Deep Learning for Physicists

Lecture #6: Convolutional neural networks | Part 2

Kosmas Kepesidis

Summary of previous lecture…

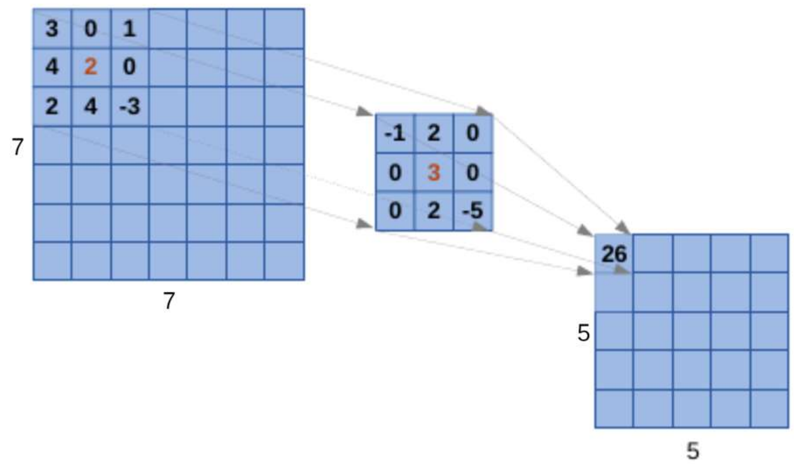
- Convolutional neural networks (CNNs) are the standard architecture for building deep networks to process image-like data

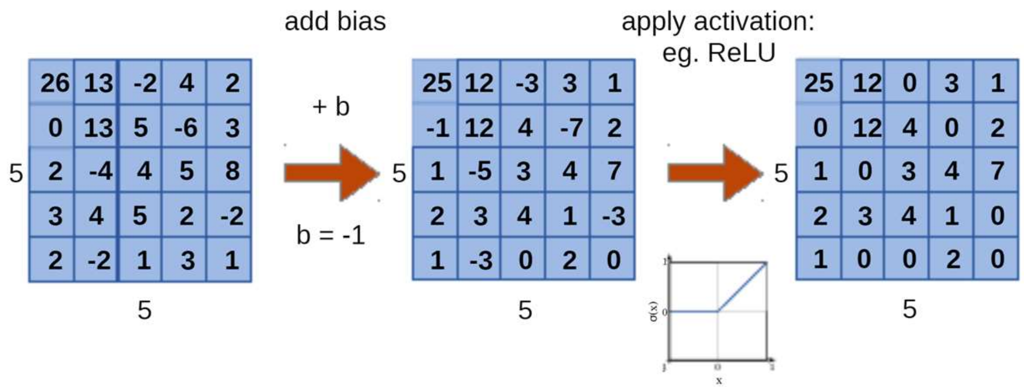
- CNNs simplify the underlying numerical problem by using symmetries that exist in images

- By sliding small filters with adaptive weights over the input, the convolutional operation can deal with variable input and output sizes

- Exploiting symmetry in data allows to reduce the total number of model parameters

