

# Neuronové sítě 2 - Konvoluční neuronové sítě - typy úloh

18NES2 -9. hodina, ZS 2024/25

Zuzana Petříčková

20. listopadu 2024

# Neuronové sítě 2 - Konvoluční neuronové sítě - typy úloh

- 1 Co jsme dělali minule
- 2 Známé architektury konvolučních neuronových sítí
- 3 Aplikace konvolučních neuronových sítí

## Co bylo minule

- Umělé neuronové sítě a zobecňování
  - Praktické ukázky - MLP (imdb), CNN (Mnist)
- Funkcionální API v Kerasu a vytváření složitějších architektur
  - Ukázka ensemble modelu
- Konvoluční neuronová síť pro klasifikaci - praktická ukázka (Cats vs. Dogs)
  - Model učený od začátku

# Dnes

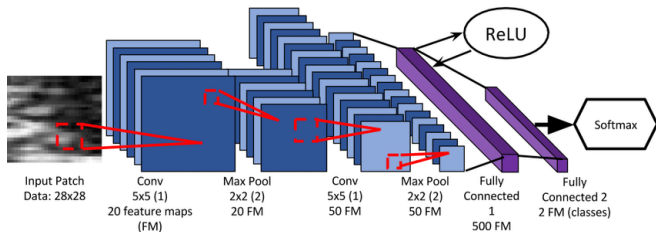
- Dokončení příkladu: Konvoluční neuronová síť pro klasifikaci - praktická ukázka (Cats vs. Dogs)
  - Model učený od začátku
  - Zlepšení schopnosti modelu zobecňovat pomocí regularizace (dropout, batch-normalization) a augmentace dat
  - Model učený pomocí přeneseného učení (transfer learning)
- Známé architektury konvolučních sítí
- Aplikace konvolučních neuronových sítí

## Známé architektury konvolučních neuronových sítí

- nejstarší typy architektur: široké, mělké, často hlubší plně propojená část, odpovídají více základnímu schématu

### LeNet 5

- jedna z původních architektur (Yann LeCun, 1998), poměrně jednoduchá, učená na MNIST

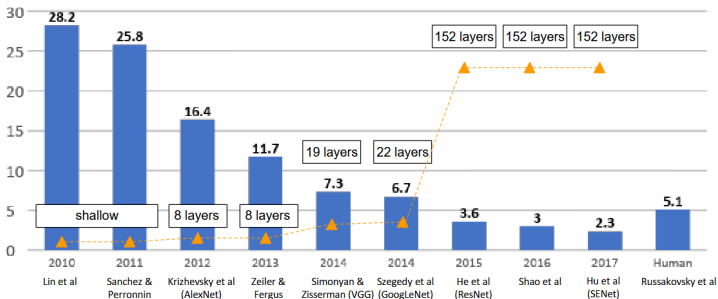


Zdroj obrázku: M. H. Yap et al., "Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks," in IEEE Journal of Biomedical and Health Informatics, vol. 22, 2018.

# Známé architektury konvolučních neuronových sítí

- ILSVRC soutěž: zásadní soutěž nad datovou sadou ImageNet

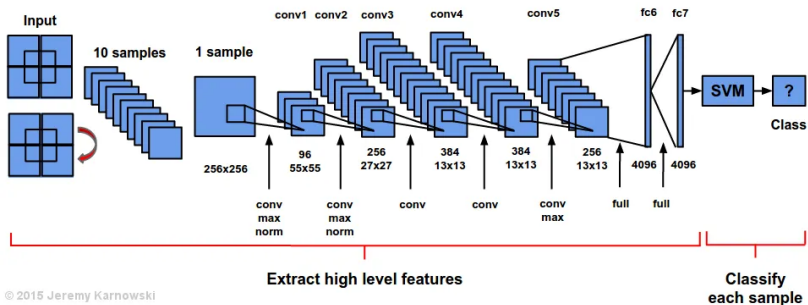
ImageNet Large Scale Visual Recognition Challenge (ILSVRC) winners



Zdroj obrázku: [https://cs231n.stanford.edu/slides/2024/lecture\\_6\\_part\\_1.pdf](https://cs231n.stanford.edu/slides/2024/lecture_6_part_1.pdf)

