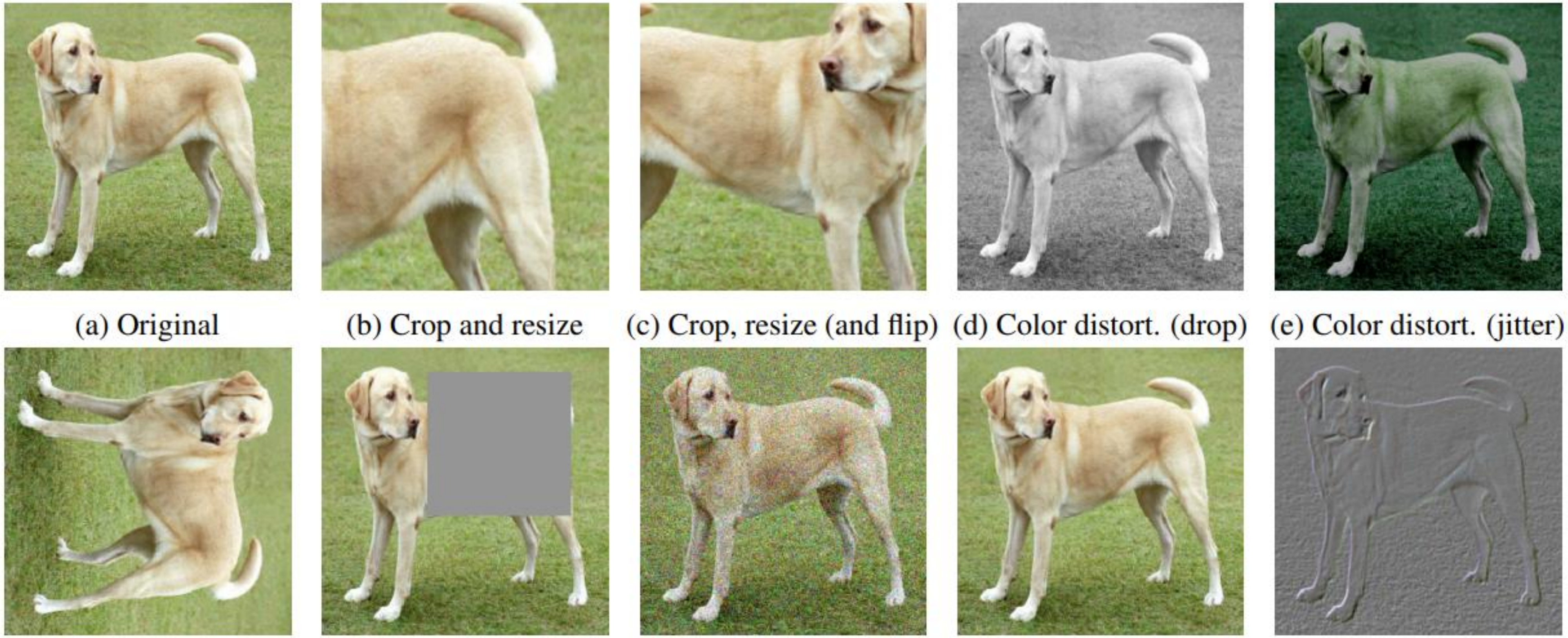
Data Augmentation

Key intuition - If I have an image with a given label, I can transform it (by flipping, rotation, etc) and the resulting image will still have the same label



Data augmentation on general purpose images

Need to exercise care with biomedical images!

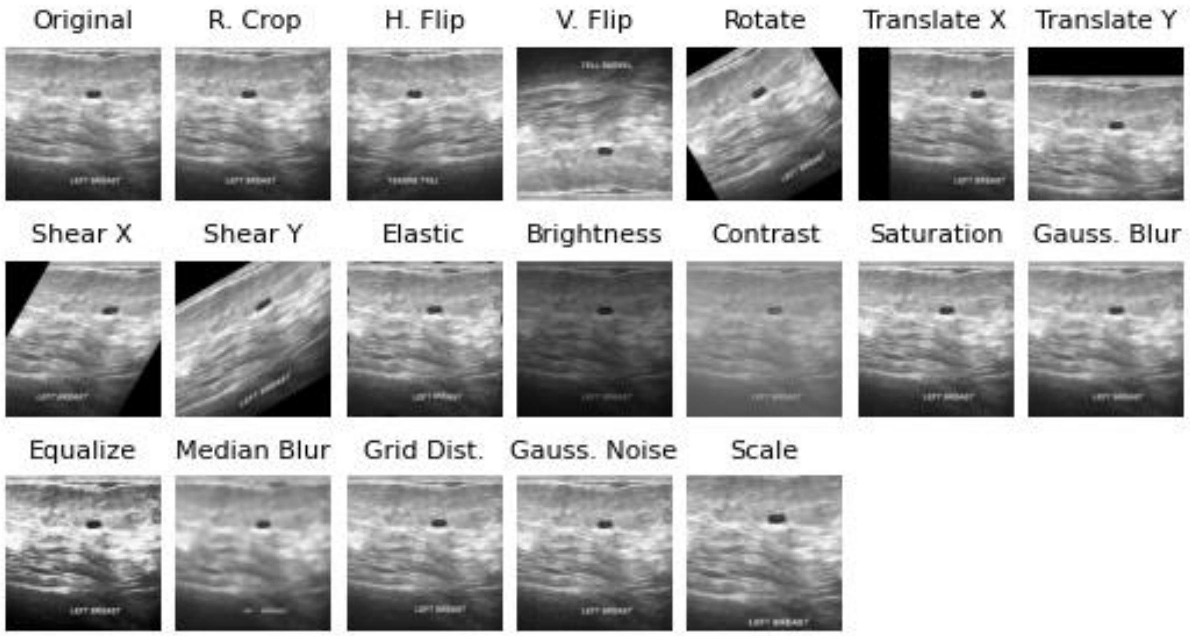


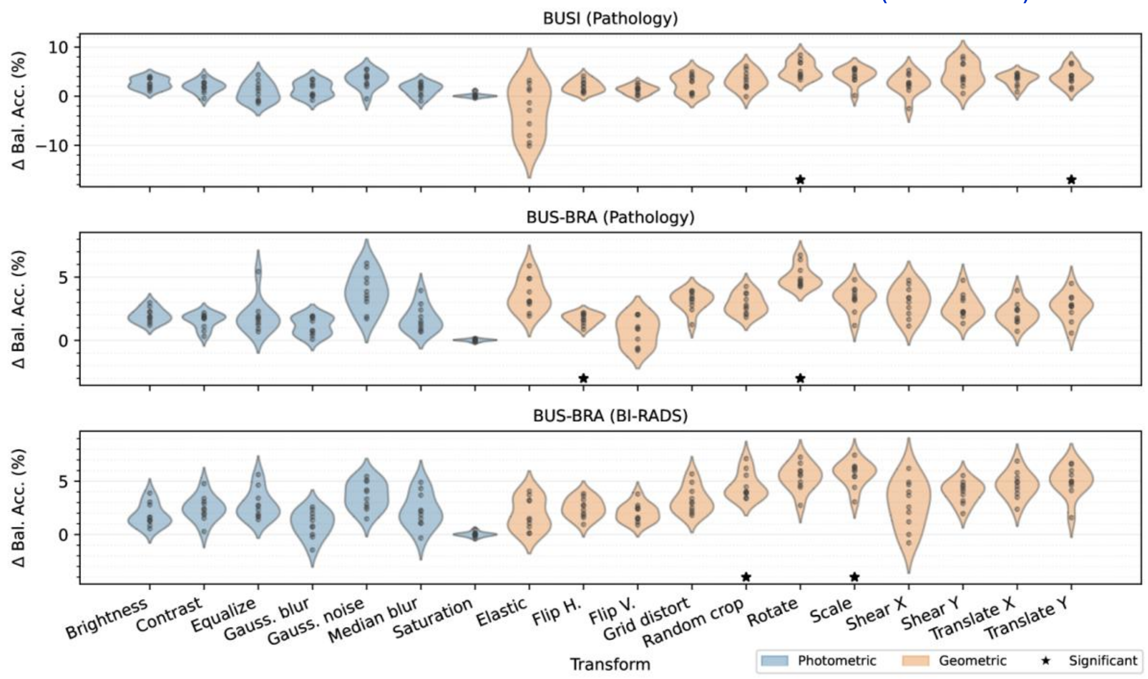
(a) Original (b) Crop and resize (c) Crop, resize (and flip) (d) Color distort. (drop) (e) Color distort. (jitter)

(f) Rotate $\{90°, 180°, 270°\}$ (g) Cutout (h) Gaussian noise (i) Gaussian blur (j) Sobel filtering

Data augmentation for biomedical images

Source (with code) [here]





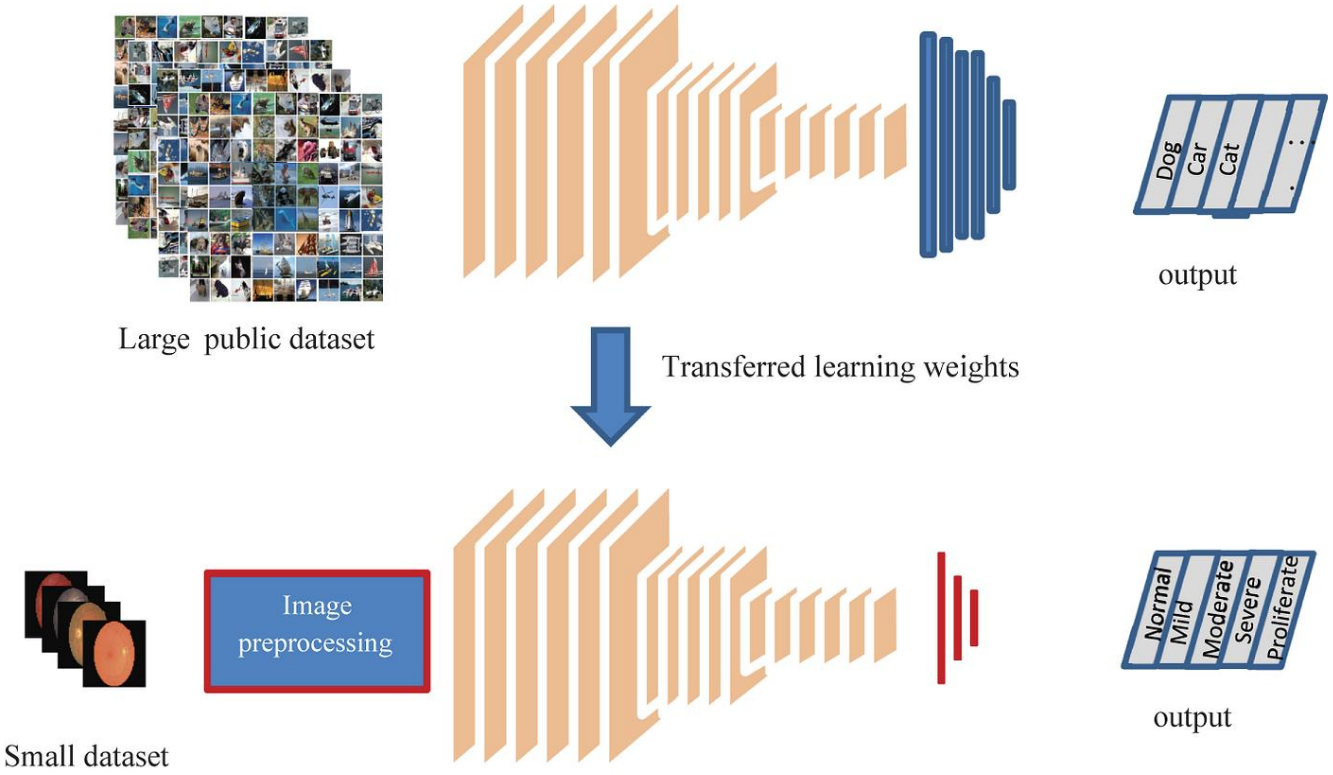
Need to personalize the augmentation technique to the imaging technology and to carefully assess the impact

