

Co jsme probírali minule

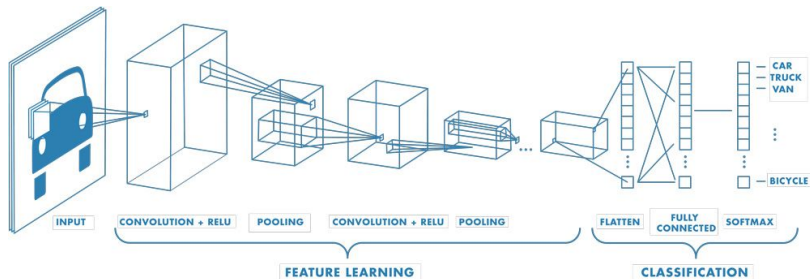
Úvod do konvolučních neuronových sítí

- 1 Operace konvoluce
- 2 Konvoluční vrstva
- 3 Architektura konvoluční neuronové sítě
- 4 Znamé modely konvolučních neuronových sítí (úvod)

Dnes

- Znamé modely konvolučních neuronových sítí (dokončení)
- Učení konvolučních neuronových sítí: přenesené učení, regularizace, předzpracování dat
- Varianty a aplikace konvolučních neuronových sítí (jen přehled)

Architektura konvoluční neuronové sítě



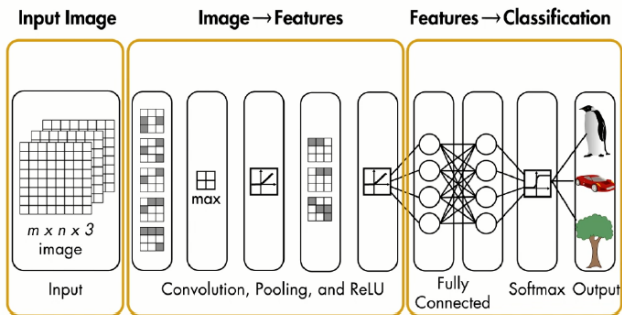
Části konvoluční neuronové sítě

- Konvoluční vrstvy pro extrakci příznaků
- Flattening vrstva - převede data na vektor čísel
- Vrstevnatá neuronová síť pro klasifikaci

Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Architektura konvoluční neuronové sítě



Části konvoluční neuronové sítě

- Konvoluční vrstvy pro extrakci příznaků
- Flattening vrstva - převede data na vektor čísel
- Vrstevnatá neuronová síť pro klasifikaci

Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Učení konvoluční neuronové sítě

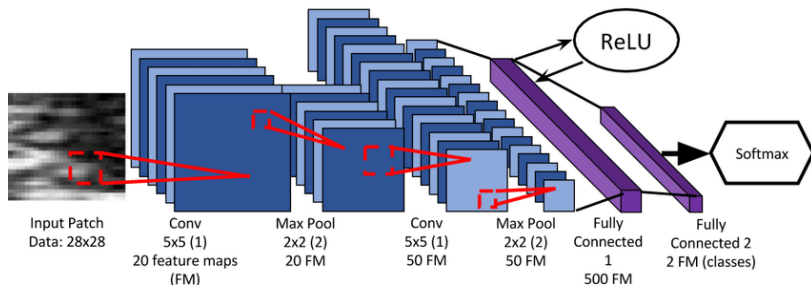
- nějaká varianta algoritmu zpětného šíření (např. SGD)
- mini-batch učení, model potřebuje k naučení větší množství dat
- velké množství parametrů

Jak zvolit vhodnou architekturu v praxi?

- neoptimalizujeme počet vrstev a neuronů
- vybereme z dostupné literatury topologii osvědčenou pro daný typ problému

LeNet 5

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



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.