# AlexNet

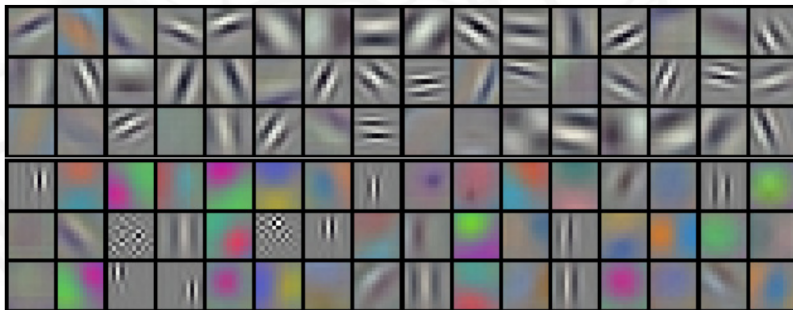
- první CNN-vítěz soutěže ILSVRC (2012) nad ImageNet-úspěšnost 84,7 % v top-5),
- již výrazně složitější (8 vrstev, 61M parametrů, 2 GPU 5-6 dnů)



Zdroj obrázku: <https://medium.com/@jkarnows/alexnet-visualization-35577e5dcd1a>

# AlexNet

- 96 filtrů  $11 \times 11 \times 3$  v první konvoluční vrstvě u AlexNet

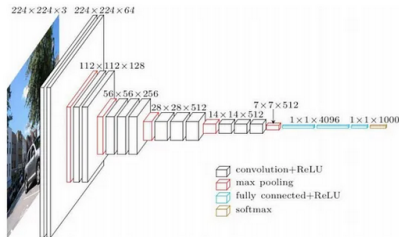
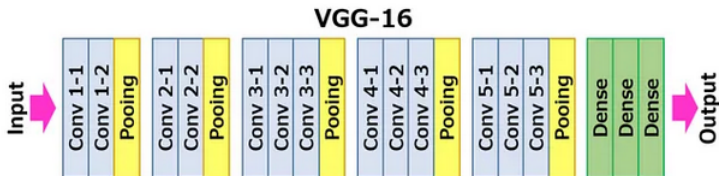


Zdroj obrázku: Alex Krizhevsky et al.: "ImageNet Classification with Deep Convolutional Neural Networks", Figure 3



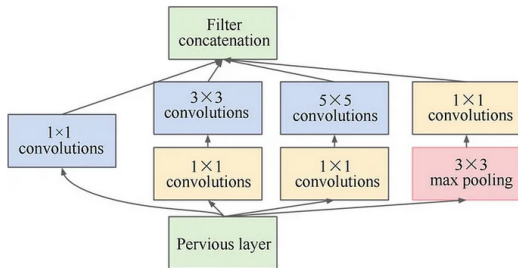
## VGGNet

- účastník ILSVCR 2014, rodina modelů (např. VGG16 - 16 vrstev, VGG19 - 19 vrstev) - úspěšnost 92,7 % v top-5
- pyramidová architektura



## GoogLeNet (Inception v1), 2014

- vítěz soutěže ILSVRC 2014, 22 vrstev  
- úspěšnost 93,33 % v top-5
- revoluční prvek: Inception bloky -  
extrakce příznaků v různých škálách,  
paralelismus, efektivita



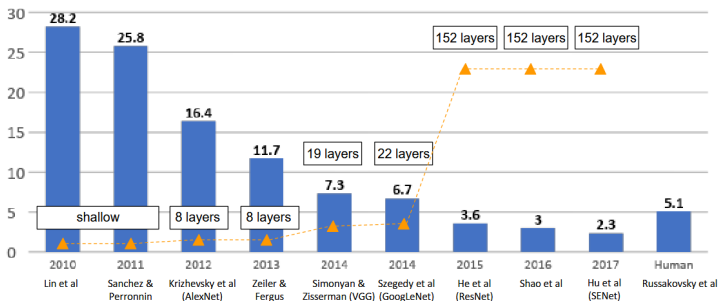
The illustration is from the original paper[1]



## Známé architektury konvolučních neuronových sítí

- ILSVRC soutěž po roce 2014: výrazně hlubší a užší architektury (omezení redundance: depthwise separable convolutions), modularita

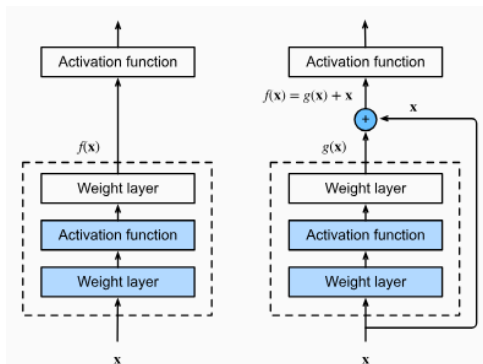
ImageNet Large Scale Visual Recognition Challenge (ILSVRC) winners



