

Aplikace konvolučních neuronových sítí  
Klára Petrovičová  
485705

Předmět: PB016 Umělá inteligence I - podzim 2019

## Obsah

Úvod do konvolučních neuronových sítí .....	3
Princip konvolučních neuronových sítí .....	3
Konvoluce .....	3
Konvoluční vrstva .....	4
Aktivační funkce .....	4
Pooling vrstva .....	4
Ztrátová vrstva .....	5
Hyperparametry .....	5
Aplikace konvolučních neuronových sítí .....	5
Objevování nových léků .....	5
Biomarkery stárnutí.....	6
Rozpoznávání přirozené řeči .....	6
Analýza videí.....	7
Rozpoznávání fotek .....	7
Rozpoznávání obličejů.....	8
Dlouhodobé předpovědi .....	8
Zdroje .....	10

## Úvod do konvolučních neuronových sítí

V současné době roste potřeba zpracovávat tzv. „velká data“ a to nejen v oblastech informatiky. Pro zpracovávání objemných dat se hledají nové metody, které pracují rychle a s co největší přesností. Například obrazová data se v poslední době nejvíce zpracovávají pomocí konvoluční neuronové sítě (anglicky convolutional neural network (CNN, nebo ConvNet).

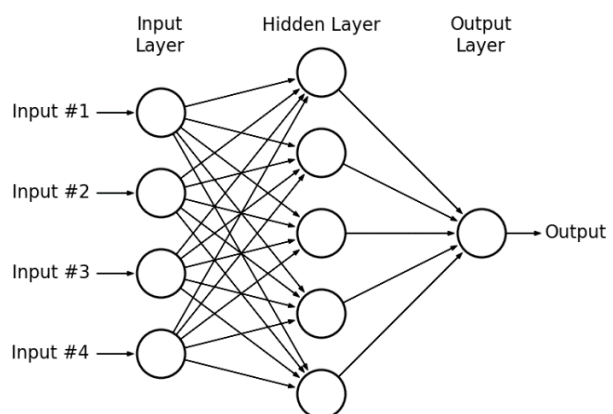
CNN patří mezi hluboké neuronové sítě a pracují nejčastěji s fotografiemi nebo s videem, ale dají se využít i k rozpoznávání přirozené řeči. Usnadňují tak analýzu v oblastech, kde je například potřeba ručně vyplnit nějakou tabulku – rozpoznávání znaků abecedy a písma ručně psaného textu. Pro každou aplikaci poté vznikají specializované modely neuronových sítí. K velkému využití konvolučních neuronových sítí dochází také v průmyslu a to od automobilního – při rozvoji autonomních řízení, až po medicínu – automatické vyhodnocování snímků pacienta.

Nejdříve se v jednoduchosti seznámíme z čeho se skládá konvoluční neuronová síť, abychom rozuměli, jak funguje a jak může být aplikována. Potom si ukážeme konkrétní aplikace s jejich podrobnějším popisem.

## Princip konvolučních neuronových sítí

Dalo by se říct, že CNN, je upravená verze vícevrstvého perceptronu. Ten sestává ze vstupní vrstvy, „skryté“ vrstvy a výstupní vrstvy a je plně propojen, což může mít za následek tak zvaný overfitting, který je způsoben tak, že vytvoříme model, který funguje přesně s určitým trénovacím setem, ale když použijeme nový dataset získáme špatnou přesnost, protože model není optimalizovaný.

Konvoluční neuronové sítě využívají hierarchického schématu na datech tak, že shromažďují více komplexních dat za využití menších a jednodušších schémat.



Obrázek 1: vícevrstvý perceptron

Podobně jako neuronové sítě, vychází i ty konvoluční z biologických procesů a využití propojení mezi neurony, které si tak předávají nějakou informaci. Konvoluční neuronové sítě vychází ze zrakové soustavy savců, v jejichž zrakové mozkové kůře reagují jednotlivé kortikální neurony na stimul v určité přijímající oblasti. Ty se překrývají a tvoří tak zrakové pole.