Pros & cons of augmentation techniques in medical imaging

Source here

Approach	Pros	Cons
Augmentations with geometric transformations	Augmentations with geometric transformations (e.g., scaling, rotation, shearing, and flipping) are effective in increasing the amount and the diversity of the datasets by adding modified versions of existing images. Because the generated new images are identifiable and resembling	The new images may not represent enough variations in the location, shape and pathology of a lesion since the contents of the real and the newly produced images are similar
Scaling, rotation	Since lesions can appear at different positions and scales, adding rotated and scaled images into the training sets help the models in learning	
Scaling	Image augmentation using scaling can produce realistic images since lesions can vary in size. Also, the usage of augmented images obtained by scaling allows network models to learn properties without depending on the original scale	
Translation	Augmentation using translation provides the prevention of positional bias. In other words, training a network model with augmented images using translation enables the network to learn geographically invariant features. The network model do not focus on properties in a single spatial location	
Flipping	They can preserve the labels and features of the reconstructed images. The label-preserving techniques are helpful in classification tasks. For example, after the translation of an image showing HGG in a classification problem, the image remains intact	Label-preserving property may not always be provided since profile information can be on the left or right side of the images. For example, breast profiles are mostly on the leftward of mammography images in some datasets. Therefore, flipping may not conserve the labels for the images in those datasets
Modification of contrast or brightness, intensity normalization, histogram equalization, noise addition and blurring, and sharpening	Medical images are acquired with various imaging modalities and technologies, and they can be diversified in pixel intensities. Therefore, generating new images by changing the intensity values of existing images (e.g., modification of contrast or brightness, intensity normalization, histogram equalization, noise addition and blurring, and sharpening) can help to augment medical images and provide an improvement in the diversity of datasets	Augmentation by changing the intensity values, particularly noise addition and blurring techniques may disturb the quality of the images (i.e., disintegrate images' features) and change the lesions' characteristics, causing the generation of unnatural images and low accuracy
Random erasing, random cropping	Augmentations with random erasing or cropping are to improve the robustness of a network. Also, they are applied without depending on any parameter	Augmentations with random erasing, cropping, and color space transformations may lead to the loss of significant information
Color space transformations	Through color space transformations, biases in an image due to various illumination can be prevented	

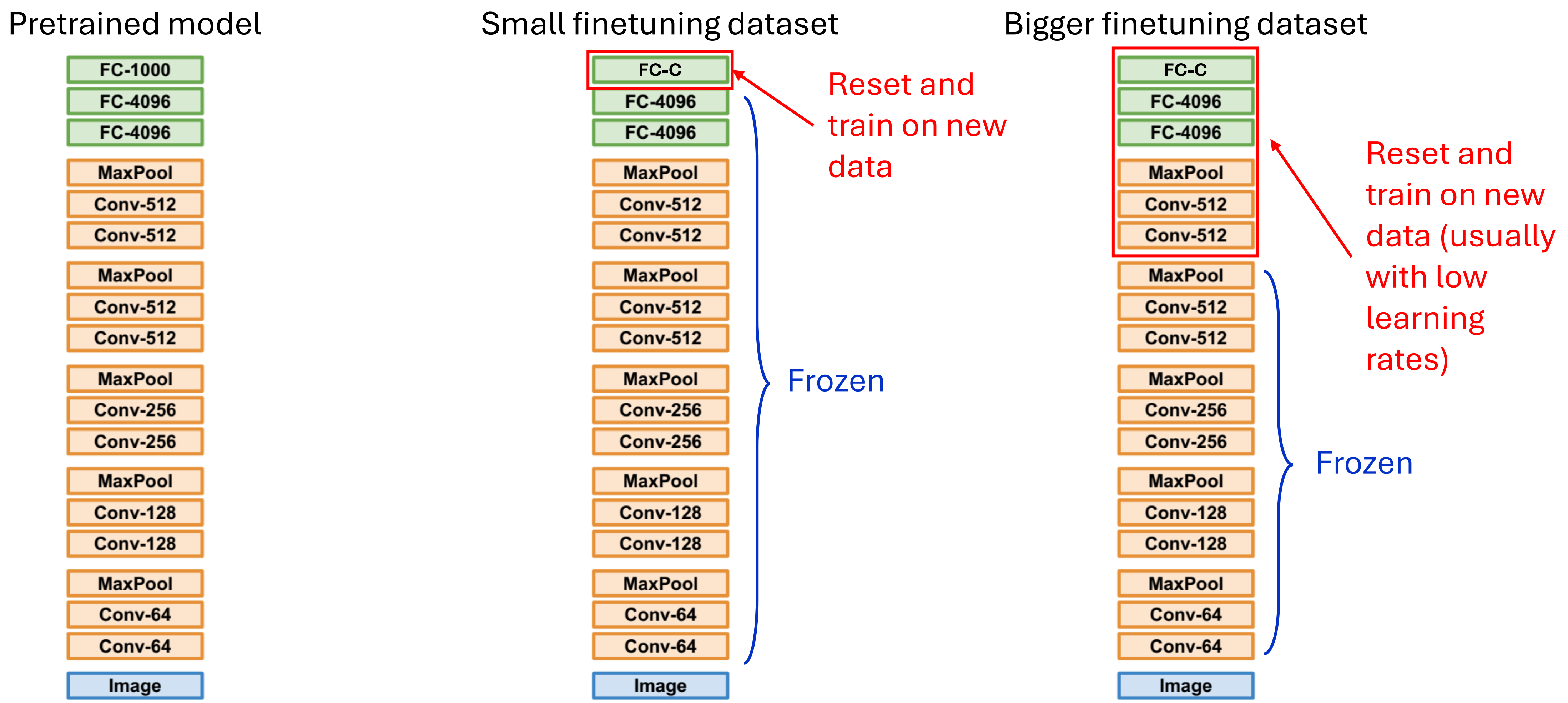
A comprehensive catalogue of augmentation techniques for medical images



Use (part of) a model trained (pretrained) by someone on large dataset as a "feature-extractor" on problems with fewer data, fine tuning only the predictor part

Transfer learning

Transfer Learning – Rule of Thumb



Biomedical Image Processing Tasks

Fundamental tasks in medical imaging

Classification

- Detect presence/absence of disease (e.g., tumor vs. normal)
- Multi-class scenarios (e.g., different tumor types)

Regression

- Predict continuous outcomes (e.g., tumor volume, disease progression)
- Often used in quantitative imaging biomarkers

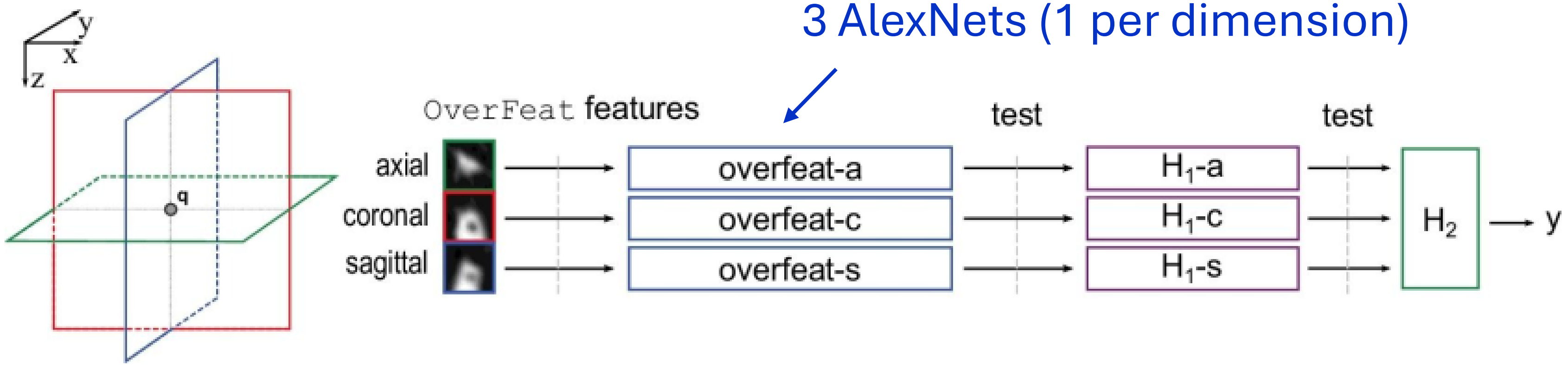
Segmentation & Detection

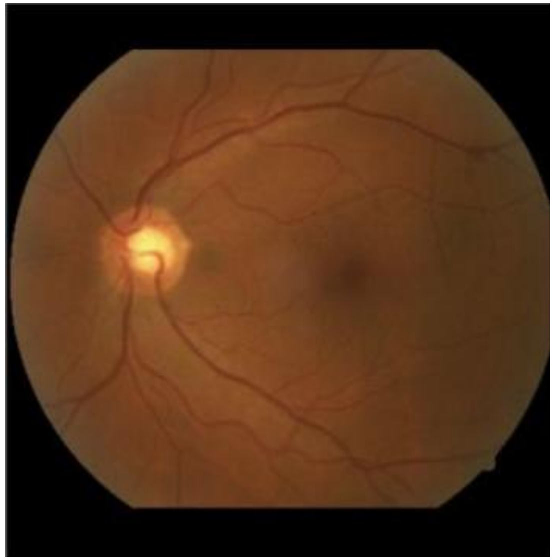
- Delineate structures (tumours, organs) at pixel/voxel-level or detect each instance of a category item in an image (e.g. cell counting)
- Essential for measuring size, shape, location, and number
- Forms basis for surgical or radiotherapy planning

Biomedical Image Classification – Early Days

3D CT scans for nodule classification (Ciompi et al, Med. Img. Analysis 2015)

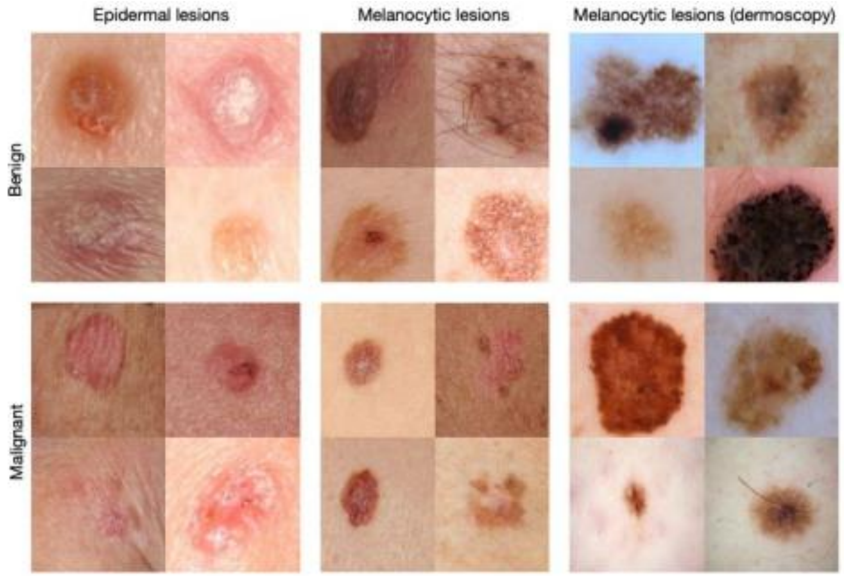
3 AlexNets (1 per dimension)