Zdroj obrázku: [https://cs231n.stanford.edu/slides/2024/lecture\\_6\\_part\\_1.pdf](https://cs231n.stanford.edu/slides/2024/lecture_6_part_1.pdf)

## ResNet, 2015

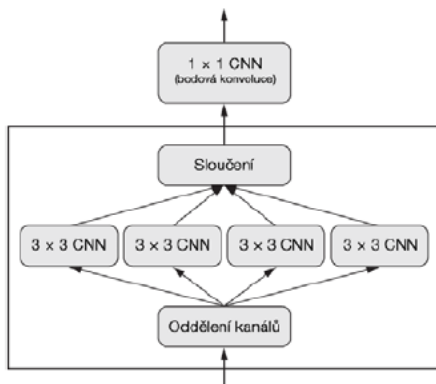
- vítěz soutěže ILSVRC 2015, 152 vrstev - úspěšnost 96,4 % v top-5
- revoluční prvek: skip connections - řeší problém tiché pošty
- residuální bloky:



Zdroj obrázku:

## Hlubkově oddělitelné konvoluce (Xception, 2017)

- další revoluční prvek na cestě ke zvýšení efektivity a zeštíhlení vrstev
- prostorová konvoluce pro každý kanál nezávisle, výstupní kanály se nakonec smíchají pomocí bodové konvoluce



## Další populární architektury

- Inception v2, v3 (2015, 2016)
- DenseNet: (Gao Huang, 2016)
- MobileNet (Google, 2017), EfficientNet (2019), SqueezeNet (2016)- další úspora výpočetních prostředků
- NASNet (Neural Architecture Search Network, 2017) - model se naučil vhodnou architekturu sám, genetické algoritmy, reinforcement learning
- ConvNeXt, 2020
- ...

### **V Kerasu:**

- <https://keras.io/api/applications/>

## Další populární benchmarkové datové sady

### **CIFAR-10, CIFAR-100**

<https://www.cs.toronto.edu/~kriz/cifar.html>

- CIFAR-10: 60000 barevných obrázků, 10 tříd (auta, psi, lodě,...)
- CIFAR-100: 100 tříd
- malé obrázky, dataset populární pro rychlé testování a porovnání modelů

### **COCO (Common Objects in Context)**

- detekce objektů, segmentace a popis obrázků



# Aplikace konvolučních neuronových sítí

- klasifikace obrázků - klasifikační hlava
- regresní úloha - regresní hlava

**Příklad regrese: predikce sklonu číslic**

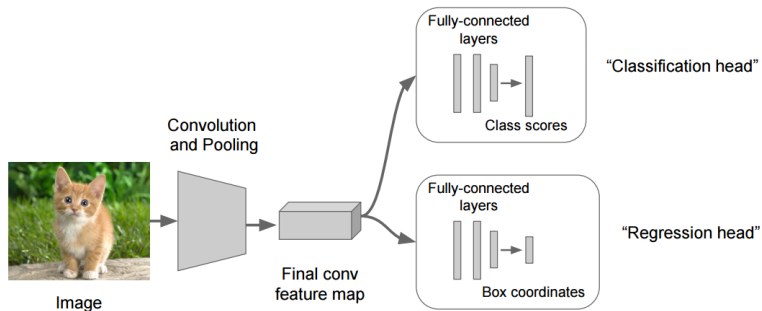


Zdroj : <https://matlabacademy.mathworks.com>



# Aplikace konvolučních neuronových sítí

- klasifikace obrázků - klasifikační hlava
- regresní úloha - regresní hlava



Zdroj : <https://i.stack.imgur.com/FGrD1.png>

# Aplikace konvolučních neuronových sítí

## Zpracování 2D obrazu

- klasifikace obrázků - klasifikační hlava
- regresní úloha - regresní hlava
- detekce objektů - detekční hlava
- segmentace obrazu - segmentační hlava
- restaurování obrazu, stylizace obrazu, generování obrazu - architektura encoder-decoder

## Jiný typ dat

- analýza videa (3D konvoluce) - akční detekce (např. sport. záznam)
- sekvenční data (1D konvoluce) - časové řady, audio data, omezeně i přirozený jazyk)