Summary of previous lecture…

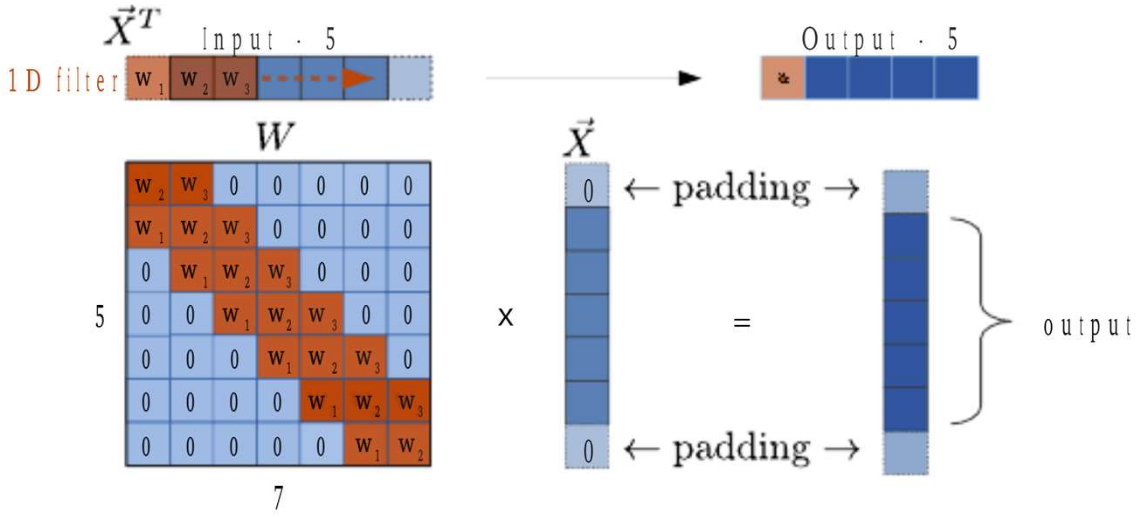




Summary of previous lecture…

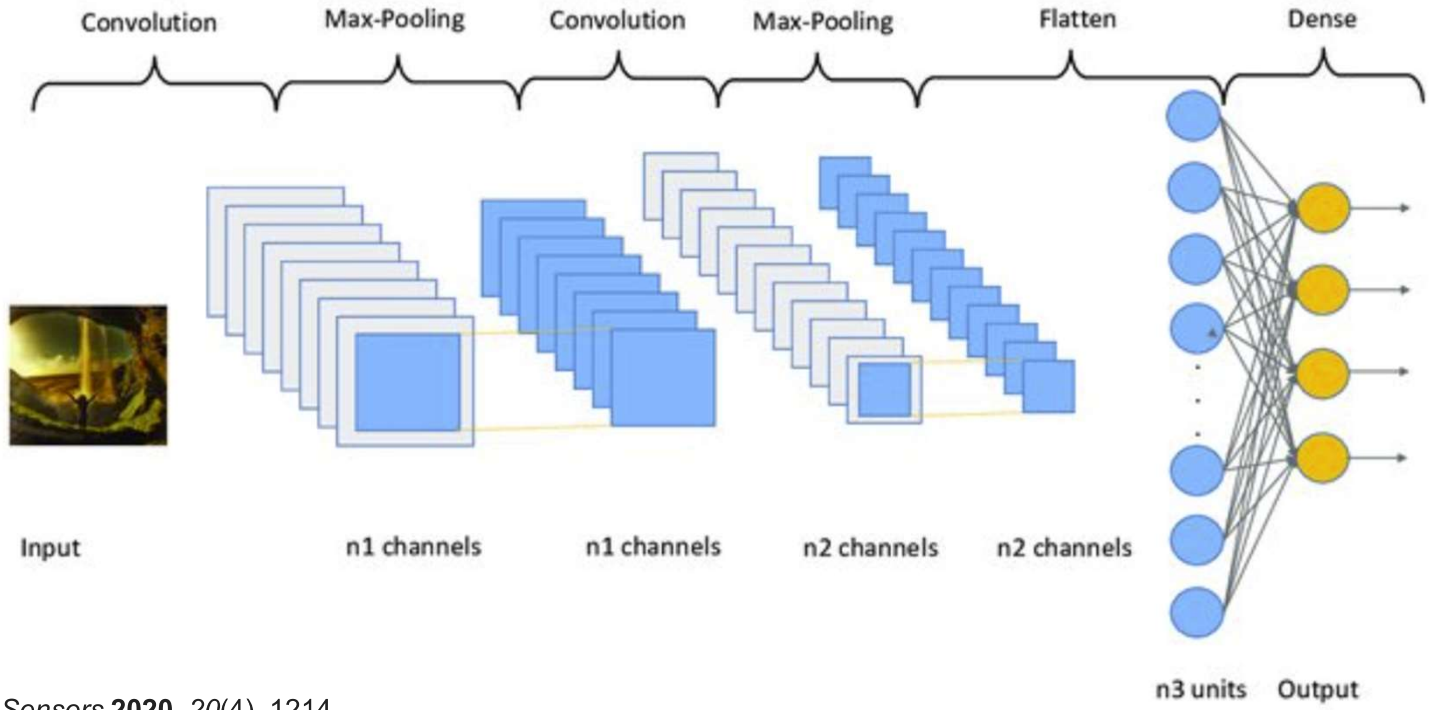
- Exploiting symmetry in data allows to reduce the total number of model parameters

- Advantages:
  1. Simplified optimization (training)
  2. Reducing the chances of overfitting



Summary of previous lecture...

- Adding fully-connected layer:



Sensors **2020**, *20*(4), 1214

Outline

- Convolutional neural networks (CNNs) | Part 1
  - Convolutions of image-like data
  - Convolutional layers
  - Multi-dimensional convolutions
  - Important operations in CNNs
  - Short- and long-range correlations
  - CNNs vs. fully-connected networks

Convolutional neural networks (CNNs) | Part 2

- Reconstruction tasks
- Advanced concepts
- Applications in physics

Convolutional neural networks (CNNs)
Part 2

Reconstruction tasks

Reconstruction tasks

- In principle, two types of general tasks can be performed using CNNs:

    1. Classification: e.g. classify detector's response as signal or background
    2. Regression: e.g. extract the energy of a measures particle

- Such operations can be performed also at the pixel level – examples include:

    - ➢ Semantic segmentation
    - ➢ Pixel-wise regression
    - ➢ Object localization (involves more than one pixel)

Reconstruction tasks

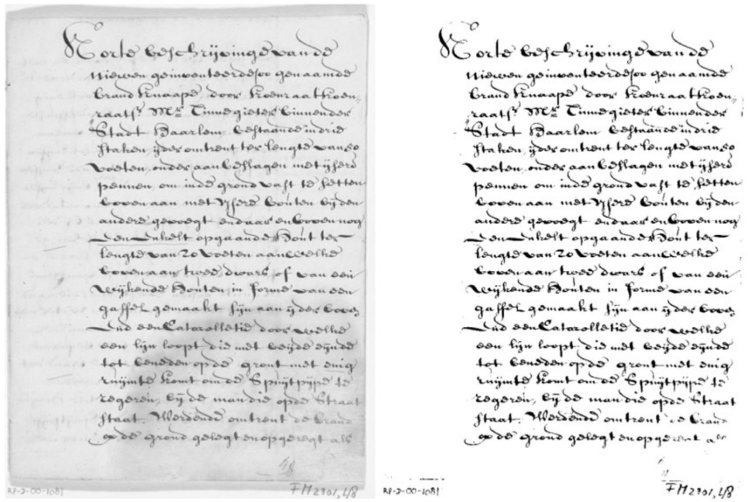
- **Semantic segmentation**

  ➢ In other words: **pixel-wise classification**
  ➢ Network output do not consist of nodes but of maps, one for each class
  ➢ The output maps has the resolution of the input image

- Example: image binarization

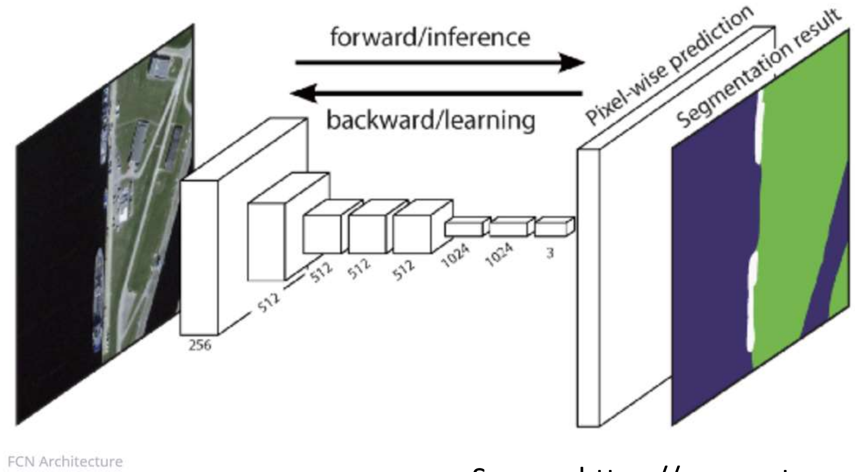
Fig. 1. Gray Scale (Multi Tone) Image   Fig. 2. Binarized Image (Two Tone Image)