LeNet 5

Datová sada MNIST

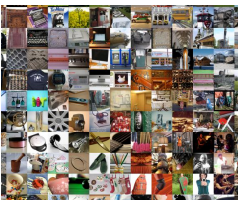
- Oblíbená benchmarková datová sada pro porovnávání různých modelů (přesnost klasifikace)
<http://yann.lecun.com/exdb/mnist/>
- 60000 označovaných trénovacích obrázků ručně psaných číslic, 10000 testovacích (od jiných lidí)
- obrázky 28x28, vycentrované a normalizované na danou velikost
- LeNet 5 dosáhla na MNIST chyby 0,8%

0 0 0 0 0 0 0 0 0 0
 1 1 1 1 1 1 1 1 1 1
 2 2 2 2 2 2 2 2 2 2
 3 3 3 3 3 3 3 3 3 3
 4 4 4 4 4 4 4 4 4 4
 5 5 5 5 5 5 5 5 5 5
 6 6 6 6 6 6 6 6 6 6
 7 7 7 7 7 7 7 7 7 7
 8 8 8 8 8 8 8 8 8 8

ImageNet Large-Scale Visual Recognition Challenge (ILSVRC, 2010-2017)

Datová sada ImageNet

- 16 milionů barevných obrázků z 20 tisíc kategorií
- soutěž odstartovala boom konvolučních neuronových sítí v rozpoznávání obrazu

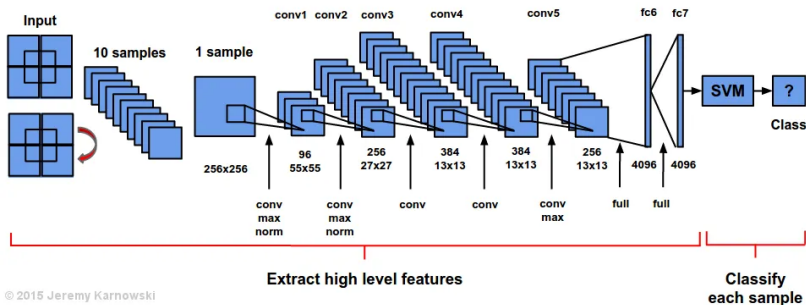


Zdroj obrázku:

https://cs.stanford.edu/people/karpathy/cnnembed/cnn_embed_full_1k.jpg

AlexNet

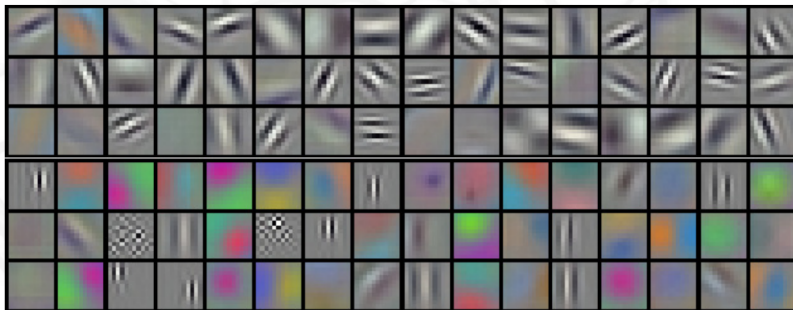
- Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton (vítěz soutěže ILSVRC 2012 - úspěšnost 84,7 % v top-5),
- již výrazně složitější (61M parametrů, 2 GPU 5-6 dnů)



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

AlexNet

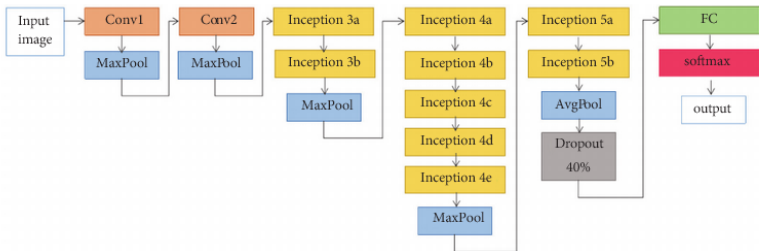
- 96 filtrů $11 \times 11 \times 3$ v první konvoluční vrstvě u AlexNet



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

GoogLeNet (Inception v1), 2014

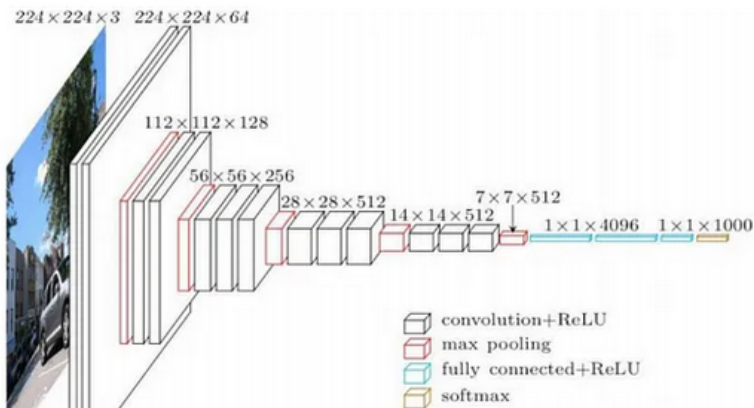
- vítěz soutěže ILSVRC 2014 - úspěšnost 93,33 % v top-5



Zdroj obrázku: Zhang, Keke et. al.: Can Deep Learning Identify Tomato Leaf Disease?. Advances in Multimedia. 2018.