# Aplikace konvolučních neuronových sítí

## Other Computer Vision Tasks

### Semantic Segmentation



GRASS, CAT,  
TREE, SKY

No objects, just pixels

### Classification + Localization



CAT

Single Object

### Object Detection



DOG, DOG, CAT

Multiple Object

### Instance Segmentation



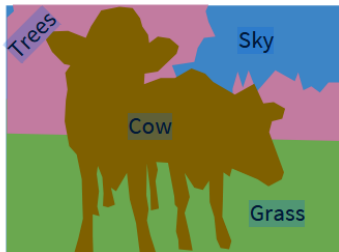
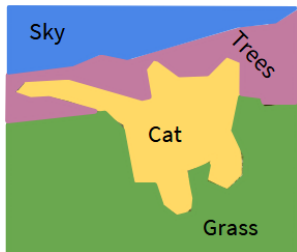
DOG, DOG, CAT

This image is CC0 public domain

Zdroj : [https://cs231n.stanford.edu/slides/2017/cs231n\\_2017\\_lecture11.pdf](https://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf)

## CNN a sémantická segmentace

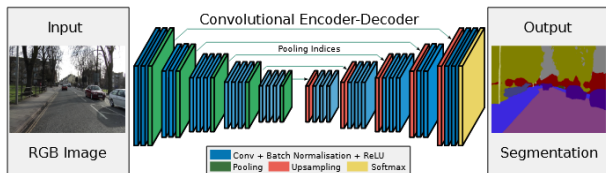
- každý pixel obrázku chceme označit labelem třídy



# Aplikace konvolučních neuronových sítí - sémantická segmentace

## SegNet, U-net,...

- enkodér-dekodér architektura
- enkodér: klasická CNN, extrakce příznaků
- latentní prostor: kompaktní reprezentace
- dekodér: obnovuje obrázek do původních rozměrů:  
transponované konvoluční vrstvy (a příp. unpooling vrstvy)



# Aplikace konvolučních neuronových sítí

## Other Computer Vision Tasks

### 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

Zdroj : [https://cs231n.stanford.edu/slides/2017/cs231n\\_2017\\_lecture11.pdf](https://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf)

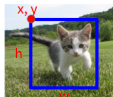
# CNN a Detekce objektů

## Jeden objekt - dvě hlavy:

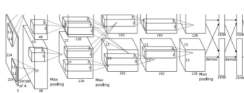
- klasifikace
- nalezení okenka  $\rightarrow$  regrese

### Object Detection: Single Object

(Classification + Localization)



[This image is CC0 public domain](#)



Vector:  
4096

Fully  
Connected:  
4096 to 1000

Class Scores

Cat: 0.9

Dog: 0.05

Car: 0.01

...

**Multitask Loss**

Fully  
Connected:  
4096 to 4

Box  
Coordinates  
( $x, y, w, h$ )

Correct label:  
Cat

Softmax  
Loss

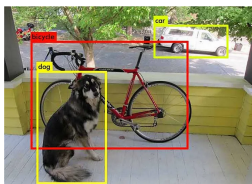
+

Loss

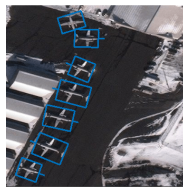
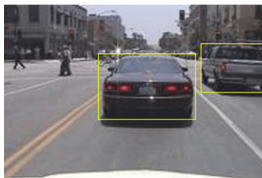
Correct box:  
( $x', y', w', h'$ )

Treat localization as a  
regression problem!

# CNN a Detekce více objektů



Source: Blog by Matthias Hellmann



Zdroj : <https://matlabacademy.mathworks.com>

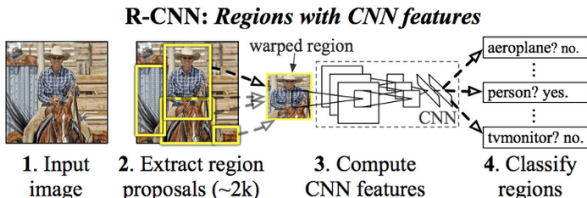
Triantafyllidou, D. et al: A Fast Deep Convolutional Neural Network for Face Detection in Big Visual Data.



# CNN a Detekce více objektů

## R-CNN (Region-based Convolutional Neural Network)

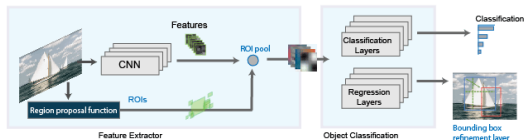
- 1 obrázek je rozdělen na různé oblasti (ROI) pomocí algoritmu selektivního hledání
- 2 každá oblast je zvětšena na stejnou velikost a předložena předučené konvoluční neuronové síti (VGG-16 v původním článku).
- 3 Nakonec jsou použity dvě hlavy: klasifikátor pro klasifikaci každé oblasti a regresor pro doladění jeho hranic.