Source: madhavuniversity.edu.in



Reconstruction tasks

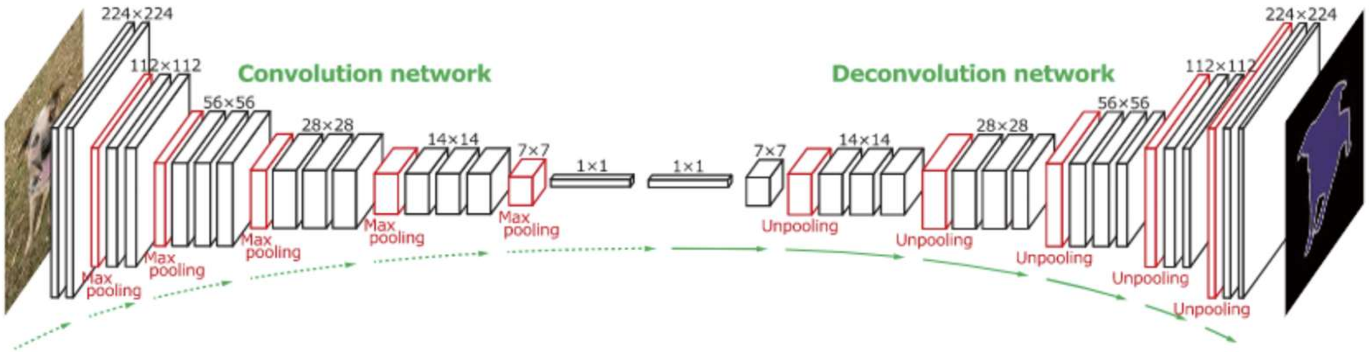
- **Semantic segmentation**
  - ➢ In other words: **pixel-wise classification**
  - ➢ Network output do not consist of nodes but of maps, one for each class
  - ➢ The output maps has the resolution of the input image



Source: https://nanonets.com/

Reconstruction tasks

- **Semantic segmentation**
  - ➢ In other words: **pixel-wise classification**
  - ➢ Network output do not consist of nodes but of maps, one for each class
  - ➢ The output maps has the resolution of the input image



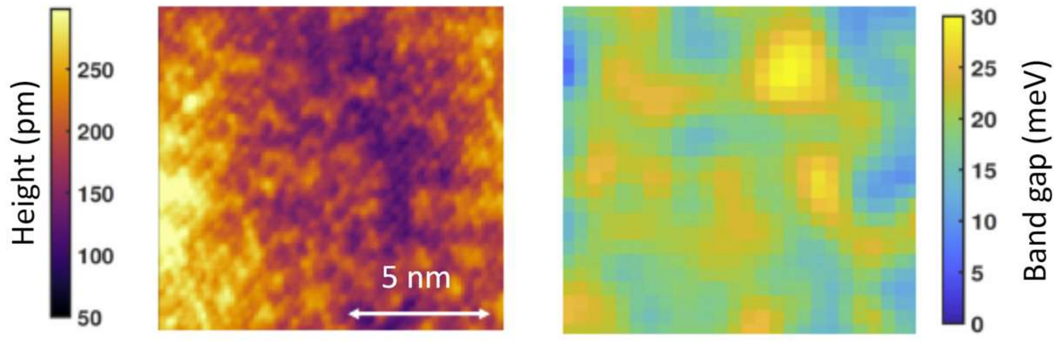
Source: H. Noh et al. (2015)

Reconstruction tasks

- **Pixel-wise regression**

  - Regression of this kind has been of lower importance in applications
  - In physics, there are a lot of emerging applications

- Example: Scanning probe microscopy of the same area of a $MnSb_2Te_4$ epitaxial film
  - From topographic image to local band gap



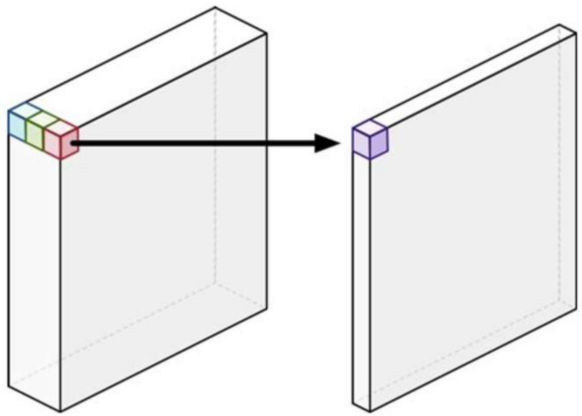
Advanced concepts

Advanced concepts

- In the last years several improvements have been proposed in for CNNs
- We review here some of these improvements that are potentially very relevant to physics applications

Advanced concepts

- **Point-wise convolution**:
  - Special case of convolution operations
  - Act only in the feature domain, using filter sizes of (1×1), each feature map is scaled by a single weight value
  - This transformation can be seen as a fully-connected layer applied to the features space
  - Used for producing low-dimensional representations of the data and reduce the computing effort
    - often called **bottleneck layers**



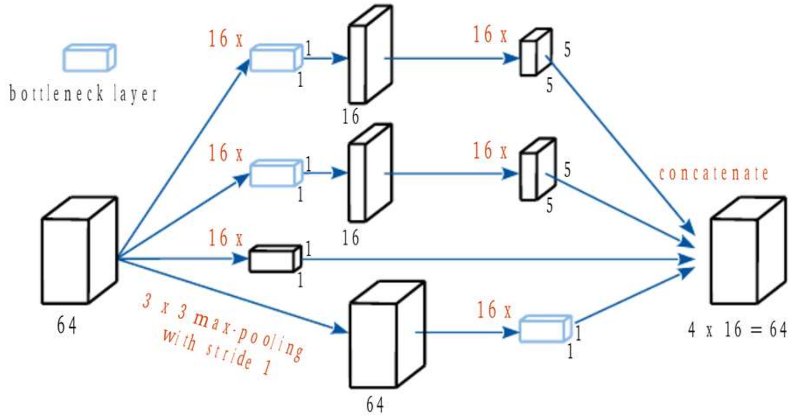
Source: International Journal of Advanced Computer Science and Applications ,Vol. 10, No. 01, 2019

Advanced concepts

- **Locally-connected convolutions**:
  - Similar to the normal convolution operations
  - The difference is that the filter weights are not shared across the image
  - Each local patch of the input is processed by a different filter – translational invariance is broken
  - Useful only when “local” correlations are very strong

Advanced concepts

- **Inception network**:
  - ➢ The size of the filter determines the size of structures that can be learned
  - ➢ Inception technique can make a layer sensitive to objects of different scales
  - ➢ Idea: several filters of different sizes operate in parallel
  - ➢ The learned feature maps are combined again to form a unified feature space, which is processed by the subsequent layers
  - ➢ Multiple bottleneck layers (point-wise convolutions) are used for preventing the feature maps of becoming too large