VGGNet

Karen Simonyan a Andrew Zisserman, 2014, rodina modelů (např. VGG16, VGG19)



Další populární architektury

- ResNet (Residual Network, 2015) - zavádí skip connections (řeší problém tiché pošty), vítěz ILSVRC 2015
- Inception v2, v3 (2015, 2016)
- DenseNet: (Gao Huang, 2016)
- MobileNet (Google, 2017), EfficientNet (2019)- úspora výpočetních prostředků
- NASNet (Neural Architecture Search Network, 2017) - model se naučil vhodnou architekturu sám, genetické algoritmy, reinforcement learning
- EfficientNet, 2019
- SqueezeNet
- ...

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, dataseta populární pro rychlé testování a porovnání modelů

COCO (Common Objects in Context)

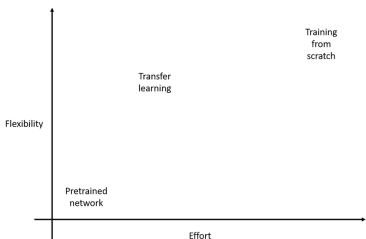
- detekce objektů, segmentace a popis obrázků

COCO 2020 Panoptic Segmentation Task



Možnosti vytvoření a učení konvolučních neuronových sítí

- učení od začátku (training from scratch)
- použití již naučeného modelu (pretrained network)
- doučení již naučeného modelu (finetuning)
- přenesené učení (transfer learning)
 - rychlé učení v rozumném čase

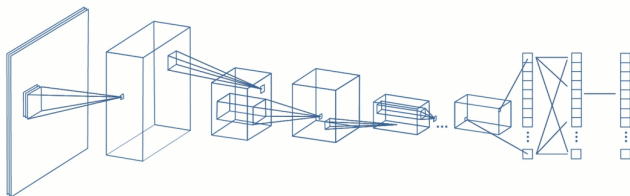


Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Přenesené učení (transfer learning)

- naučený model klasifikuje do 1000 tříd:

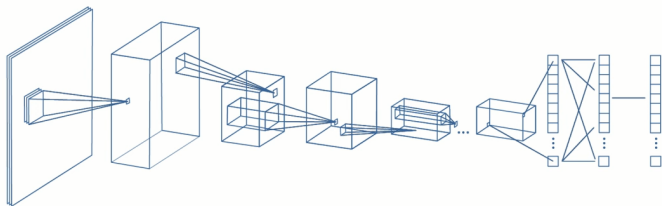


Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Přenesené učení (transfer learning)

- je třeba klasifikovat do trochu jiných tříd:

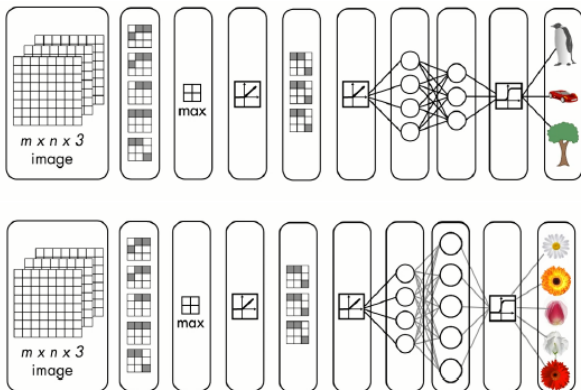


Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Přenesené učení (transfer learning)

- nahradíme klasifikační část neuronové sítě a naučíme ji na nových datech (váhy předchozích vrstev zafixujeme)