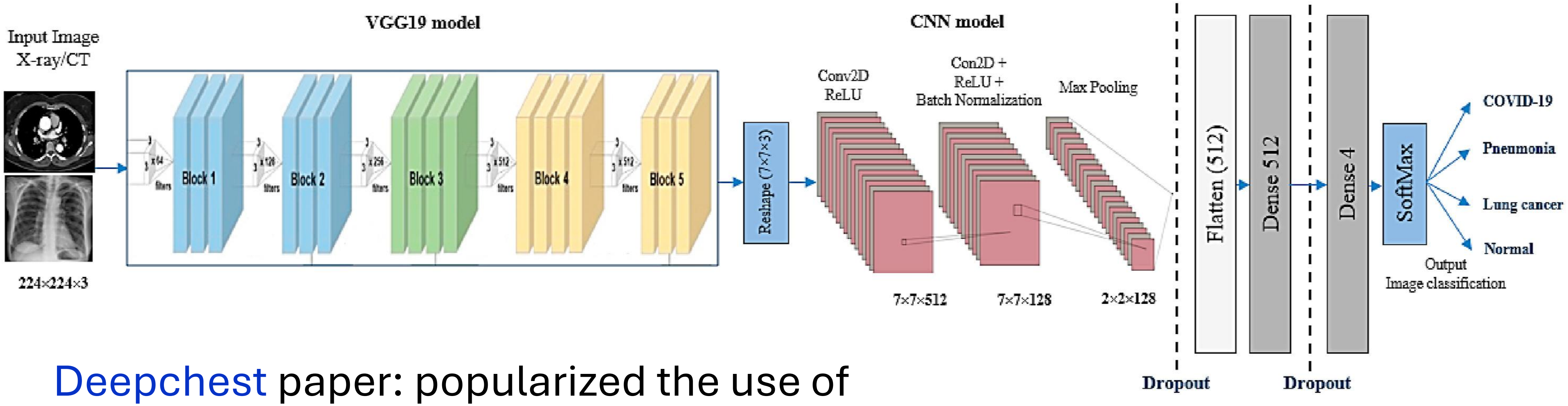
Retinal fundus binary classification starting from pretrained GoogLeNet (Gulshan et al, JAMA 2016)

Skin cancer grading in 757 classes by GoogLeNet (Esteva et al, Nature 2017)

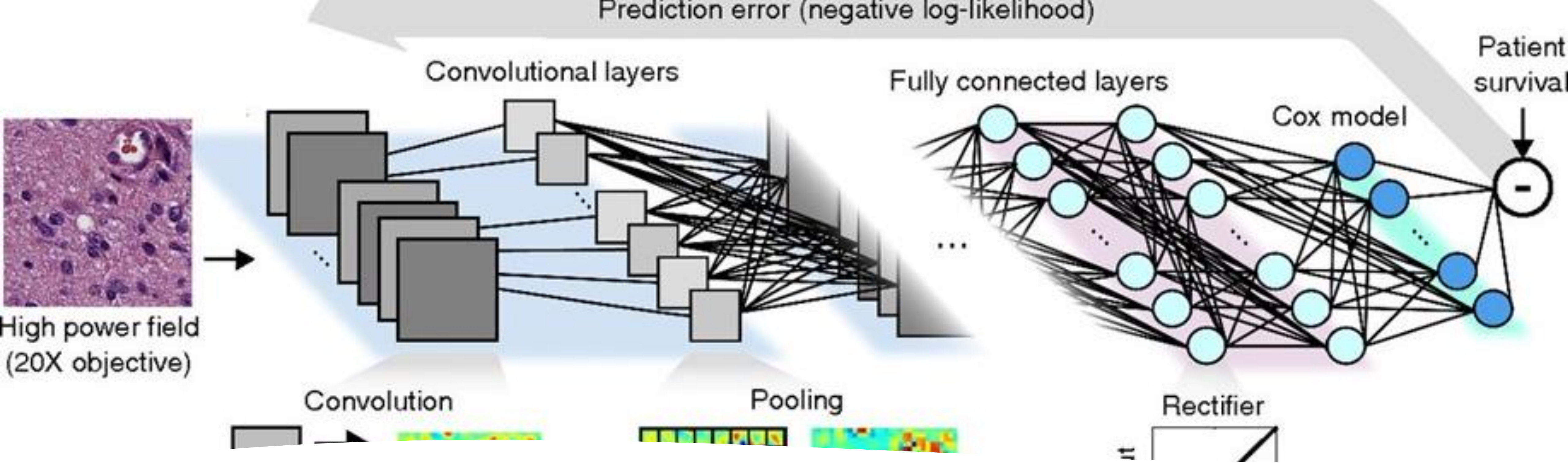


COVID and the CNN frenzy

Ibrahim et al, Computers in biology and medicine, 2021



Deepchest paper: popularized the use of CNNs in lung imaging on multi-disease diagnosis



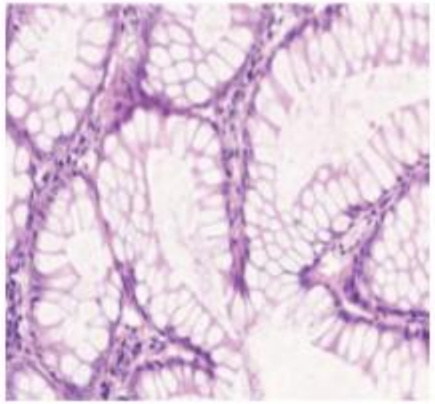
CNNs for pathology images

CNN integrated with a Cox proportional hazards model

Mobadersany et al. PNAS 2018

Advanced Tasks: Segmentation and Detection

Classification & Regression

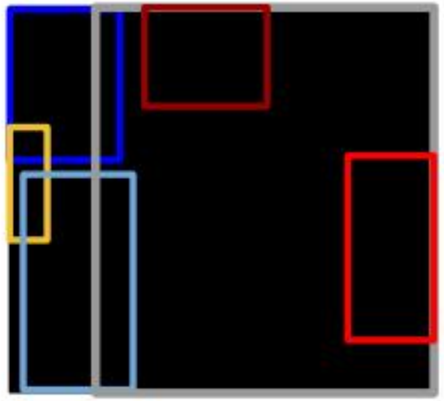


Output: one prediction per image

Semantic Segmentation

Output: one class label per pixel

Detection



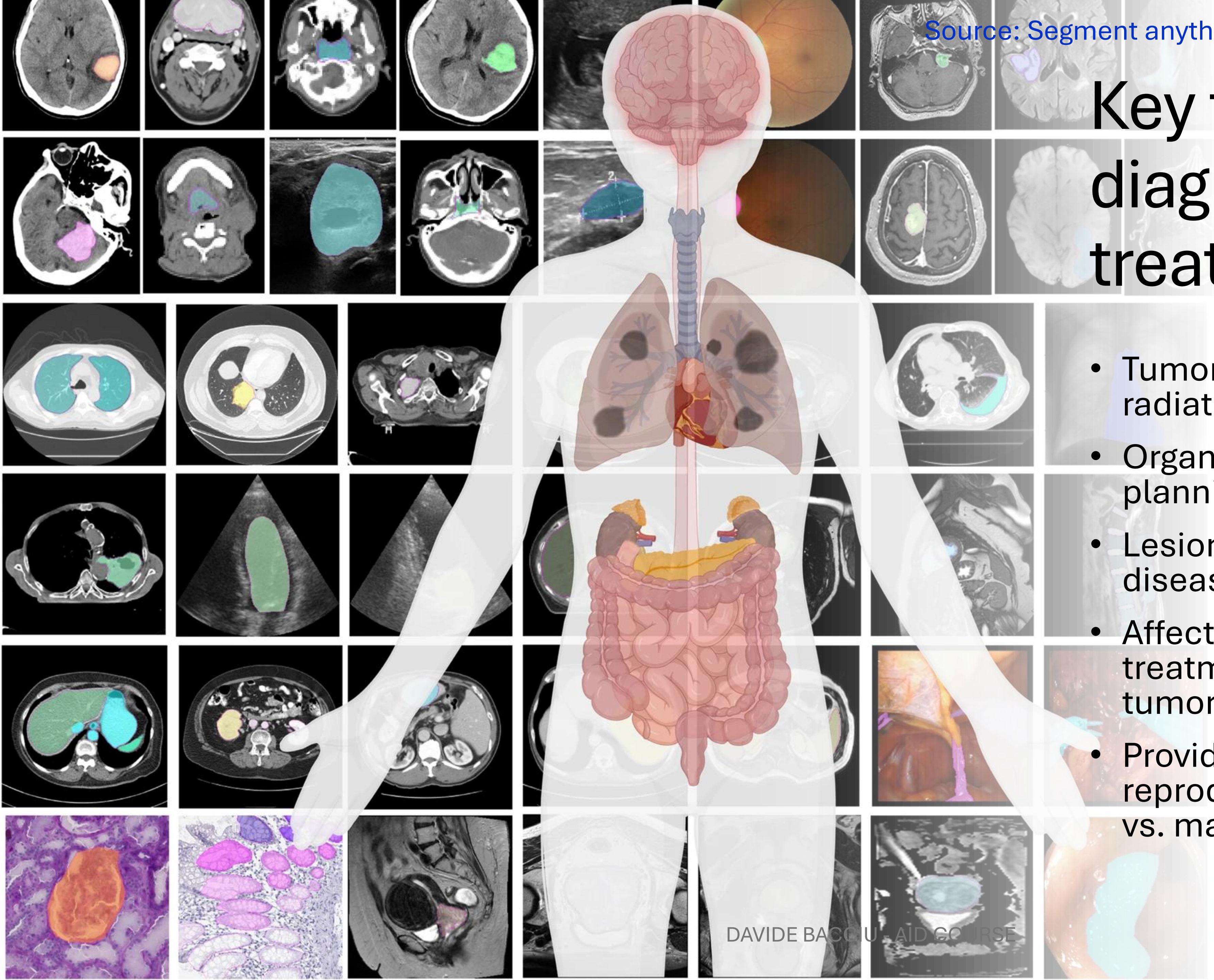
Output: one bounding around each instance of an object category of interest

Instance Segmentation

Output: category and instance label for each pixel

Need different CNN architectures!