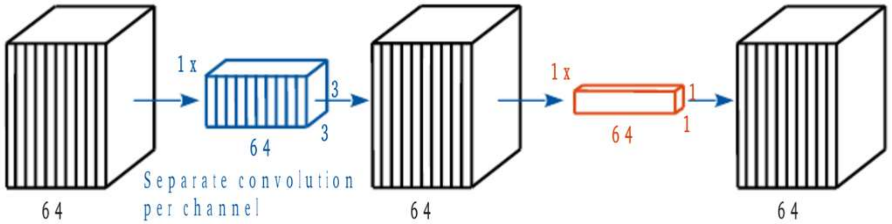
C. Szegedy et al., Going deeper with convolutions, 2015IEEE Conference on Computer Vision and Pattern Recognition (CVPR),pp. 1–9 (2015)

Advanced concepts

- **Separable convolutions**:

  - If spatial correlations are decoupled from cross-channel correlations, it is better to calculate them in separate convolutions – this is an extreme variant of **Inception**

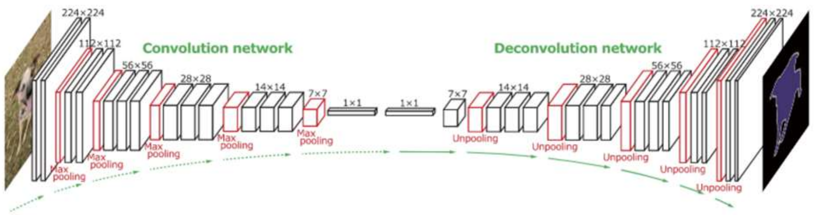
  - A recently published work of a method called **Xception**, decomposes the convolution operation into two subsequent ones (factorization)
    1. First part acts on spatial dimensions of each channel: **depth-wise** convolution
    2. Second part is a **point-wise** convolution that cross-correlates the channels



J. Hu et al., Squeeze-and-excitationnetworks, (2019)
A. G. Howard et al., Mobilenets: Efficient convolutional neural networksfor mobile vision applications, (2017)

Advanced concepts

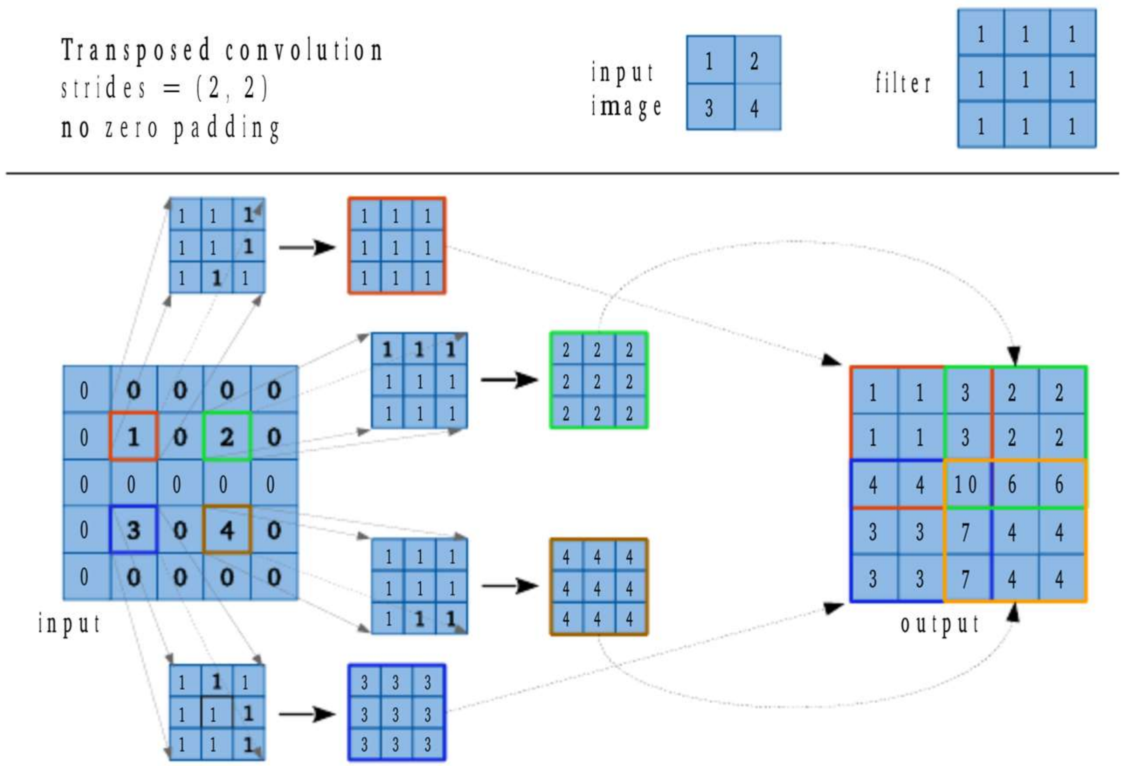
- **Transposed convolutions** or **deconvolutions**:
  - ➢ As we have seen, sometimes it is useful to reverse the order of a convolution
  - ➢ Apply filter on single pixel and output a patch
  - ➢ By sliding the filter over many pixels, several patches can be created and aggregated
  - ➢ It can be thought as if going from extracted features back to the original image
  - ➢ Many applications, such as **semantic segmentation**



Source: H. Noh et al. (2015)

Advanced concepts

- **Transposed convolutions** or **deconvolutions**:



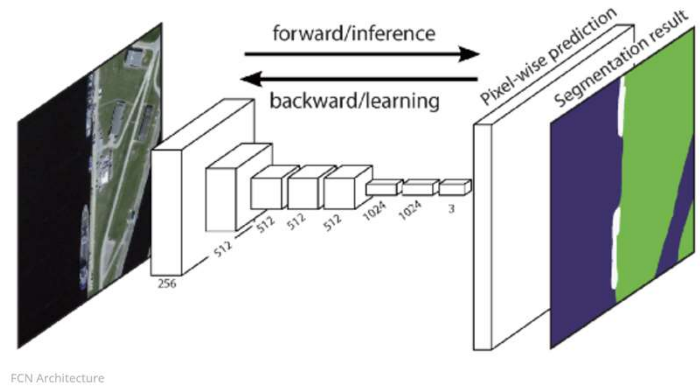
Advanced concepts

- **Upsampling**

  - Another method to increase the size images is simple **upsampling** (as previously discussed)