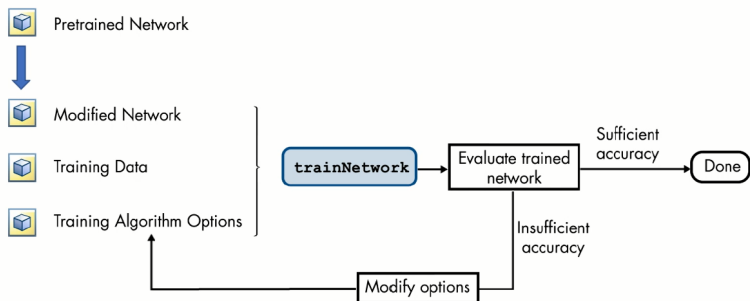


Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Přenesené učení (transfer learning)

Proces přeneseného učení je iterativní



Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Přenesené učení (transfer learning)

Lze využít celou řadu modelů předučených na ImageNet (popř. na jiných datových sadách, např. audio záznamy)

- AlexNet, 224x224x3
- GoogleNet, 227x227x3
- ResNet, 227x227x3
- ...

(podívejte se na další modely dostupné v Deep Network Designeru)

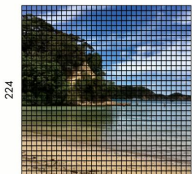
Přenesené učení a předpracování dat

- data je třeba převést na požadovanou velikost (není možné měnit první vrstvu sítě!)

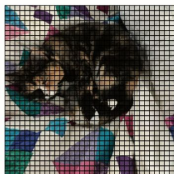


Přenesené učení a předpracování dat

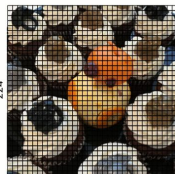
- data je třeba převést na požadovanou velikost (není možné měnit první vrstvu sítě!)



224



224



224

Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Přenesené učení a předpracování dat

- černobílé obrázky je třeba převést na 3 kanály

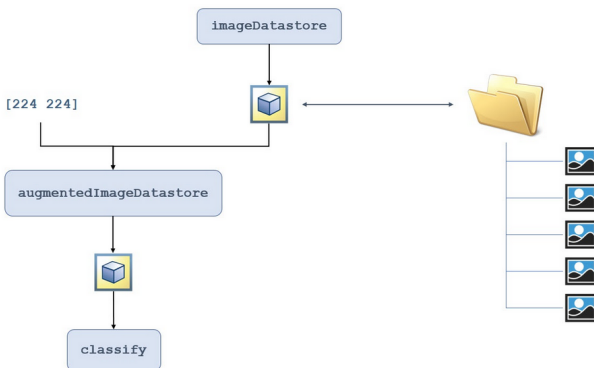


Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Přenesené učení a předpracování dat

- v Matlabu se o základní předzpracování dat postará rozšířený DataStore



Zdroj :

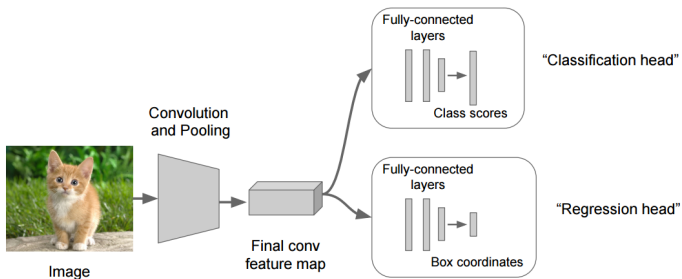
<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Přenesené učení a finetining

- model naučený s pomocí přeneseného učení můžeme dále vylepšit pomocí finetuning-u
 - 1 nejprve aplikujeme přenesené učení (váhy mimo klasifikační část zafixujeme)
 - 2 pak aplikujeme finetuning na celý model (s velmi malým parametrem učení - ideálně začneme s hodnotou, na které končil původní model)
 - 3 můžeme použít regularizaci

Přenesené učení - rozšíření

- klasifikační část neuronové sítě (tj. „hlavu“) můžeme nahradit jinou hlavou a použít naučené příznaky pro řešení jiné úlohy nad stejnými (nebo jinými) daty
 - jiná klasifikační úloha
 - regresní úloha
 - detekce objektů
 - ...