## CNN a Detekce více objektů

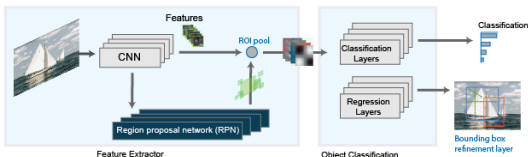
### varianta: fast R-RCNN

- rychlejší výpočet díky sdílení příznaků mezi oblastmi zájmu (ROI - region of interest)

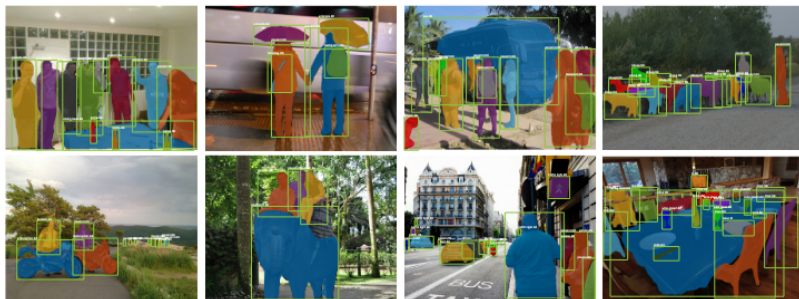


### varianta: faster R-RCNN

- automatické rozpoznávání ROI uvnitř modelu



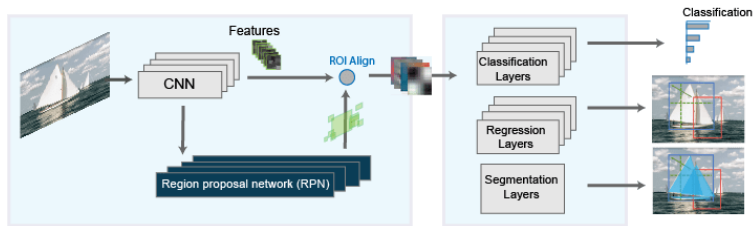
## CNN a Instančná segmentace obrazu



Zdroj : He et al., Mask R-CNN, 2017, Figure 2, <https://arxiv.org/abs/1703.06870>

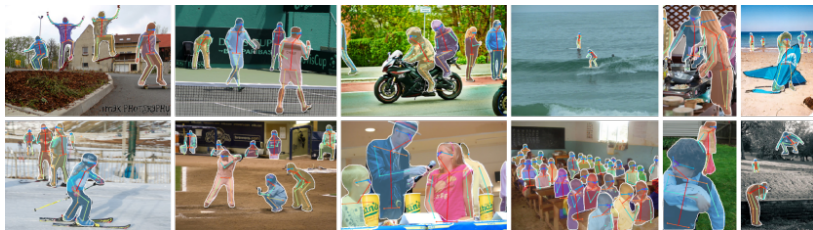
# Aplikace konvolučních neuronových sítí - segmentace obrazu

- rozšíření Faster R-CNN o další větev pro přesnou segmentaci objektu na úrovni pixelů



Zdroj : <https://www.mathworks.com/help/vision/ug/getting-started-with-mask-r-cnn-for-instance-segmentation.html>

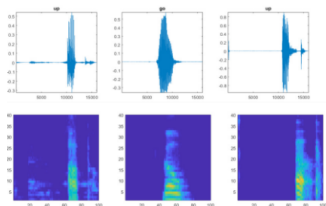
# Aplikace konvolučních neuronových sítí - human pose estimation



Zdroj : He et al., Mask R-CNN, 2017, Figure 7, <https://arxiv.org/abs/1703.06870>

## Další aplikace konvolučních neuronových sítí

- analýza videa - 3D CNN
- sekvenční data
  - audio data (např. převod na obrázek - vstupem je spektrogram)
  - časové řady - učení na historických datech, 1D konvoluce
  - zpracování přirozeného jazyka (např. je vstupem embedding matice)



Top: original audio signals. Bottom: corresponding spectrograms.

Zdroj : <https://www.mathworks.com/content/dam/mathworks/ebook/gated/deep-learning-practical-examples-ebook.pdf>