  - Can be a good replacement of deconvolution in some cases

  - Many applications, such as semantic segmentation



Source: https://nanonets.com/

Advanced concepts

- **Shortcuts**

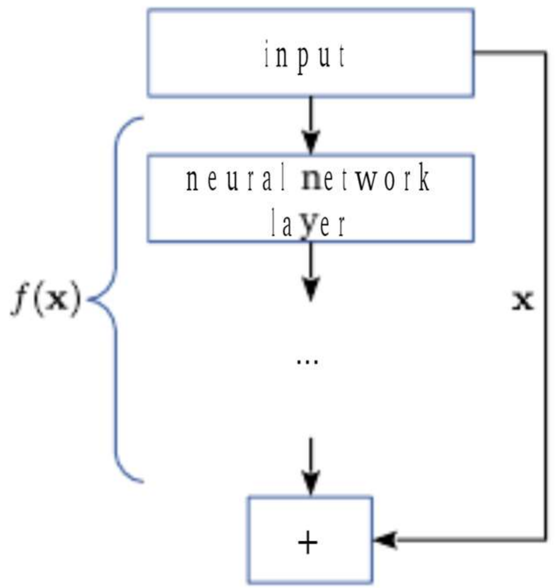
  - We have visited the concepts of **residual learning** and **shortcuts** during our discussion on the fully-connected networks – in a similar way, they can be used also for CNNs

  - Basic idea of **residual network** (ResNet) is to add small modifications to already learned features

  - Figure shows the concept of shortcut in residual networks

    - Operation is divided into two parts:
      1. Pass-on of original input (i.e. learned features of previous layers)
      2. Small change $f(x)$ learned from convolutional layers

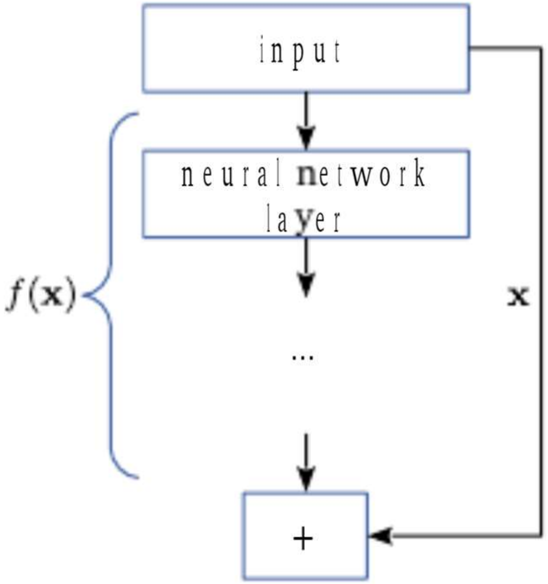
    - Output: $\mathbf{y} = \mathbf{x} + f(\mathbf{x})$



K. He, X. Zhang, S. Ren and J. Sun, Deep residual learning for imagerecognition, IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

Advanced concepts

- **Shortcuts**
  - ➢ Effectiveness of ResNet:
    - ➢ Large predictive capacity can be created using many small mappings
    - ➢ Learned features have priority – already adequately learned features are easily propagated throughout the network
    - ➢ This *bypass* also reduces potential problems with vanishing gradients
    - ➢ ResNet
      - uses shortcuts massively between layers of different hierarchies
      - it is possible to train very deep networks with more than a thousand layers
      - offer an extremely high capacity for complex problems



K. He et al., Deep residual learning for imagerecognition, IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

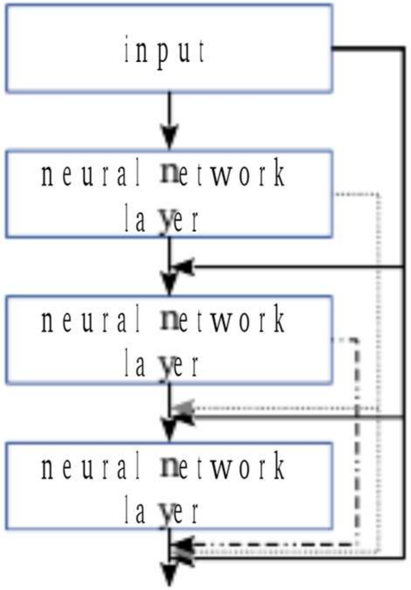
Advanced concepts

- **Shortcuts**

  - **Densely-connected** convolutions (DenseNet) – figure shows the architecture

  - Similar to ResNet but now addition has been replaced by concatenation:
    - Concatenating two feature maps (matrices) of dimensions $m \times n$ results into a feature map of dimensions $2 \times m \times n$

  - All feature maps created are linked together in every subsequent layer
    - Different levels of hierarchy are connected to each other



G. Huang et al., Densely connected convolutional networks, IEEE Conference on Computer Vision and Pattern Recognition, pp. 2261–2269 (2017)

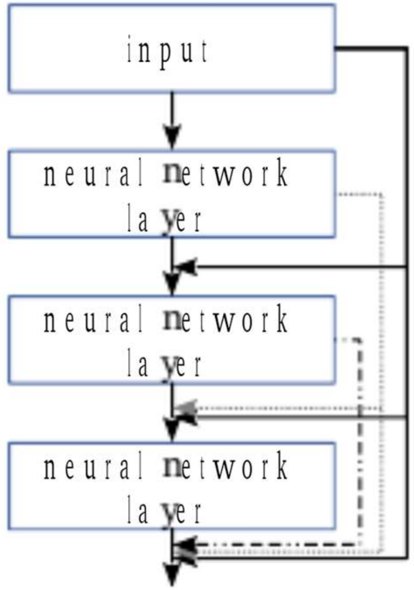
Advanced concepts

- **Shortcuts**

  - Benefits:
    - It is a way of doing **weight sharing** that generally stabilizes the network
    - Simplifies propagation of gradients

  - Problem:
    - Rapid in-crease in memory consumption

  - Solution:
    - Use of point-wise convolutions as transition layers



G. Huang et al., Densely connected convolutional networks, IEEE Conference on Computer Vision and Pattern Recognition, pp. 2261–2269 (2017)