Přenesené učení - omezení

- přenesené učení se hodí, pokud chce naučit model na podobných datech (obecný model → konkrétnější model)
- nelze použít, pokud by bylo třeba změnit vstupní vrstvy
 - větší počet kanálů (např. rentgenový snímek apod.)
 - rozdílná úroveň abstrakce
 - úplně jiná data
- někdy nezbyvá než učit model od začátku

příklad: fotografie ramene rozšířená o rentgen a MRI



<https://radiopaedia.org/cases/normal-shoulder-mri>

Konvoluční neuronová síť a regularizace

- konvoluční sítě mohou být náchylné k přeučení (obzvlášť pokud použijeme přenesené učení a nebo finetuning a máme k dispozici jen poměrně malé množství dat)
- učení může být dost pomalé

→ regularizační techniky

Konvoluční neuronová síť a regularizace

Regularizační techniky rozšiřující model

- early stopping
- L2-regularizace

$$E_n = \beta E + (1 - \beta) \frac{1}{2} \|\vec{w}\|_2^2 = \beta E + (1 - \beta) \sum_i w_i^2$$

- dropout (zhuštěný ensemble learning)
- normalizace

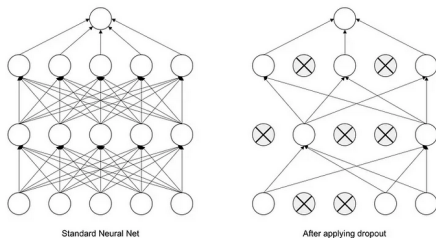
Regularizace a data

- augmentace (rozšíření) dat

Konvoluční neuronová síť a regularizace

Dropout

- spočívá v náhodném vypínání (deaktivování) některých skrytých neuronů během učení
- při testování jsou všechny neurony aktivované
- má smysl u plně propojených vrstev
- přidává se speciální dropout vrstva

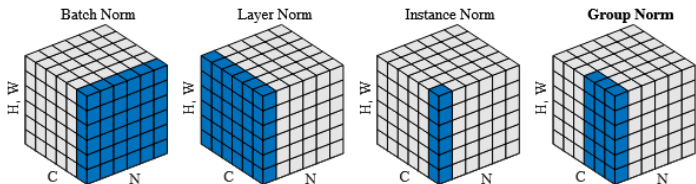


Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting", JMLR 2014

Konvoluční neuronová síť a regularizace

Normalizace vstupů jednotlivých vrstev

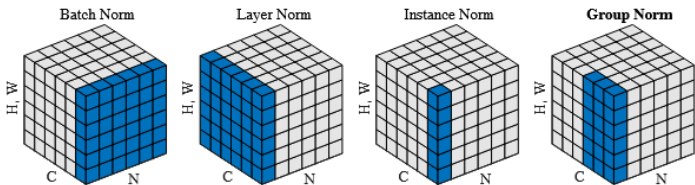
- snaha zafixovat střední hodnoty a rozptyly vstupů vrstvy
- snaha o řešení problému mizejících gradientů
- N ... batch (počet FM), C ... channels (kanály), H, W ... rozměry FM (matice příznaků)
- různé varianty:



Konvoluční neuronová síť a regularizace

Normalizace vstupů jednotlivých vrstev

- implementováno pomocí přidání další vrstvy (např. zakonvoluční vrstvy)
- rychlejší učení, menší citlivost na inicializaci vah
- obdobná technika: **weight normalization**
- robustnost k šumu v datech (nahradí Dropout)
- nevhodné pro rekurentní sítě



Konvoluční neuronová síť a regularizace

Augmentace dat

- různé (náhodné) transformace obrazu (rotace, posun, zrcadlení, zešikmení, měna rozlišení, změna jasu a kontrastu, oříznutí, přidání šumu, blur, kombinace)