## Konvoluce

Síť dostane nějaký vstup – tensor (počet fotek, jejich délka, šířka a barevná hloubka = počet bitů k určení barvy jednoho pixelu v bitmapové fotce, nebo počet bitů použit pro každou komponentu barvy u jednoho pixelu). Každý neuron zpracovává data jen z podčásti předchozí vrstvy a tahle další vrstva je menší, než předchozí. Konvoluce kombinuje matice –

hodnoty vrstev se vstupem. Při dopředném průchodu konvoluje každý filtr přes výšku a šířku vstupu a počítá skalární součin mezi filtrem (váhy) a vstupem (jeho pixely).

Tvoří se tak 2D aktivační mapa filtru. Filtry detekují rohy, tvary, barvy, čím jdeme hlouběji, tím větší detaily zkoumáme.

### Váhy

Každý neuron počítá výstupní hodnotu pomocí funkce, která je daná vektorem váhy a biasu (filtry – představují vlastnosti vstupu, např. tvar). Samotné učení potom představují iterativní změny vah a biasu, kdy se po aktivaci filtru detekuje určitá vlastnost s konkrétními parametry vztahenými na vstup. Kdybychom použili vícevrstvý plně propojený perceptron, RGB kanály by představovaly 3 milióny vah.

### Konvoluční vrstva

Konvoluce kombinuje matice – při dopředném průchodu konvoluje každý filtr přes šířku a hloubku, počítá skalární součin mezi filtrem a vstupem a produkuje 2D aktivační mapu filtrů. První vrstva extrahuje nejprimitivnější vlastnosti, jako jsou rohy, barva atd. Vrstvy postupně zjišťují detailnější vlastnosti.

### Aktivační funkce

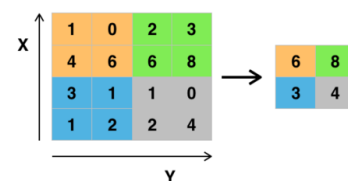
Efektivně odstraňuje negativní prvky z aktivační mapy a nastavuje je na nulu. A to bez toho aniž by ovlivňovala konvoluční vrstvu, nebo aktivační mapu. Používá se v konvoluční vrstvě a snižuje linearitu výstupních informací. Dříve se používala sigmoidea, nebo hyperbolický tangens. ReLU (rectified linear unit) byla představena v roce 2000 a její „objevitel“ Prof. Dr. Richard Hahnloser ji přisuzuje sílu kvůli tomu, že vychází jak z matematiky, tak z biologie. ReLU bere jen kladné hodnoty a ty negativní nastavuje na nulu. Oproti předchozím dvěma typům je také rychlejší, ale má nevýhodu, že neurčuje chování pro negativní vstupy.

$$f(x) = x^+ = \max(0, x)$$

Obrázek 2: funkce ReLU

### Pooling vrstva

Podobně jako konvoluční vrstva, redukuje množství dat kombinováním výstupů z neuronových clusterů v jedné vrstvě do jediného neuronu v další vrstvě a snižuje tak velikost a komplexnost vstupu a to tak, že extrahuje dominantní vlastnosti. Existují různé typy poolingů. nejpoužívanější jsou max pooling (z konvolované části vybírá jen maximální hodnoty) a average pooling (z oblasti vezme průměrnou hodnotu dané oblasti).



Obrázek 3: Max Pooling

### Ztrátová vrstva

Určuje, jak moc se po natrénování liší předpovězený výstup a skutečné labely. Je to poslední vrstva neuronové sítě.