Source: [Arxiv]

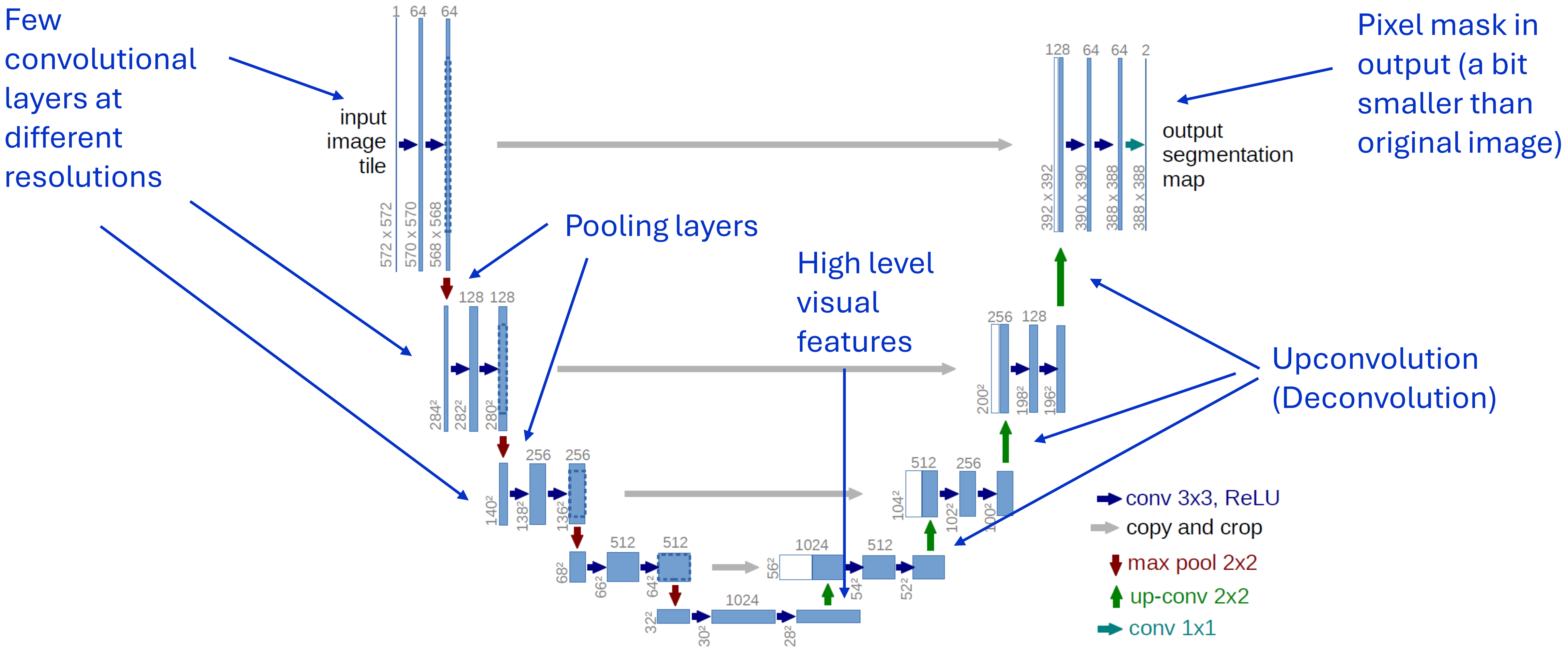


Source: Segment anything in medical images, Nature, 2024

Key tasks in diagnostics and treatment

- Tumor boundary detection for radiation therapy
- Organ delineation for surgical planning
- Lesion quantification for disease progression
- Affects prognosis and treatment strategies (e.g., tumor growth rates)
- Provides consistent, reproducible measurements vs. manual outlining

U-Net: CNN for Biomedical Image Segmentation

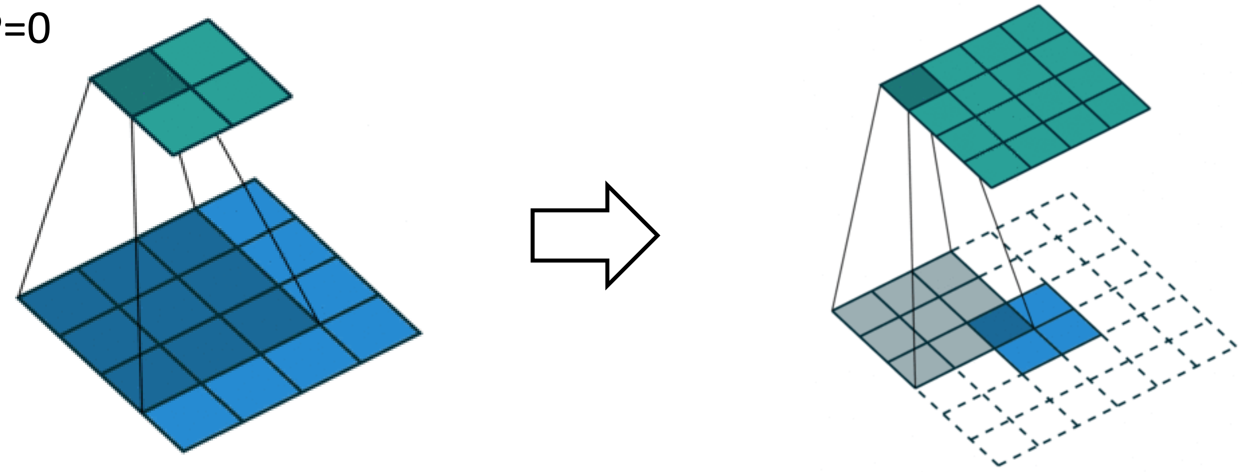


Deconvolution (Transpose Convolution)

Upsample a smaller input image (blue) into a larger output image (green)

K=3, S=1, P=0

Standard convolution

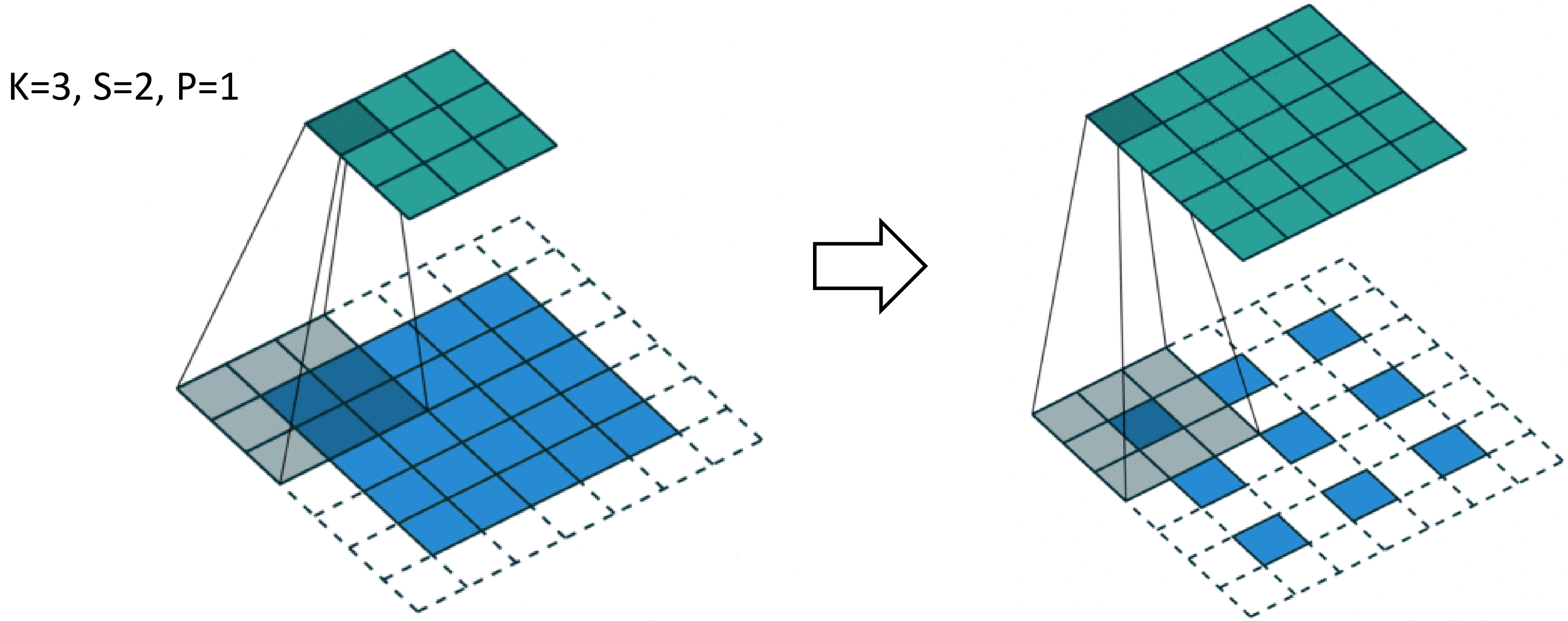


Transpose convolution

If you had no padding in the forward convolution, you need to pad much when performing transposed convolution

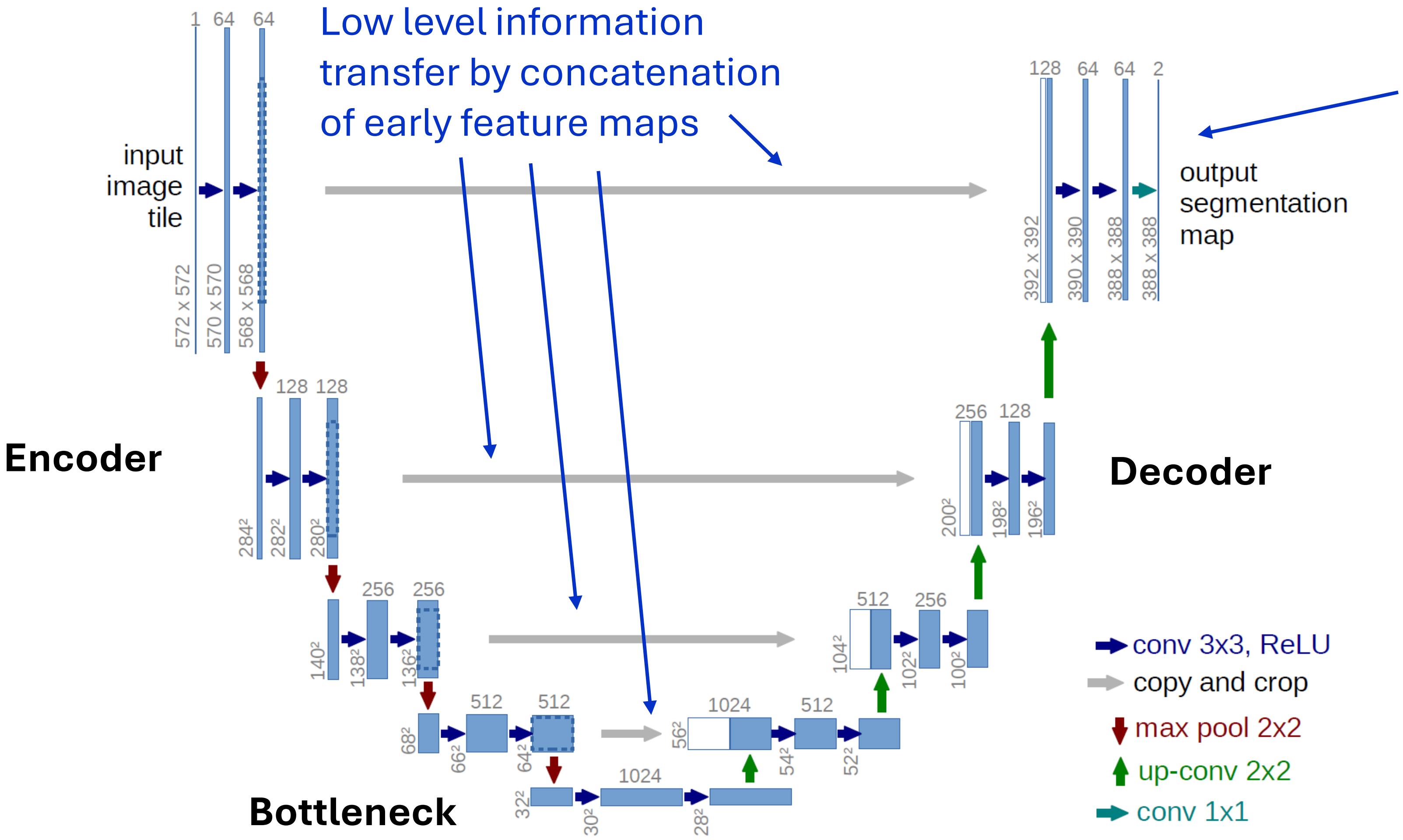
Deconvolution (Transpose Convolution)

If you have striding, you need to fill in the convolution map with zeroes to obtain a correctly sized deconvolution



https://github.com/vdumoulin/conv_arithmetic

U-Net: CNN for Biomedical Image Segmentation



Train with classification loss (e.g. binary cross entropy) on every pixel, sum over all pixels to get total loss

Key aspects of U-Nets

Handling Fewer Images

- U-Net can be trained effectively on relatively small datasets (typical in medical imaging)
- Use of heavy data augmentation is standard

Robust Feature Localization

- Skip connections preserve spatial information lost by pooling
- Helps differentiate fine boundaries (tumor edges, organ interfaces)