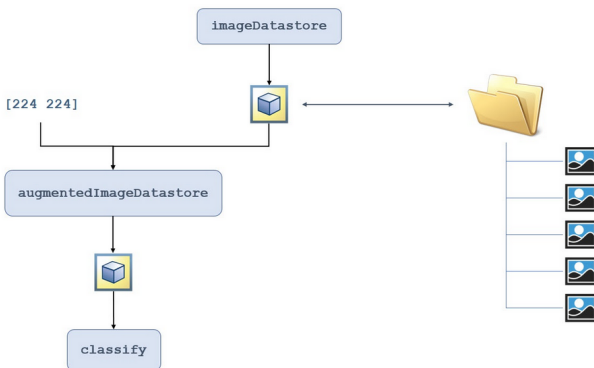


Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Augmentace dat

- v Matlabu se o základní předzpracování dat postará rozšířený DataStore

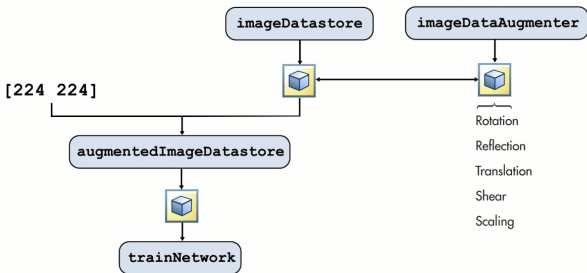


Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Augmentace dat

- v Matlabu se o základní předzpracování dat postará rozšířený DataStore

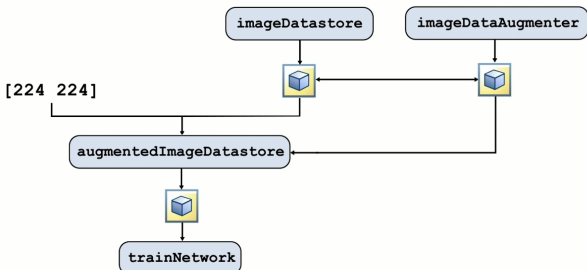


Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Augmentace dat

- v Matlabu se o základní předzpracování dat postará rozšířený DataStore



Zdroj :

<https://matlabacademy.mathworks.com/details/deep-learning-onramp/deeplearning>

Výhody a nevýhody konvolučních neuronových sítí

- ušité na míru datům, která jsou uspořádaná do mřížky
- invariance vůči posunutí, velikosti, odstínu,...
- robustní k šumu v datech
- výpočetně náročné učení, vyžaduje velké množství dat, GPU
- nebezpečí přeučení → regularizace
- adversarial examples

Konvoluční neuronové sítě - implementační detaily

- obvykle stačí velikost filtru 3×3
- často se kombinuje 2×2 max-pooling spolu s tím, že se v každé vrstvě zdvojnásobuje počet kanálů
- pokud použijeme více konvolučních vrstev, nepotřebujeme větší počet plně propojených vrstev
- mezi regularizačními technikami je nejoblíbenější batch normalizace spolu s (mírnou) L2-regularizací na konvolučních vrstvách

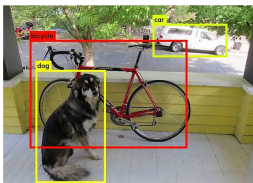
Aplikace konvolučních neuronových sítí

- klasifikace obrázků
- detekce objektů - místo klasifikační části detekční hlava
- segmentace obrazu
- analýza videa
- sekvenční data (časové řady - 1D konvoluce, audio data, přirozený jazyk)

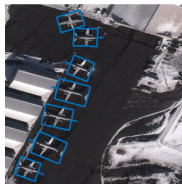
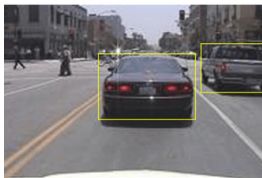
Příklad regrese: predikce sklonu číslic



Aplikace konvolučních neuronových sítí - detekce objektů



Source: Iliasz by Mathias Huberman



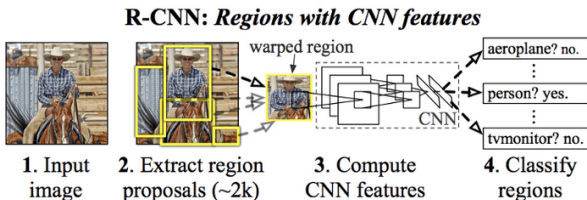
Zdroj : <https://matlabacademy.mathworks.com>

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

Aplikace konvolučních neuronových sítí - detekce objektů

R-CNN (Region-based Convolutional Neural Network)

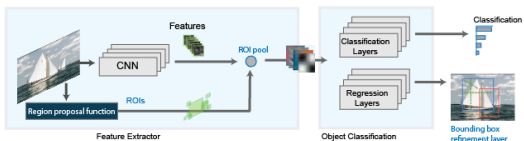
- 1 obrázek je rozdělen na různé oblasti (ROI) pomocí algoritmu selektivního hledání
- 2 každá oblast je zvětšena na stejnou velikost a předložena předučené konvoluční neuronové síti (VGG-16 v původním článku).
- 3 Nakonec jsou použity dvě hlavy: klasifikátor pro klasifikaci každé oblasti a regresor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