### Hyperparametry

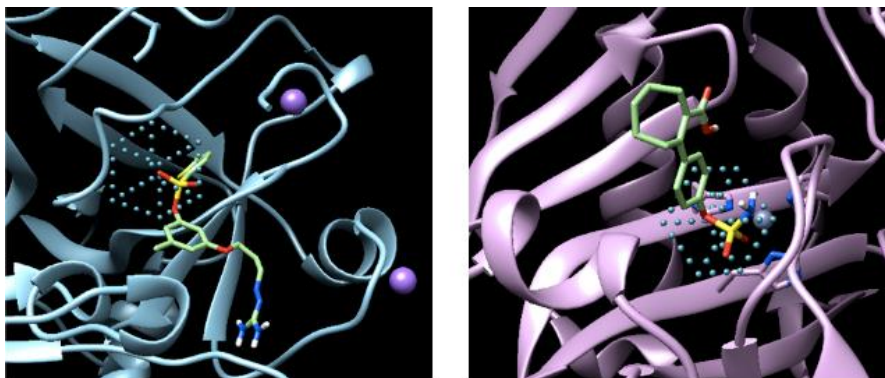
Jejich nastavování ovlivňuje ještě víc výkonost neuronové sítě. Mohou jimi být například **hloubka** – kontroluje počet neuronů ve vrstvě spojujících se do jednoho regionu v určité dimenzi. Například v první vrstvě bere nezměněnou fotku, v další vrstvě se neurony aktivují pod stejným tvarem/barvou, **padding** – kontroluje velikost výstupu - kolik neuronů se vejde do objemu vrstvy, aby byly symetricky rozprostřeny, **stride** – posouvá pixely podle prostorových umístění. Pomocí hodnot těchto parametrů se dá poté vypočítat prostorová velikost výstupu.

## Aplikace konvolučních neuronových sítí

### Vývoj nových léků

První neuronovou sítí na návrh nových medikamentů je AtomNet představený společností Atomwise. AtomNet používá konvoluční neuronové sítě, které předpovídají interakce testovaných molekul a proteinů na základě velkého množství statistických dat, čímž mohou identifikovat potencionální lék. Předchozí přístupy, které nepoužívaly neuronové sítě, byly neefektivní, vývoj takového léku trval třeba 13 let a stál až miliardy dolarů. Bylo nutno testovat tisíce sloučenin a jejich vzájemné reakce. COO Alexander Levy uvádí: "Můžete vzít interakce mezi léky a obrovským biologickým systémem a rozložit je na menší a menší interaktivní skupiny. Pokud prostudujete dostatečné množství příkladů, můžete pak provádět extrémně přesné a rychlé předpovědi." - <http://www.vekstroju.cz/jak-umela-inteligence-pomaha-s-objevovanim-novych-leku>.

Software již pomohl vyvinout léky na Ebolu a roztroušenou sklerózu.



Obrázek 4: AtomNet se učí rozpoznávat sulfonylové skupiny (vyskytující se v antibiotikách)

V Atomwise vyhodnocují 10 milionů snímků každý den. Aktuálními tématy jsou rostoucí odolnost vůči antibiotikům, pandemické viry, Alzheimerova choroba, rakovina a nemoci nervové soustavy. Neuronové sítě umožňují objevovat nejrůznější přístupy a tím, že konvoluční neuronové sítě jdou po vrstvách k větším detailům, jsou vhodné pro tento výzkum při rozpoznávání chemických struktur. AtomNet objevuje znaky, jako jsou aromaticita (vlastnosti určující aromatické uhlovodíky (areny)),  $sp^3$  uhlíky (čtyři jednoduché vazby odpovídající  $sp^3$  hybridní stav uhlíku) a vodíkové vazby.

### Biomarkery stárnutí

National Health and Nutrition Examination Survey (NHANES) je výzkumný program zabývající se zdravotním stavem v průběhu života. Kombinují rozhovory, fyzické pokusy a laboratorní testy dětí a dospělých a zkoumají vliv stravy a životního stylu na celkovou životosprávu.

Mimo to zkoumají faktory jako například bydliště, původ, rasu, stravu a fyziologické parametry. Analýzou těchto dat potom předpovídají riziko vzniku nemocí u jedinců a skupin. Mohou tak také stanovit standardy, co se týče průměrné váhy, výšky a krevního tlaku.