2D vs. 3D U-Net Variants

- 2D: Processes slices independently, good if GPU memory is limited
- 3D: Captures volumetric context but more memory-intensive

Assessing Image Segmentation

Intersection-over-Union (IoU) – a.k.a. the Jaccard Index

pixels included in both target and prediction maps

$$IoU = \frac{target \cap prediction}{target \ \cup prediction}$$

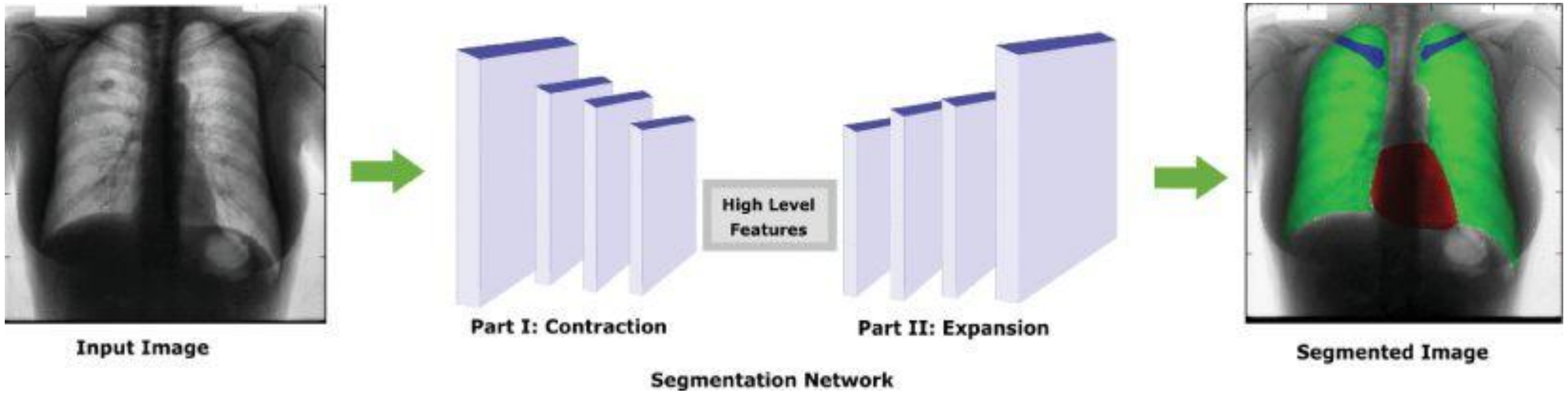
Total # pixels in the union of both masks

Can be computed at many levels of resolution

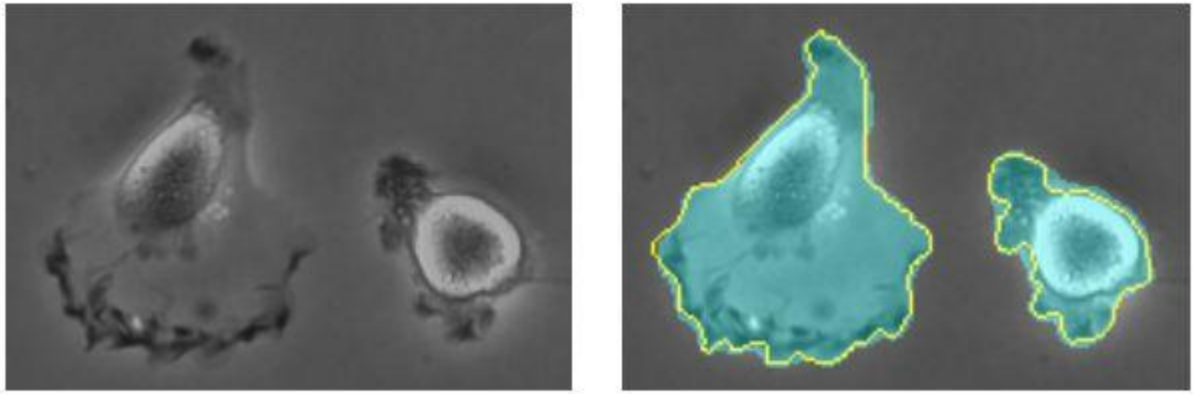
- Averaged over all masks
- On individual masks
- At image level

U-Net Example Application

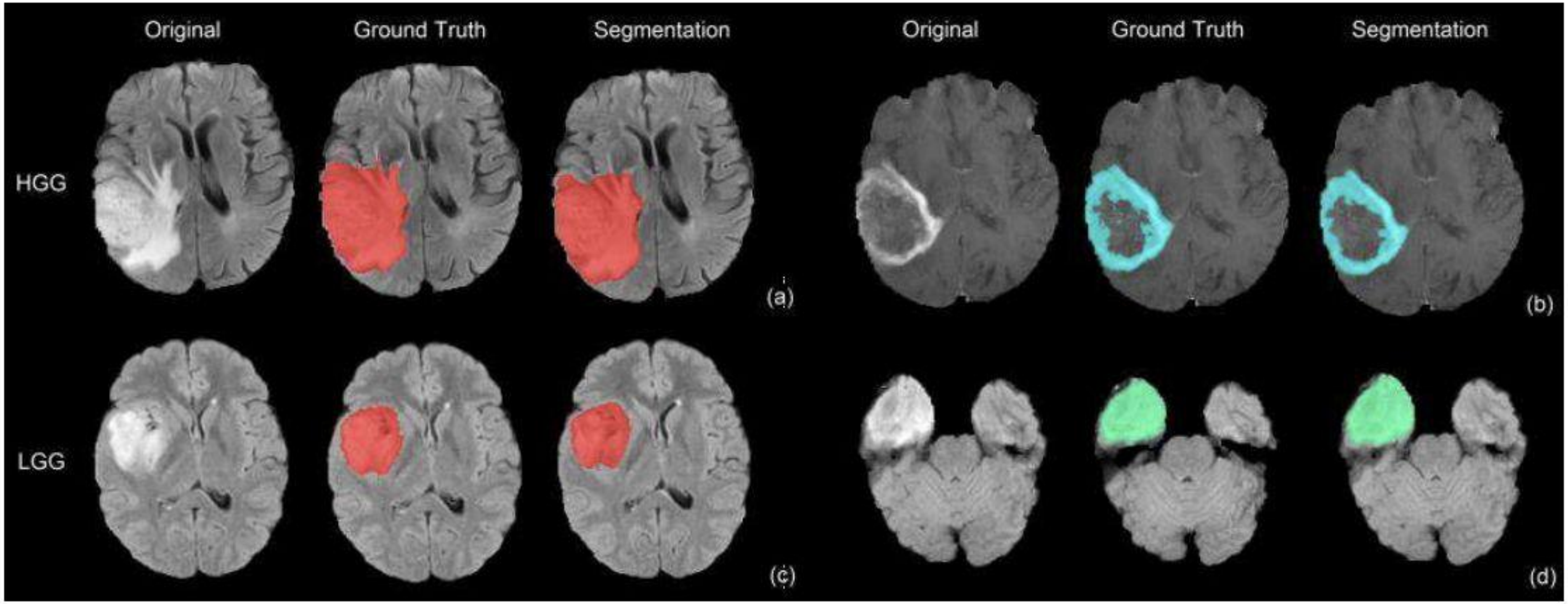
Multiclass segmentation of lungs, clavicle and heart (Novikov et al, IEEE TMI 2018)



Electron Microscopy (EM)
cell segmentation
(Ronneberger et al, 2015)