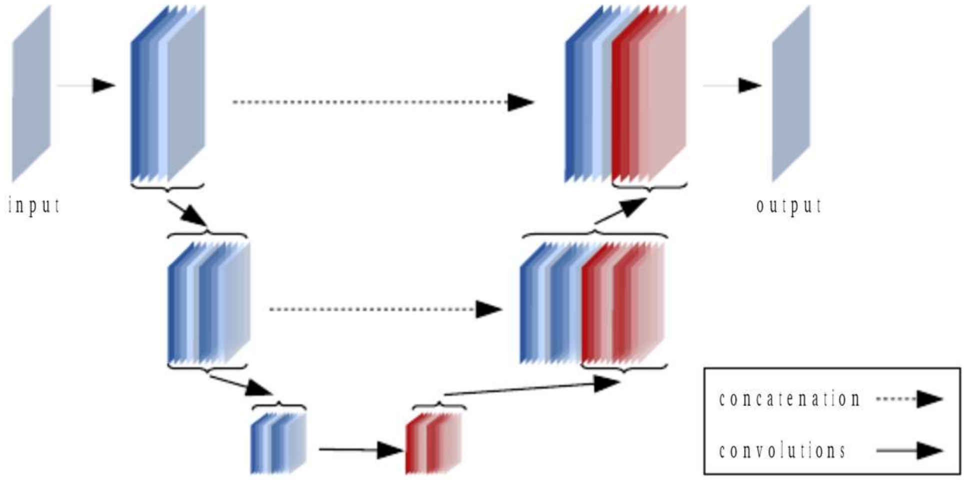
Advanced concepts

- **Shortcuts**

➢ Common tasks in computer vision:
1. Global image classification
2. Identification of objects in images
3. Image segmentation

➢ The problem of distinguishing objects and identify their locations is computationally expensive, since the network need to examine all images pixel by pixel
- Example: answering medical questions using X-ray images

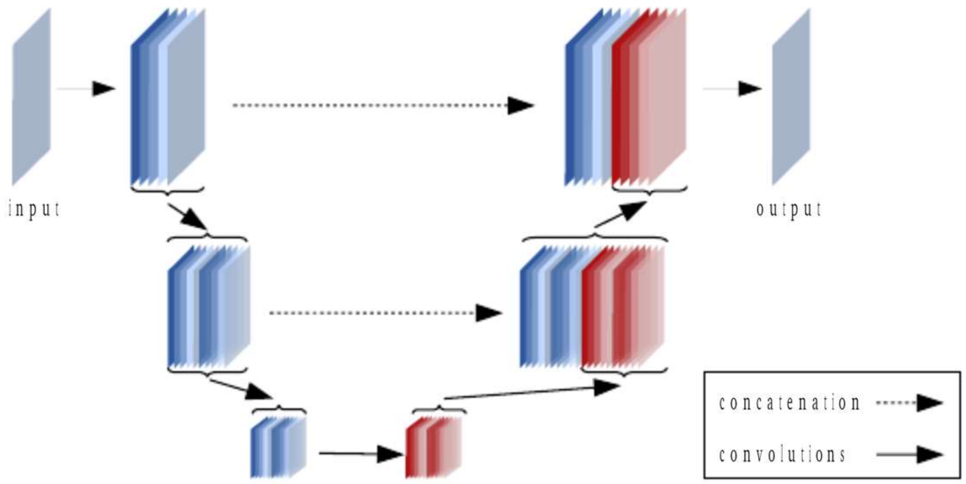
➢ An architecture called **UNet** offers a way for precise object localization and object classification using concatenation operations



O. Ronneberger et al., U-net: Convolutional networksfor biomedical image segmentation, (2015)

Advanced concepts

- **Shortcuts**



➢ **UNet** consists of two parts, a contracting and an expanding part

➢ Contracting part – similar to regular CNN
  - A series of convolutions and a pooling operation can achieve features capable of object identification

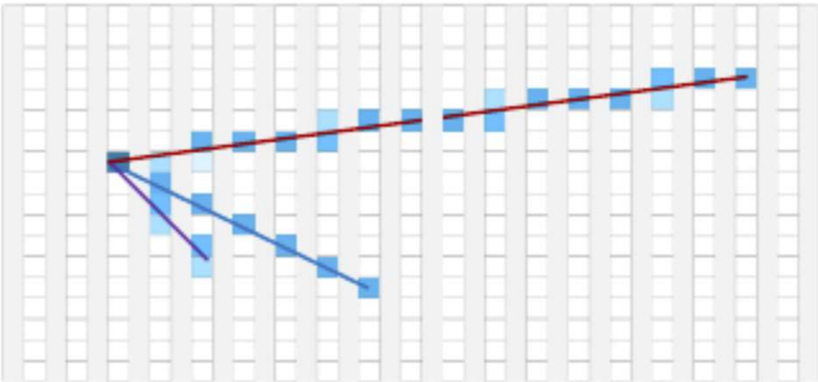
➢ Expanding part – *transposed convolutions*:
  - After transposed convolutions (or upsampling) feature maps are concatenated with same-resolution maps of the contracting part
  - This mechanism is capable of localization

O. Ronneberger et al., U-net: Convolutional networksfor biomedical image segmentation, (2015)

Applications in physics

Applications in physics

- In particle physics experiments, it is common to graphically display collision effects on a grid
- Events of interaction between particles and detectors are recorded, in a way that reveals particle trajectories



- Important questions: determine which particle is observed (electron, muon, etc.)

Applications in physics

- Revealing the identities of particles corresponds to a classification problem
  - typically multiclass classification, where the classes correspond to particle types
  - Trajectories can be used as hints for particles identity
  - Sometimes the signals themselves are sensitive enough to differences in particles
  - This type of data has similarities to images that can be directly processed using **CNNs**