Aplikace konvolučních neuronových sítí - detekce objektů

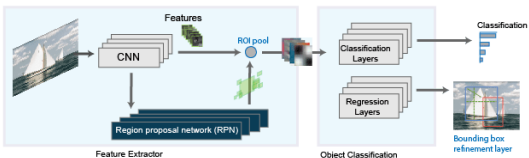
varianta: fast R-RCNN

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

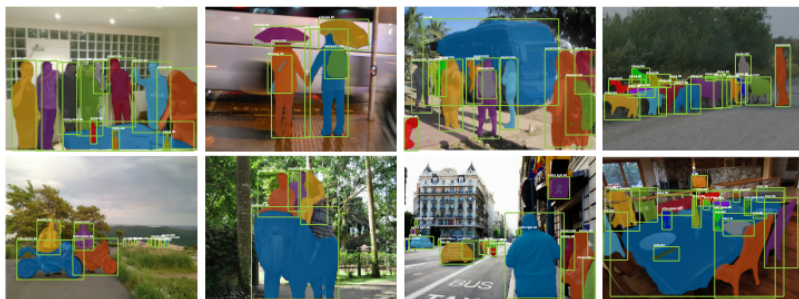


varianta: faster R-RCNN

- automatické rozpoznávání ROI uvnitř modelu



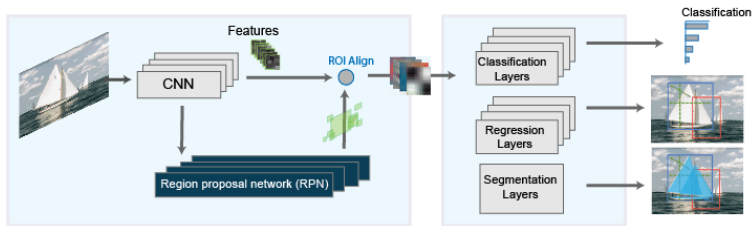
Aplikace konvolučních neuronových sítí - segmentace obrazu



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

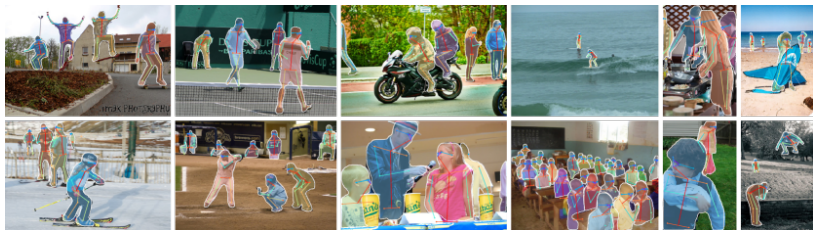
Aplikace konvolučních neuronových sítí - segmentace obrazu

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



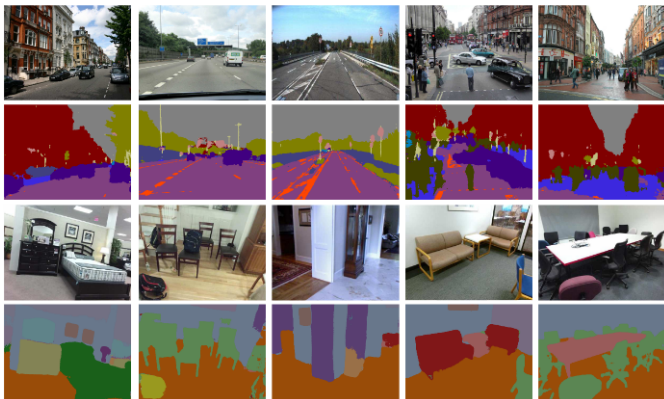
Zdroj : <https://www.mathworks.com/help/vision/ug/getting-started-with-mask-r-cnn-for-instance-segmentation.html>