Konvoluční neuronové sítě pracují dobře s velkými daty v průběhu času. Mohou tak tedy vyhodnocovat pacientova klinická data a tělesnou aktivitu a dělat tak lepší závěry, než jakých je schopen lékař. Jsou kombinovány s Cox-Gompertz proporčním hazard modelem, což je statistický model. Tato kombinace přístupů je použita pro určování biomarkerů stárnutí a zjišťování vlivů jednotlivých faktorů na stárnutí a smrt.

Odhadnutí věku osoby na základě biomarkerů je další výzva, kterou se výzkum zabývá. Na základě hluboké konvoluční krátkodobé paměti (ConvLSTM) předvídají věk osoby, jak na základě biomarkerů, tak chytrého zařízení, které monitoruje fyzickou aktivitu, tyto data potom porovnávají s NHANES data setem. Tento výzkum také pomáhá s rozvojem mHealth, což je mobilní rozhraní pro měření zdravotního stavu a získávání dat, a to i za pomoci chytrých doplňků jako jsou hodinky.

Tento výzkum již zjistil, že větší riziko cukrovky, nadváhy a zhoršení mentálního stavu je následkem nedostatku sportovní aktivity. -

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6684608/>.

### Rozpoznávání přirozené řeči

Konvoluční neuronové sítě bývají často využívány pro rozpoznávání přirozené řeči. Na toto téma již vznikla řada projektů. Rozvoj všudypřítomných chytrých zařízení ve formě smartphonů, tabletů a notebooků vyvolal potřebu lepších dotazovacích systémů. Na popularitě získávají chytrí asistenti, jakou jsou Siri od firmy Apple, Alexa od Amazonu a Google Assistant od Googlu. Tyto systémy potřebují parsovat informace a získávat z nich důležité informace, které využívají při vyhledávání na internetu. Mapují věty přirozené řeči na jejich hlubší význam, při čemž samozřejmě záleží na kontextu.

Jsou různé typy rozpoznávání významu. Například mělké, které rozpoznává fakta v mluveném projevu a dává jim váhy podle toho, jakou budou hrát roli v celkovém významu. Tento systém využívá například Alexa od Amazonu, která funguje jako Cloud servis, protože správná interpretace je výpočetně náročná. Systém rozebere větu na individuální zvuky a porovnává je s databází, aby se zjistilo, jaké slovo nejvíce odpovídá. Klíčová slova se určují podle jejich smyslu a vah. Například pokud slyší slovo počasí, nebo teplota, otevře aplikaci o počasí. Pokud nám chce něco říct Alexa, postup se opakuje v opačném pořadí.

### Analýza videí

Video je daleko komplexnější, než fotografie, protože se musí klasifikovat více dimenzí. Jeden způsob je vzít čas a prostor jako ekvivalentní dimenze a udělat konvoluci na obou z nich. Druhý způsob je spojit dvě konvoluční neuronové sítě jednu pro prostorový a druhý pro dočasný stream. Výzkumy, které se zabývají analýzou videí pomocí neuronových sítí, často využívají jako data set videa z YouTube.

Analýza videí se využívá například při rozpoznávání tváří v reálném čase z kamerových záznamů. Tato analýza je však výpočetně velmi náročná. NVIDIA na toto téma udělala výzkum. GPU se při trénování hlubokých neuronových sítí používají častěji, než CPU, protože jsou rychlejší a energeticky úspornější. NVIDIA tyto výkony GPU porovnávala při nastavení různých parametrů, využití různých frameworků (Caffe a Intel's Deep Learning Framework). Při výzkumu porovnávali výkony GPU při trénování – kdy použijeme hodně vstupů ve velkých dávkách a při vyhodnocování – natrénovaná síť dostává nové neznámé vstupy, které vyhodnocuje. Zjistili, že GPU jsou stejně výhodné pro oba procesy.

Jedna z knihoven, která se zabývá klasifikací a analýzou videí v reálném čase je OpenCV.