Segmentation of tumors in brain MR image slices

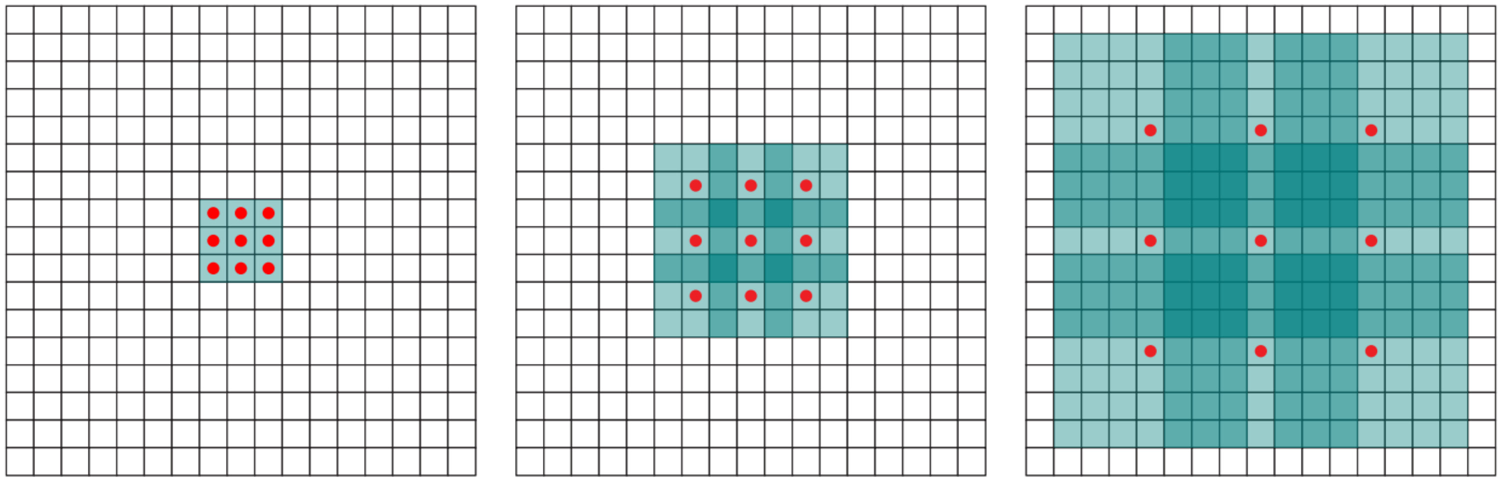


Dong et al, MIUA, 2017

A-Trous/Dilated Convolutions

3x3 Convolution Dilated by 1

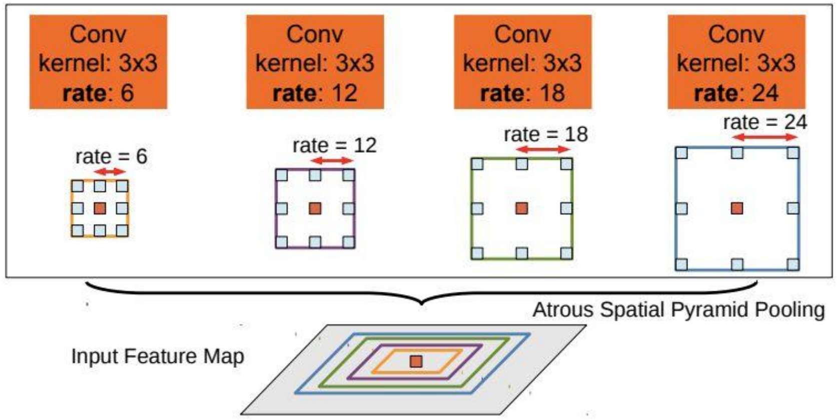
Standard 3x3 Convolution



3x3 Convolution Dilated by 3

Yu et al, ICLR 2016

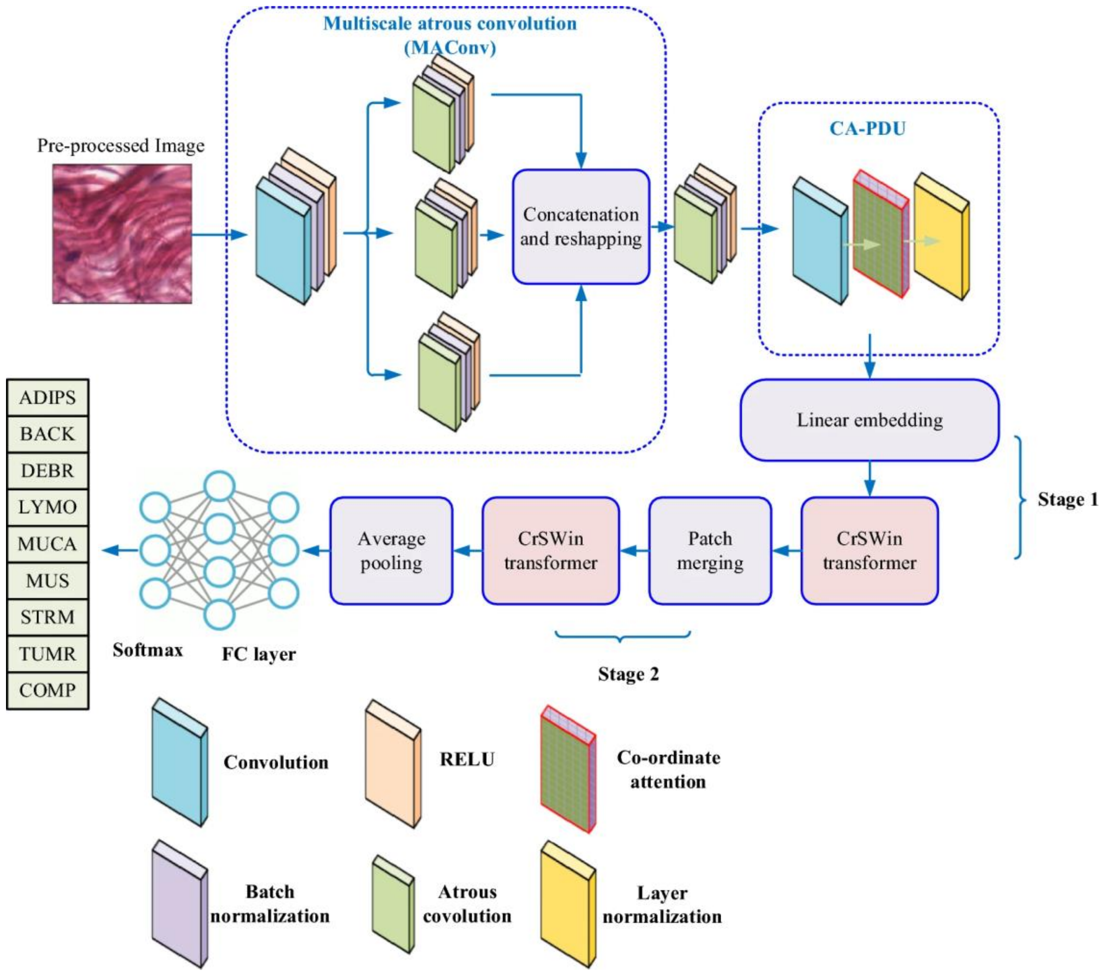
Chen et al, TPAMI 2017



Parameter efficient and can be combined in multiple filter scales to obtain a multiresolution representation

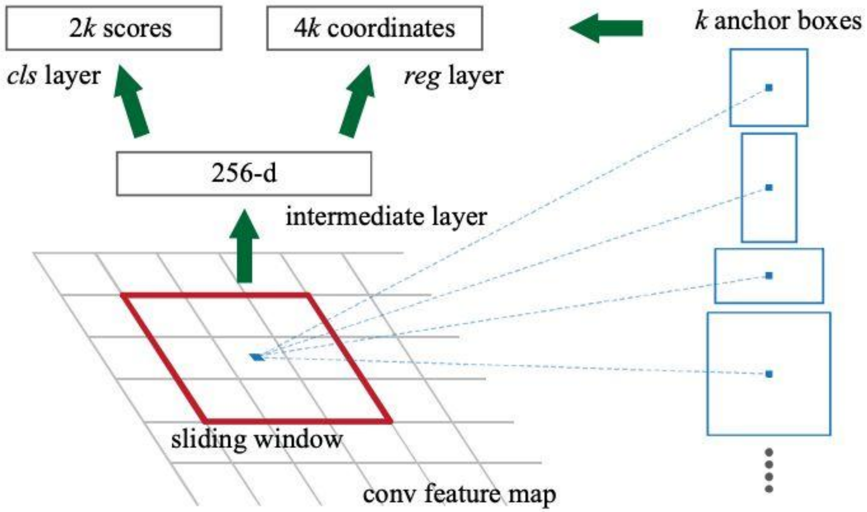
Histopathological tissue segmentation

Multi-resolution a-trous convolutions for categorization of colorectal tissues

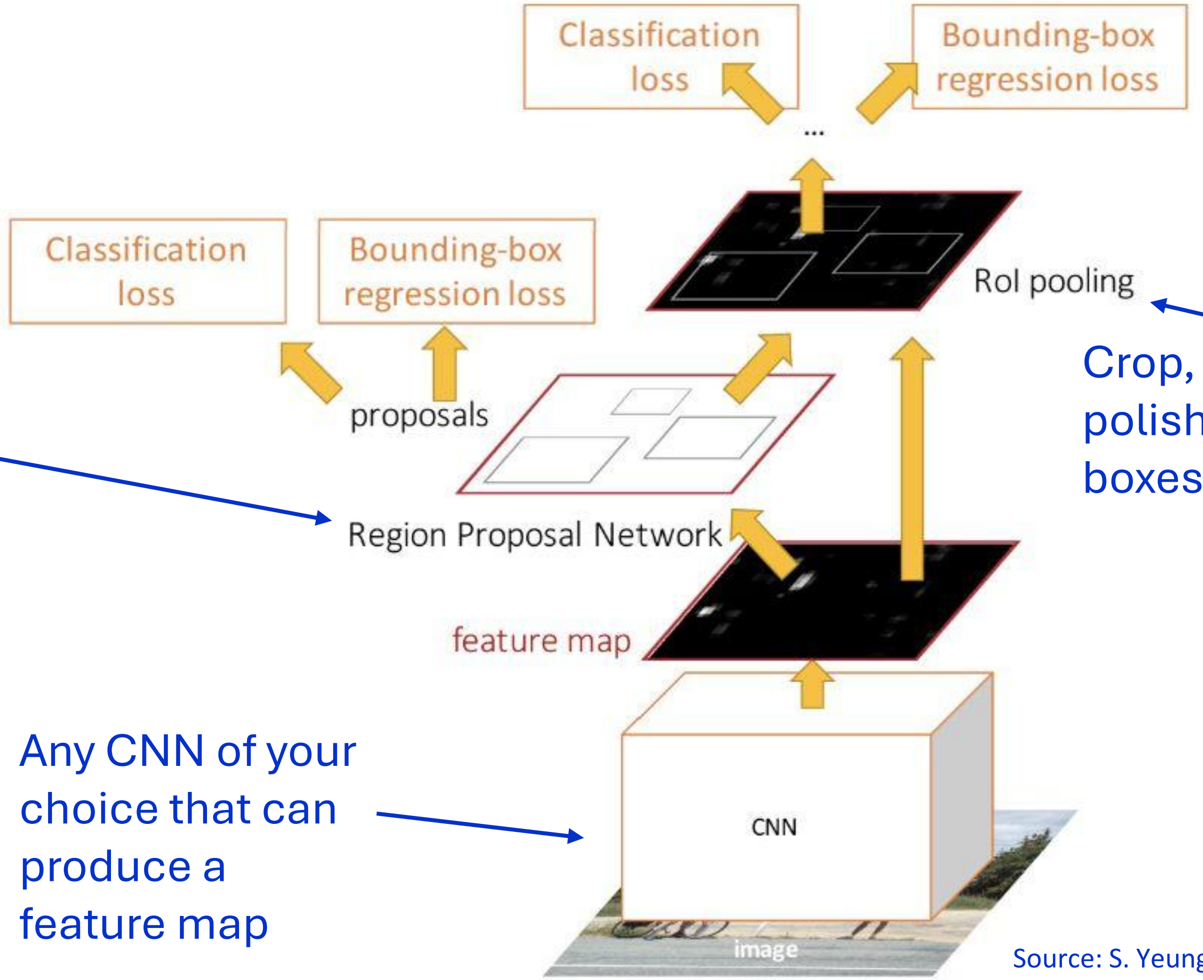


Khalid et al, Scientific Reports, 2024

Detection: Faster R-CNN



Generate bounding boxes proposals
- x,y position
- size
- confidence

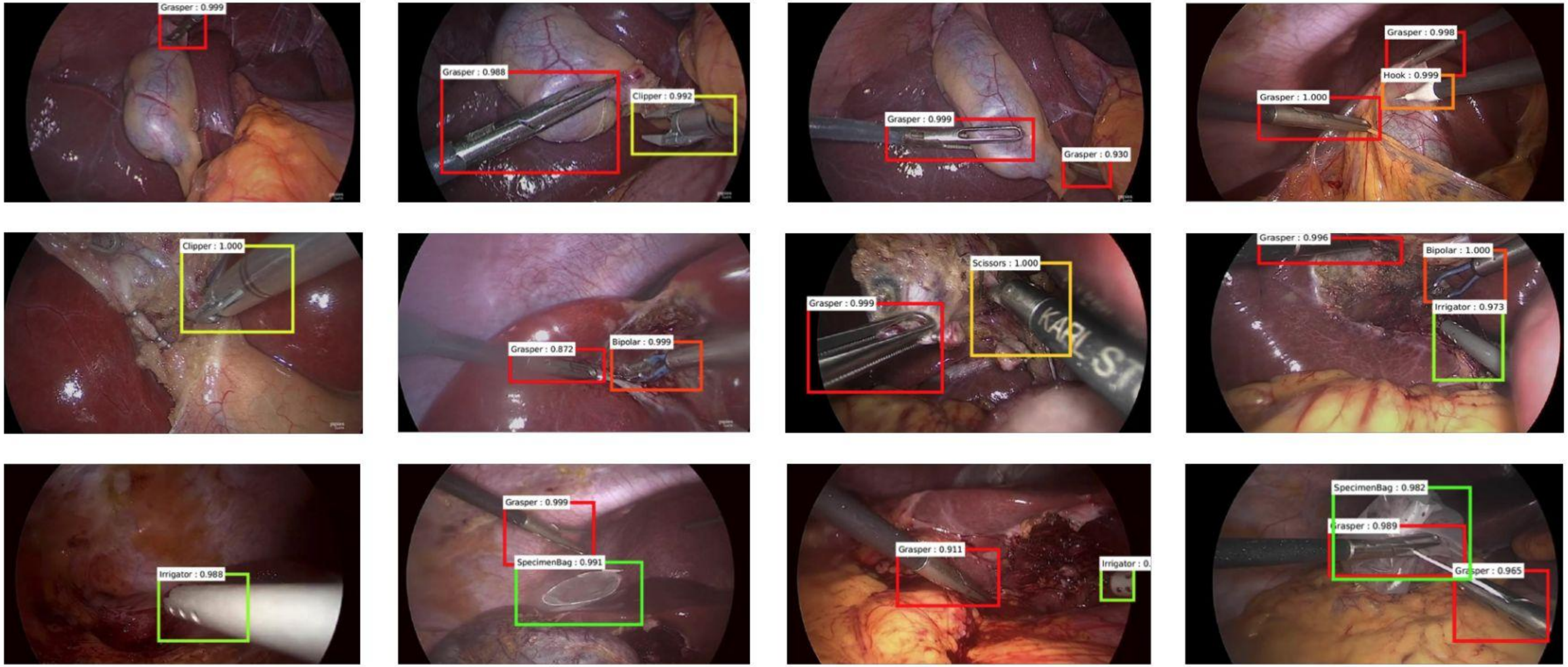


Any CNN of your choice that can produce a feature map

Crop, fuse and polish bounding boxes proposals

Source: S. Yeung, BIODS 220

Detection in surgical video frames



Jin et al, WACV 2018

More CNNs for detection

YOLO - Region proposal generation + region classification are fused into a single stage