Aplikace konvolučních neuronových sítí - human pose estimation



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

Aplikace konvolučních neuronových sítí - sémantická segmentace

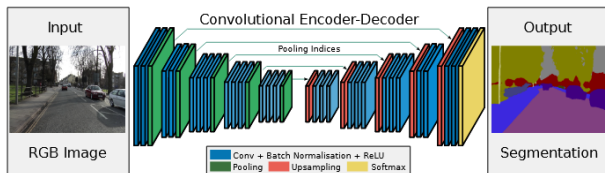


Zdroj : SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation <https://arxiv.org/pdf/1511.00561>

Aplikace konvolučních neuronových sítí - sémantická segmentace

SegNet

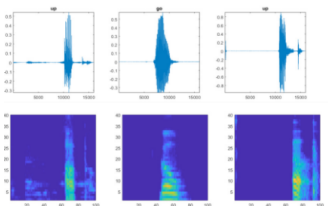
- enkodér-dekodér architektura
- enkodér: klasická CNN
- dekodér: obnovuje obrázek do původních rozměrů:
transponované konvoluční vrstvy a unpooling vrstvy



Zdroj : SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation <https://arxiv.org/pdf/1511.00561>

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>

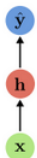
Zajímavé odkazy

- vizualizace konvoluční neuronové sítě
<https://poloclub.github.io/cnnexplainer/>
- Matlab onramp kurzy k hlubokému učení
<https://matlabacademy.mathworks.com/#ai>
<https://matlabacademy.mathworks.com/details/deep-learning-with-matlab/mldl>

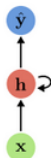
Rekurentní neuronová síť a zpracování posloupností

Jednoduchá rekurentní síť

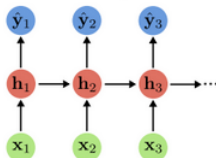
Feedforward
Neural Network



Recurrent Neural Network (RNN)
with feedback connection



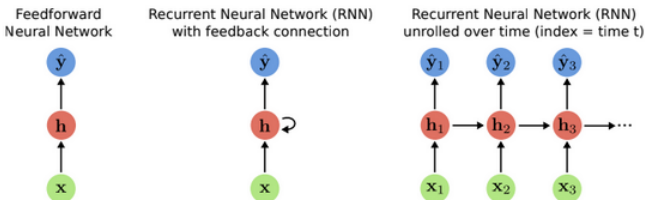
Recurrent Neural Network (RNN)
unrolled over time (index = time t)



- myšlenka: neuron si udržuje svůj vnitřní stav: $h_t = f_h(h_{t-1}, x_t)$
- výstup neuronu: $y_t = f_y(h_t)$
- je možné zpracovat posloupnosti libovolné délky
- f_y, f_h se v čase nemění

Rekurentní neuronová síť a zpracování posloupností

Jednoduchá rekurentní síť



- vnitřní stav: $h_t = \tanh(W_h h_{t-1} + W_x x_t + b_h)$
- výstup neuronu: $y_t = W_y h_t + b_y$
- obvyklá přenosová funkce je \tanh
- váhy a prahy W, b se v čase nemění

Rekurentní neuronová síť a zpracování posloupností

Typy úloh nad sekvenčními daty

- one-to-one - např. obyčejná klasifikace
- many-to-one - např. klasifikace sentimentu, rozpoznání akce na videu
- one-to-many - např. vygenerovat slovní popis k obrázku
- many-to-many - strojový překlad, object tracking

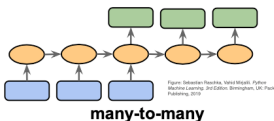
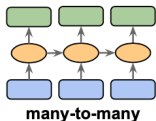
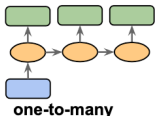
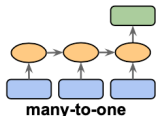
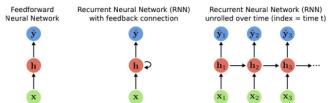


Figure: Sebastian Raschka, "Natural Language Processing with Python: Analyzing Text with the Open Source Tools NLTK, Gensim, and Spacy", Packt Publishing, 2014.

Rekurentní neuronová síť a zpracování posloupností

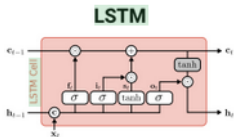
Algoritmus učení: backpropagation through time (BPTT)



- pro celou sekvenci najednou

Znamé modely RNN

- LSTM (Long Short Term Memory)
- GRU (Gated recurrent unit)



Hochreiter, 1997