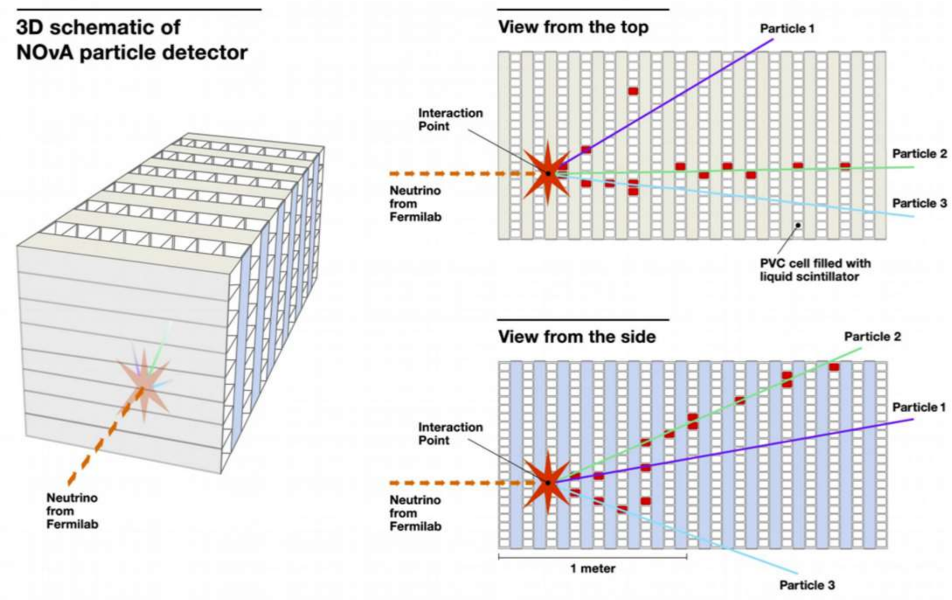
- Example: neutrino oscillations (NOvA)
  - Neutrinos captured by detectors after a flight of 800 km from the generation site (FermiLab)
  - The type of neutrino was determined by detecting an electron or muon as a collision biproduct
    - muons leave a $\sim$ 10 m long narrow trace of signals
    - electron traces are less long and broader

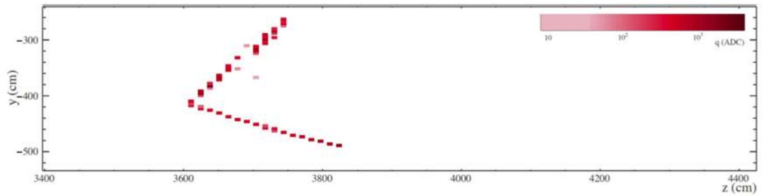
  - Binary-classification problem!

A. Aurisano et al., A convolutional neural network neutrino event classifier, Journal of Instrumentation 11, 09 (2016)

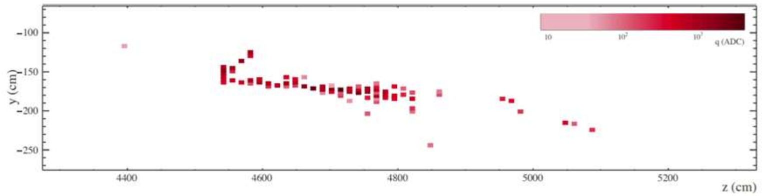
Applications in physics

- Example: neutrino oscillations

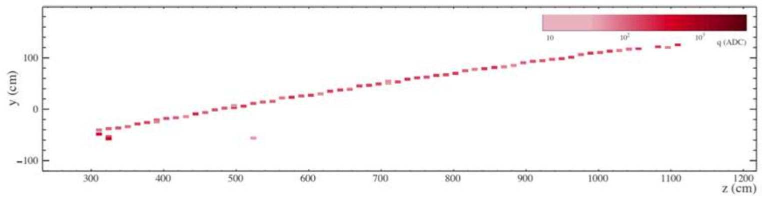




(a) A ν_e CC QE electron plus one hadron signature where the upward-going shower-like prong with multiple hit cells on each plane corresponds to an electron and the downward-going track-like prong with approximately one hit per plane correspond to a proton.



(b) A ν_e CC RES electron plus hadron shower signature with a characteristic electron shower and short prongs which could correspond to multiple hadrons.

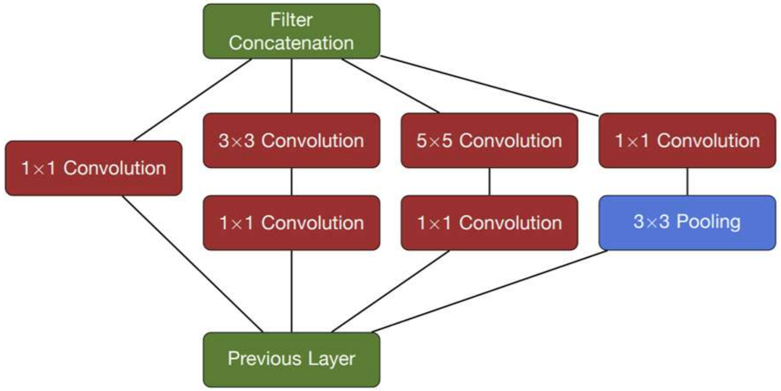


(c) A ν_μ CC QE muon plus one hadron signature with a long track-like prong with lower charger-per-cell corresponding to a muon and a short prong with larger charge-per-cell corresponding to a proton.

A. Aurisano et al., A convolutional neural network neutrino event classifier, Journal of Instrumentation 11, 09 (2016)

Applications in physics

- Example: neutrino oscillations
    - ➢ Using the **inception** concept (discussed previously), convolutional filters were trained on the track patterns of electrons and muons
    - ➢ The network improved the identification of electrons by about 30% over previously-used methods

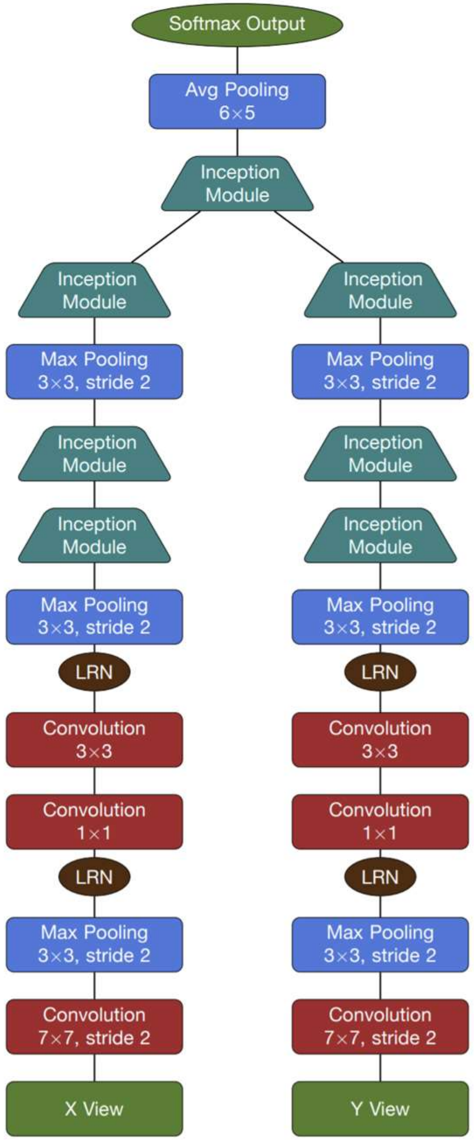


A. Aurisano et al., A convolutional neural network neutrino event classifier, Journal of Instrumentation 11, 09 (2016)

Applications in physics

- Example: neutrino oscillations

    - ➢ Using the **inception** concept (discussed previously), convolutional filters were trained on the track patterns of electrons and muons
    - ➢ The network improved the identification of electrons by about 30% over previously-used methods