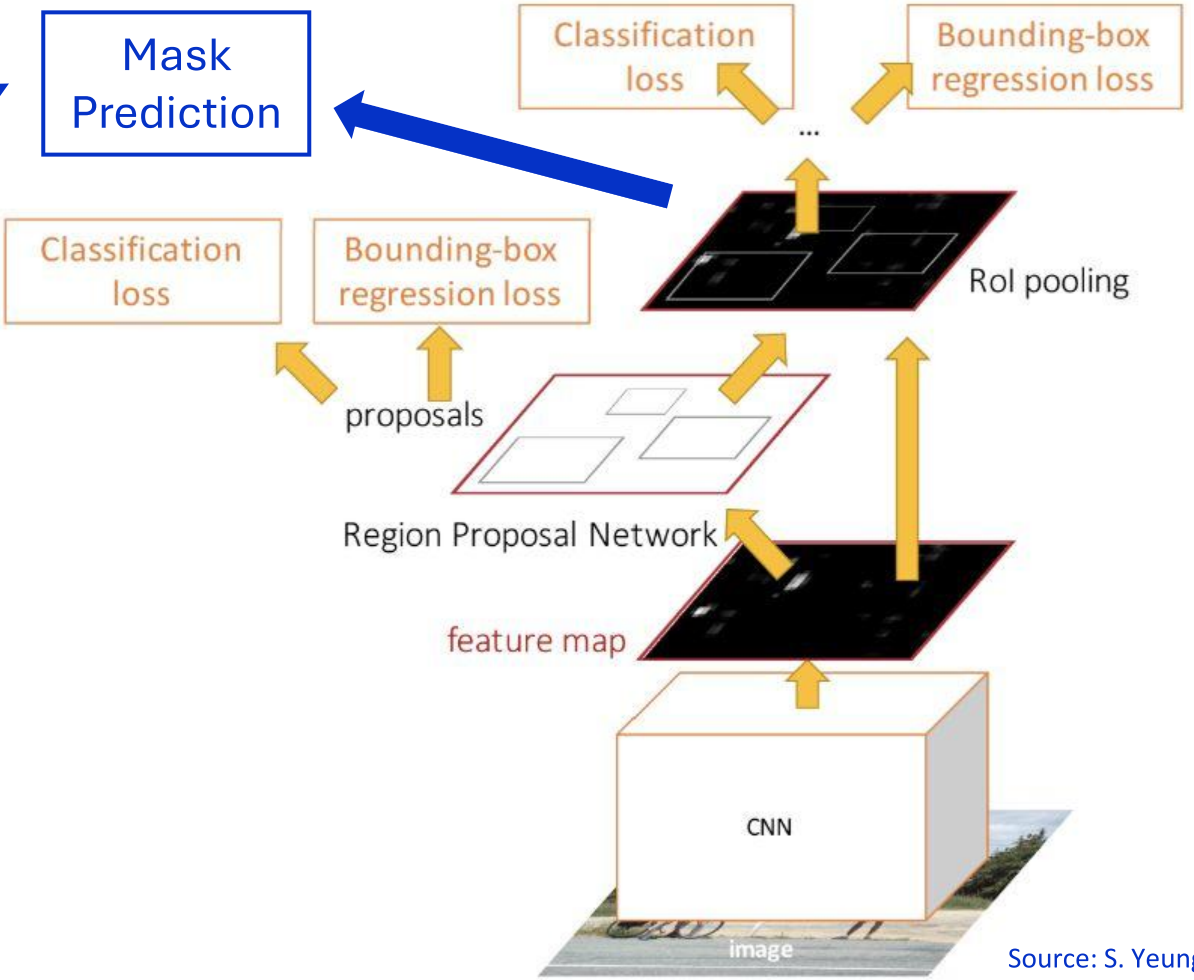
- Faster, but lower performance than two-stage (R-CNN)
- Struggles more with class imbalance relative to two-stage networks that filter only top object candidate boxes for the second stage

RetinaNet: single-stage detector that uses a “focal loss” to adaptively weight harder examples over easy background examples

- Faster than R-CNN on some benchmark tasks, while being more efficient

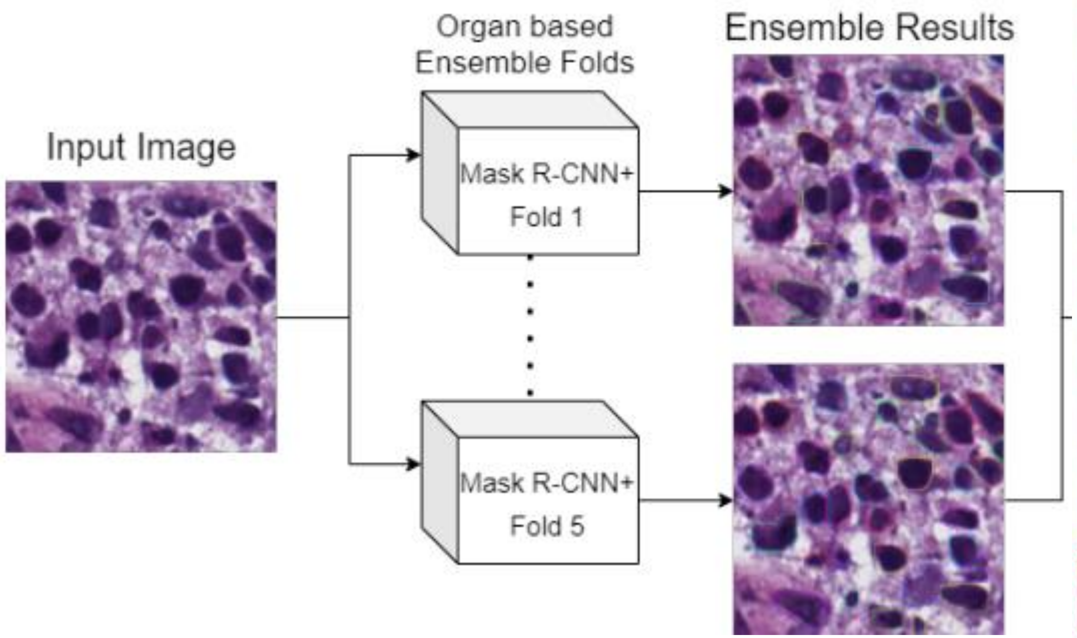
Instance Segmentation: Mask R-CNN

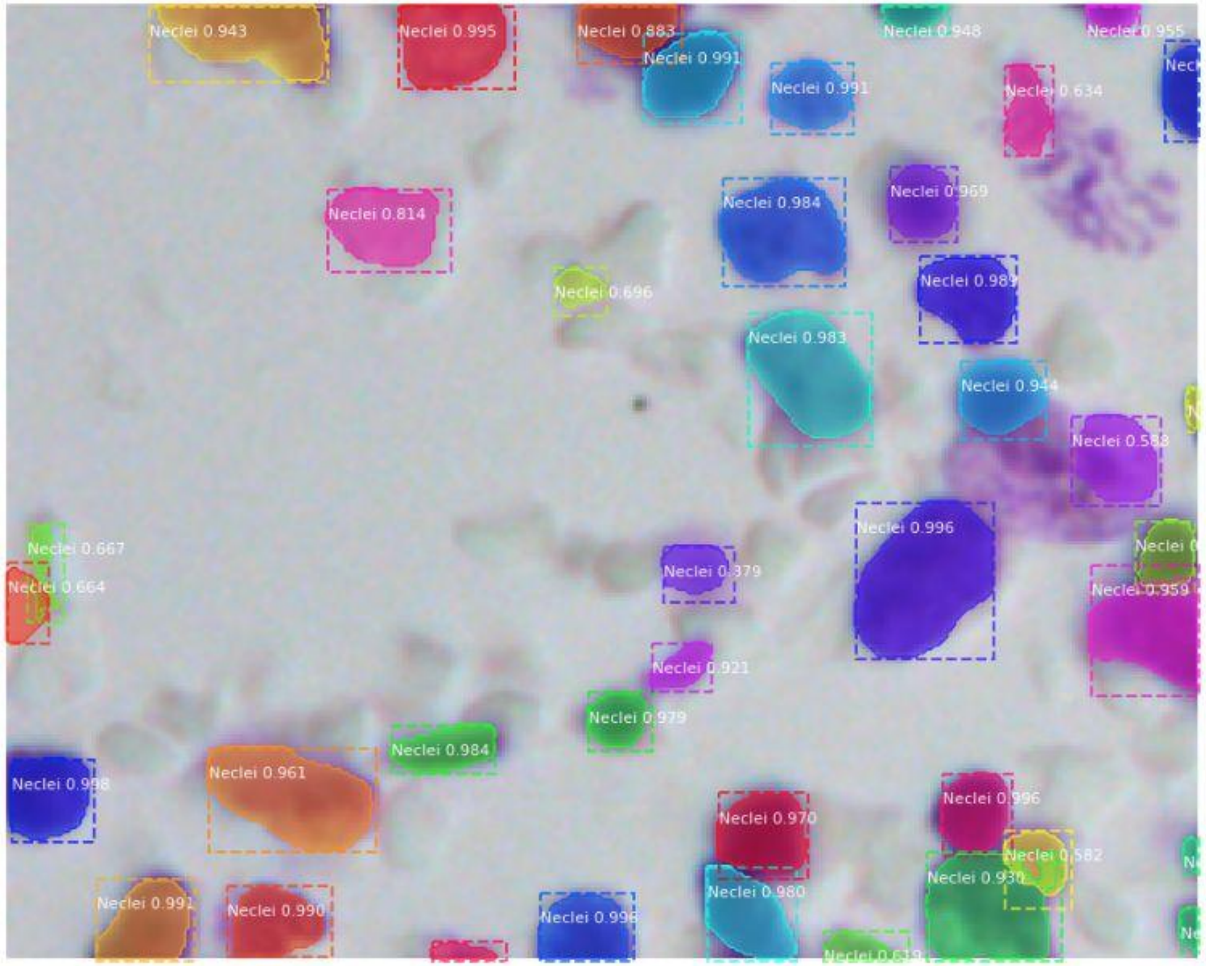
Adds a small mask network that operates on each region of interest to predict a segmentation map rather than just a bounding box