### Rozpoznávání fotek

Konvoluční neuronové sítě jsou nejčastěji využívány pro rozpoznávání a klasifikaci fotografií a obrázků. Neuronové sítě nyní zvládnou rozpoznávání lidských tváří i z nejrůznějších úhlů, nebo jen z částí.

Databáze MNIST (Modified National Institute of Standards and Technology database) je velká databáze ručně psaných čísel, která slouží primárně k trénování nejrůznějších vyhodnocovacích systémů. Ručně psané číslovky jsou od amerických studentů a ze statistického úřadu. Databáze obsahuje 60 000 trénovacích obrázků a 10 000 testovací obrázků. Od zveřejnění data setů vzniklo velké množství vědeckých publikací, kdy se vědecké týmy snaží pracovat na snížení výpočetní chyby. V původní vědecké práci využívali support-vector machine, které měly chybu 0.8 %. Výpočet chyby u MNIST data setu se nyní pohybuje asi okolo 0.23 %. Konvoluční neuronové sítě vykazují lepší výsledek, než někteří lidé.

Type	Classifier	Error rate (%)
Linear classifier	Pairwise linear classifier	7.6 <sup>[9]</sup>
K-Nearest Neighbors	K-NN with non-linear deformation (P2DHMDM)	0.52 <sup>[19]</sup>
Boosted Stumps	Product of stumps on Haar features	0.87 <sup>[20]</sup>
Non-linear classifier	40 PCA + quadratic classifier	3.3 <sup>[9]</sup>
Support-vector machine (SVM)	Virtual SVM, deg-9 poly, 2-pixel jittered	0.56 <sup>[21]</sup>
Deep neural network (DNN)	2-layer 784-800-10	1.6 <sup>[22]</sup>
Convolutional neural network (CNN)	6-layer 784-40-80-500-1000-2000-10	0.31 <sup>[15]</sup>
Convolutional neural network	6-layer 784-50-100-500-1000-10-10	0.27 <sup>[24]</sup>
Convolutional neural network	Committee of 5 CNNs, 6-layer 784-50-100-500-1000-10-10	0.21 <sup>[17][18]</sup>
Deep neural network	2-layer 784-800-10	0.7 <sup>[22]</sup>
Deep neural network	6-layer 784-2500-2000-1500-1000-500-10	0.35 <sup>[23]</sup>
Convolutional neural network	Committee of 35 CNNs, 1-20-P-40-P-150-10	0.23 <sup>[8]</sup>

Obrázek 5: ML metody a jejich výpočetní chyba

Další známá konvoluční neuronová síť se nazývá AlexNet. Tato síť vyhrála v roce 2012 ImageNet Large Scale Visual Recognition Challenge. ImageNet je obrovská databáze navržena speciálně pro rozpoznávání a klasifikaci objektů. Obsahuje 20 000 kategorií, ve kterých je několik stovek obrázků. V roce 2015 vyhrála síť AlexNet soutěž very deep CNN od Microsoftu a měla 100 vrstev. V roce 2018 udělala společnost ImageNet novou výzvu a to rozpoznávání 3D objektů. Pokrok v této oblasti by mohl zlepšit navigaci autonomních aut.

Jednu z nejlepších přesností má GoogLeNet – s chybou 0.06656. Síť má přes 30 vrstev a její přesnost se hodně blíží lidské. Ale i ty nejlepší algoritmy mají problém rozpoznat malé objekty a velké detaily jako jsou například květiny. Ale oproti lidem jsou lepší v rozpoznávání druhů a ras.

### Rozpoznávání obličejů

S rozpoznávání lidských tváří je spojen zajímavý NVIDIA projekt, který vytváří fotografie lidí, kteří neexistují. Fotky jsou generovány na stránce [thispersondoesnotexist.com](http://thispersondoesnotexist.com). Tyhle projekty často otevírají diskuze, jak se dají neuronové sítě zneužít.

Další kontroverzní využití našli v Číně v jejich systému sociálního kreditu, kdy pomocí kamerového systému sledují občany a podle jejich chování jim strhávají přidělené body.