A. Aurisano et al., A convolutional neural network neutrino event classifier, Journal of Instrumentation 11, 09 (2016)

Applications in physics

- Example: quarks and gluons

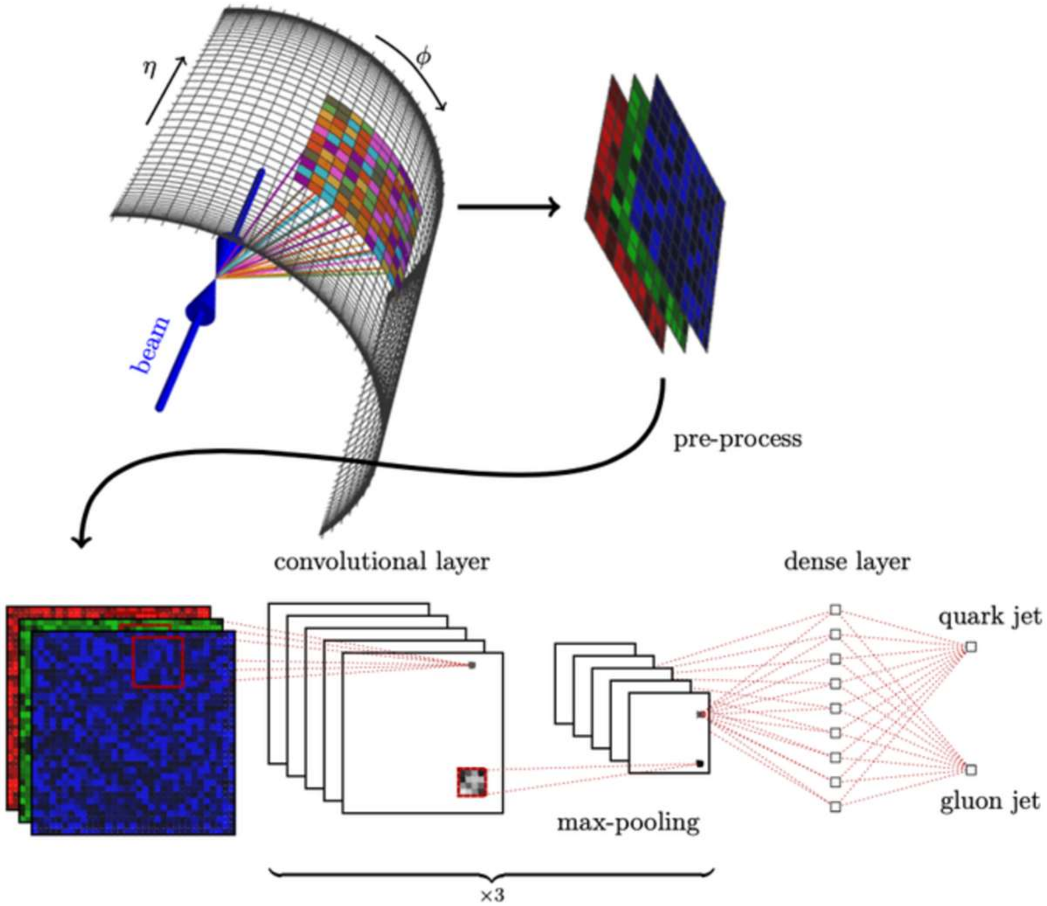
    - Particle collisions at LHC (CERN) produce particle jets
    - These jets hit the detectors:
        - Jet thickness is smaller (with fewer particles) if it originates from *up/down quark*
        - Jet thickness is larger (with more particles) if it originates from a *gluon*

    - Particle jets were projected on a effective cylindrical coordinates (n, ϕ)
    - Idea: use physical quantities as colors on an image (RGB channels)
        - Green: transverse momenta of charged particles
        - Blue: transverse momenta of neutral particles
        - Red: particle charge

P. T. Komiske et al., Deep learning in color: Towards automated quark/gluon jet discrimination, Journal of High Energy Physics 01 (2017) 110

Applications in physics

- Example: quarks and gluons



P. T. Komiske et al., Deep learning in color: Towards automated quark/gluon jet discrimination, Journal of High Energy Physics 01 (2017) 110

Applications in physics

- Example: pixel-wise segmentation of neutrinos
  - For *neutrino* measurements in a LArTPC (liquid argon time projection measurements), the **UNet** architecture has been proven to be efficient for pixel-wise segmentation of particle tracks using 2D projections of underlying physics events

C. Adams et al., Deep neural network for pixel-level electromagnetic particle identification in the microboone liquid argon time projection chamber, Physical Review D 99, 9 (2019)

Applications in physics

- Example: fluorescence microscopy

  - **Diffraction limit** is a physical barrier that restricts the resolution of image details to almost half the optical wavelength

  - Super-resolution methods have been developed

  - Example: single-molecule localization microscopy (SMLM)
    - Permits the creation of high-resolution images by integrating many localization outcomes into single images
    - CNNs have been successfully used for the localization problem
    - Turning the localization problem into a classification one
      - Example: is there an object (such as molecule) within a pixel/voxel or not

L. Schermelleh et al., Super-resolution microscopy demystified, Nature Cell Biology 21, pp. 72–84 (2019)

Summary

- Convolutional neural networks (CNNs) are the standard architecture for building deep networks to process image-like data

- CNNs simplify the underlying numerical problem by using symmetries that exist in images

- By sliding small filters with adaptive weights over the input, the convolutional operation can deal with variable input and output sizes

- Exploiting symmetry in data allows to reduce the total number of model parameters

Summary

- By stacking convolutional layers, the receptive field of view increases, which allows to extract features of different hierarchies

- More complex architectures involve shortcuts and filters that are sensitive to different scales – they can help build very deep models and improve the performance

- Using batch normalization in CNNs can increase the models’ convergence and performance, especially when training very deep models