Source: S. Yeung, BIODS 220

Mask R-CNN ensembles for cell nuclei segmentation





Bancher et al, MICCAI-COMPAY 2021

Other predictive tasks in medical imaging

Registration

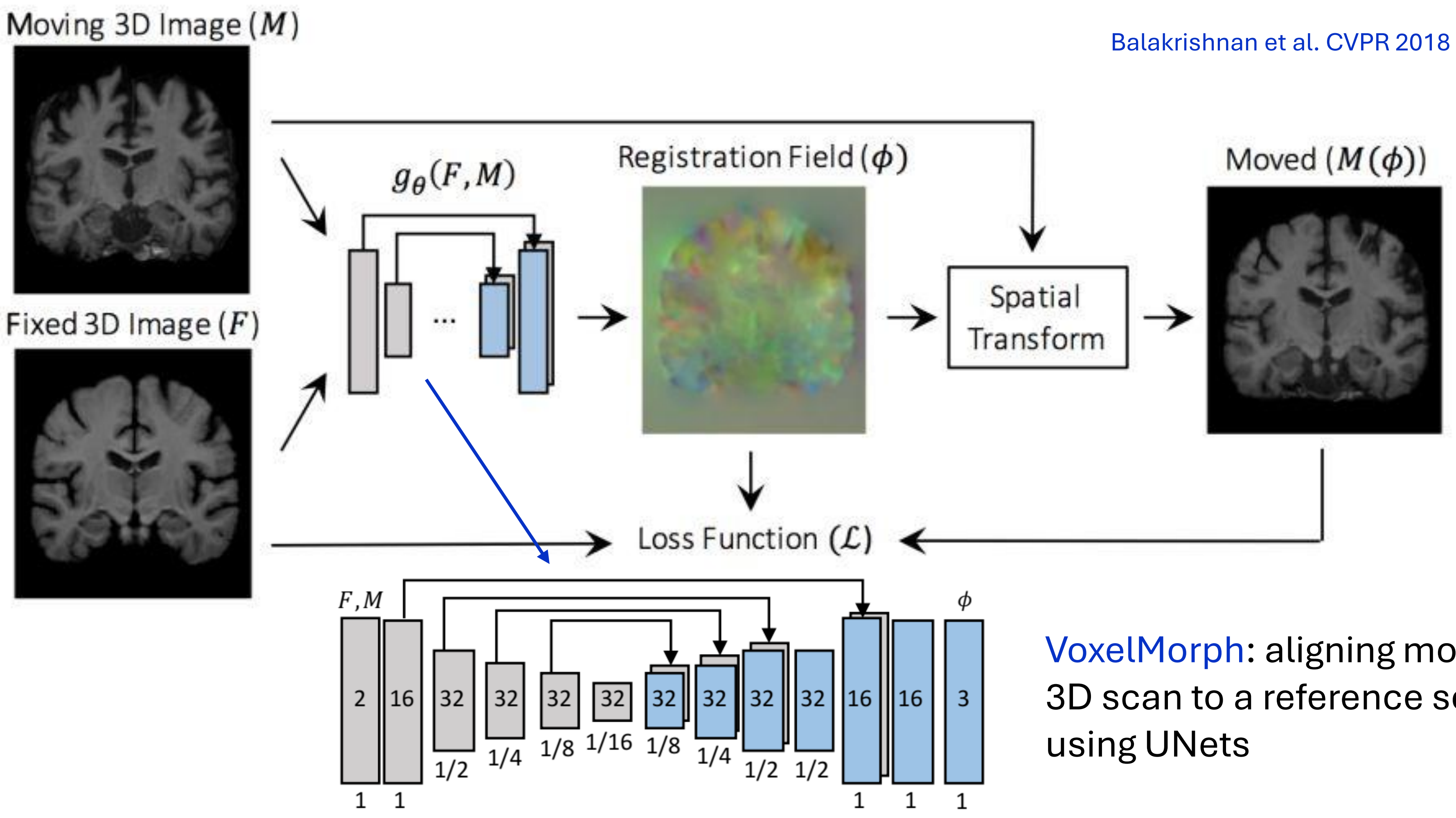
- Align images from the same or different modalities (e.g., CT-MRI fusion)
- Correct for patient movement and acquisition differences
- Essential for multimodal data fusion and longitudinal studies
- Enables precise anatomical mapping and improved diagnosis

Enhancement

- Improve image quality by reducing noise and artifacts
- Enhance contrast and resolution to reveal fine anatomical details
- Critical for revealing subtle pathologies and aiding diagnosis
- Often used as a preprocessing step for better downstream analysis

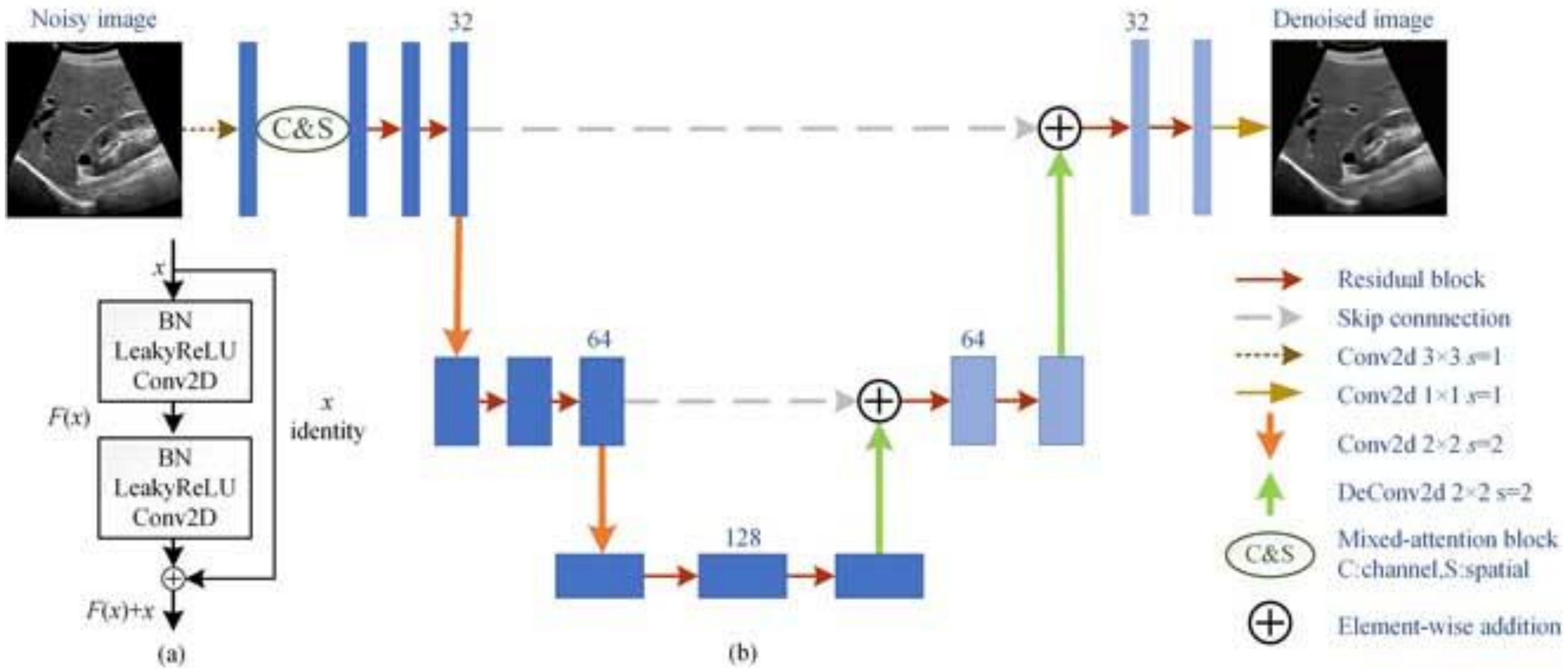
Image Registration

Balakrishnan et al. CVPR 2018



VoxelMorph: aligning moving 3D scan to a reference scan using UNets

Image Enhancement



Lan & Zhang IEEE Access 2020

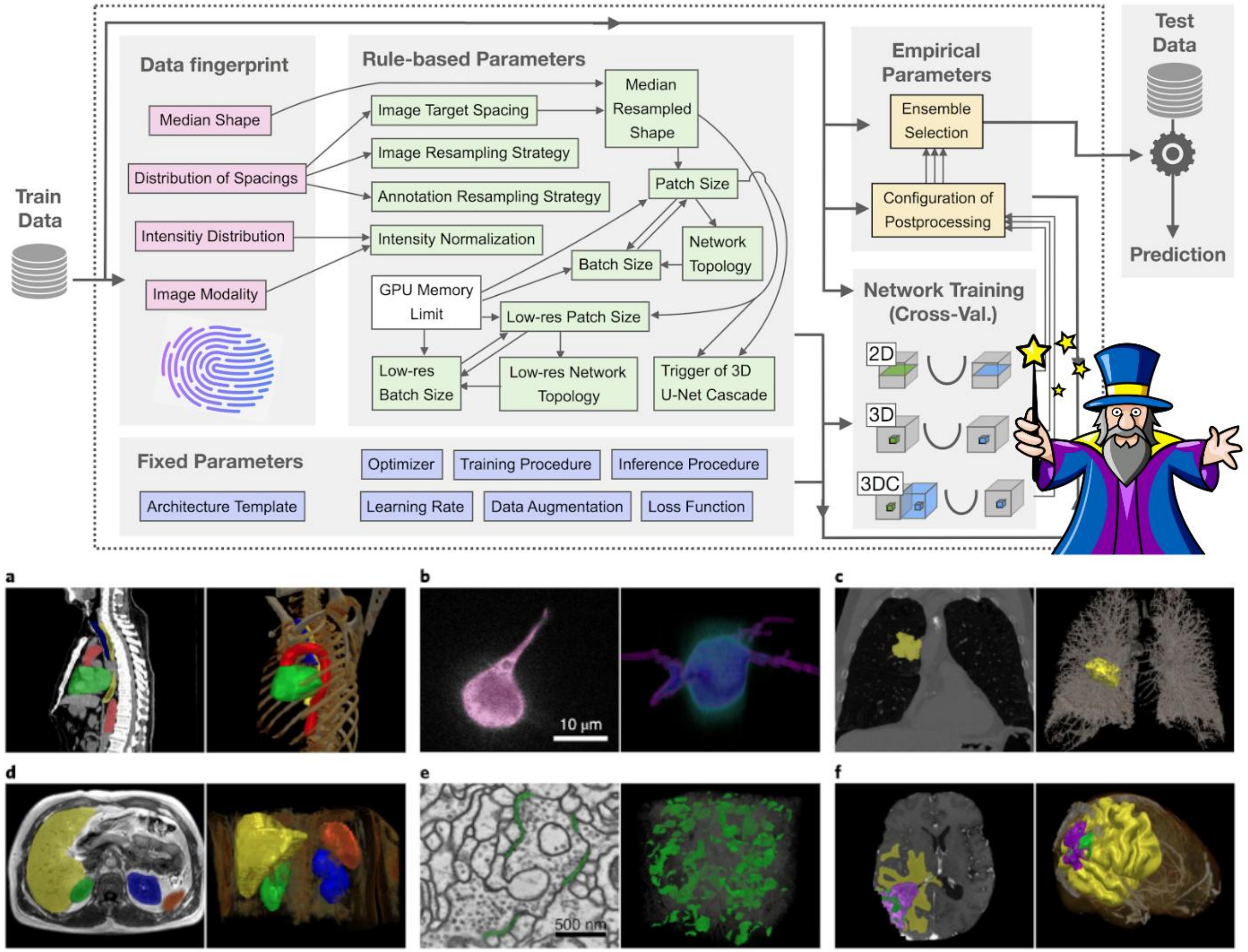
Wrap-up

Towards Engineered Machine Vision for Biomedical Practitioners

nnU-Net (https://github.com/MIC-DKFZ/nnUNet)

- Self-configuring U-Net framework
- Automatically adapts architecture and hyperparameters
- Top performance in multiple segmentation challenges

Part of a new wave of approaches targeting broad generalizability with little model tuning (foundation models)



Take home lessons

- Medical Imaging
  - Diverse modalities and acquisition methods (X-ray, CT, MRI, Ultrasound, Pathology, …)
  - AI tasks: classification, regression, segmentation, detection, registration, enhancement
- Deep Learning Foundations
  - CNNs: convolution, pooling, various CNN architectures
  - Encoder-decoder structures (U-Net): dilated convolutions, a-trous convolutions, residual connections
- Segmentation
  - Critical for accurate delineation (tumors, organs)
  - Specialized architectures and evaluation (and loss) functions
- Practical Considerations
  - Data preprocessing, augmentation, handling imbalance
  - Training strategies, metrics, clinician-in-the-loop validation
- Looking Ahead
  - Highly engineered and generalizable pre-trained models
  - Robustness and explainability

Next Lecture

- Tomorrow: In-itinere lab test

Next week:

- (Tue) Guest lecture on machine vision for brain imaging
  - AI Meets Psychiatry: fMRI-Based Multi-Disorder Diagnosis
  - Elisa Ferrari, CEO Quantabrain
- (Wed) Lab tutorial
- (Thu) Lab tutorial