Kamery rozpoznají člověka, který například přechází na červenou na semaforu, najdou jej v data setu a strhnou mu body. To, jak systém funguje a jestli vůbec není veřejně známo.

Nicméně rozpoznávání lidí kamerovými systémy může sloužit k dopadení zločinců, nebo sledování podezřelých aktivit v prostorech jako jsou například letiště.

I v této oblasti se často využívá knihovna OpenCV.

### Dlouhodobé předpovědi

Do nedávné doby byly na dlouhodobé předpovědi používány rekurentní neuronové sítě, ale ukázalo se, že konvoluční fungují na vyhodnocování ještě lépe. Jejich architektura je založena na tom, že pracují s předáváním informací po vrstvách a to bez ztráty rozlišení.



Učí se z dlouhého časového úseku a pracují s velkým množstvím vzorů, ze kterých pak mohou vyvodit závěr na základě zkušenosti s podobnými daty.

### Detekce a předpověď podvodů

Nezákonné chování představuje pro telefonní společnosti problém, protože zhoršuje kvalitu telekomunikačních služeb. Tyto aktivity se často vyskytují například v online bankovníctví, e-komerci a dalších finančních operacích. U telefonních společností je to například podvádění při platbě telefonního kreditu.

Společnosti potřebují efektivní algoritmy, které nebezpečné chování předpoví. V nedávné studii od Neural Technologie bylo odhadnuto, že společnosti ztratí díky těmto nezákonným činnostem v průměru \$249 miliard dolarů.

Nové technologie (<https://www.sciencedirect.com/science/article/pii/S1877050918301182>) na detekci podvodů kombinují využití konvolučních neuronových sítí v oblastech jako je rozpoznávání objektů, fotek, rozpoznávání přirozené řeči atd. Používají se hluboké vícevrstvé sítě, které extrahují prvky z dat. Podobně jako u všech předchozích příkladů využití neuronových sítí, pracují i tyto se vzory - chování uživatele a detekují tak zvláštní chování. U zákazníka zaznamenávají průměrný počet hovorů za den, platby, kredit, četnost SMS zpráv atd. čímž generují uživatelský profil. Pokud jeho chování začne vybočovat, mohou detekovat případný podvod a to ještě dřív, než k němu dojde.

## Zdroje

<http://www.vekstroju.cz/umela-inteligence-pomaha-proti-zlocinu>

<https://www.atomwise.com/2015/12/02/introducing-atomnet-drug-design-with-convolutional-neural-networks/>

<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5980076/>

<https://www.quora.com/How-does-Alexa-work>

[https://www.nvidia.com/content/tegra/embedded-systems/pdf/jetson\\_tx1\\_whitepaper.pdf](https://www.nvidia.com/content/tegra/embedded-systems/pdf/jetson_tx1_whitepaper.pdf)

<https://analyticsindiamag.com/why-convolutional-neural-networks-are-the-go-to-models-in-deep-learning/>

<https://www.sciencedirect.com/science/article/pii/S1877050918301182>

<https://arxiv.org/pdf/1404.7296.pdf>

obrázek 1 - [https://www.researchgate.net/figure/A-hypothetical-example-of-Multilayer-Perceptron-Network\\_fig4\\_303875065](https://www.researchgate.net/figure/A-hypothetical-example-of-Multilayer-Perceptron-Network_fig4_303875065)

obrázek 2 - [https://en.wikipedia.org/wiki/Rectifier\\_\(neural\\_networks\)](https://en.wikipedia.org/wiki/Rectifier_(neural_networks))

obrázek 3 - [https://en.wikipedia.org/wiki/Convolutional\\_neural\\_network](https://en.wikipedia.org/wiki/Convolutional_neural_network)

obrázek 4 - <https://www.atomwise.com/2015/12/02/introducing-atomnet-drug-design-with-convolutional-neural-networks/>

obrázek 5 - [https://en.wikipedia.org/wiki/MNIST\\_database](https://en.wikipedia.org/wiki/MNIST_database)