**ŽILINSKÁ UNIVERZITA V ŽILINE**

FAKULTA RIADENIA A INFORMATIKY

DIPLOMOVÁ

PRÁCA

BRANISLAV TOMKA

**Implementácia konvolučnej hlbokej neurónovej siete pre diagnostiku rakoviny kože**

Vedúci práce: Ing. Lukáš Falát, PhD.

Registračné číslo: 988/2019

Žilina, 2020

**ŽILINSKÁ UNIVERZITA V ŽILINE**

FAKULTA RIADENIA A INFORMATIKY

BAKALÁRSKA

PRÁCA

ŠTUDIJNÝ ODBOR:

Inteligentné Informačné Systémy

BRANISLAV TOMKA

**Implementácia konvolučnej hlbokej neurónovej siete pre diagnostiku rakoviny kože**

Žilinská univerzita v Žiline

Fakulta Riadenia a Informatiky

Školiace pracovisko KMME

Žilina, 2020

#### [Čestné vyhlásenie](Pokyny_pre_vypracovanie_ZP.docx#Čestné_vyhlásenie)

Čestne prehlasujem, že som prácu vypracoval samostatne s využitím dostupnej literatúry a vlastných vedomostí. Všetky zdroje použité v diplomovej práci som uviedol v súlade s predpismi.

Súhlasím so zverejnením práce a jej výsledkov.

...........................................

V Žiline, dňa ............ Branislav Tomka

#### [Poďakovanie](Pokyny_pre_vypracovanie_ZP.docx#Poďakovanie)

„*Moje poďakovanie patrí školiteľovi diplomovej práce Ing. Lukáš Falát, PhD. za odbornú pomoc, pripomienky a usmerňovanie pri tvorbe práce.*“

[ABSTRAKT V ŠTÁTNOM JAZYKU](Pokyny_pre_vypracovanie_ZP.docx" \l "Abstrakt)

TOMKA, Branislav: ***Implementácia konvolučnej hlbokej neurónovej siete pre diagnostiku rakoviny kože***. [Diplomová práca]. – Žilinská univerzita. Fakulta Riadenia a Informatiky; Katedra makro a mikroekonomiky. – Školiteľ/Vedúci: Ing. Lukáš Falát, PhD. – Stupeň odbornej kvalifikácie: inžinier. – Žilina: FRI UNIZA, 2020. Počet strán .

[**Kľúčové slová**](Pokyny_pre_vypracovanie_ZP.docx#Kľúčové_slová)**:**

[ABSTRAKT V CUDZOM JAZYKU](Pokyny_pre_vypracovanie_ZP.docx" \l "Abstrakt_CJ)

**Key words**

[Obsah](Pokyny_pre_vypracovanie_ZP.docx#Obsah)

Zoznam obrázkov 8

Zoznam tabuliek 9

Zoznam skratiek 10

Úvod 11

1 Rakovina kože 12

1.1 Štatistiky 12

1.2 Prevencia 13

1.2.1 Metóda ABCDE 13

Zoznam použitej literatúry 15

Zoznam príloh 16

Prílohy 17

# [Zoznam obrázkov](Pokyny_pre_vypracovanie_ZP.docx" \l "Zoznam_obrázkov)

# [Zoznam tabuliek](Pokyny_pre_vypracovanie_ZP.docx" \l "Zoznam_obrázkov)

# [Zoznam skratiek](Pokyny_pre_vypracovanie_ZP.docx" \l "Zoznam_skratiek)

# [Úvod](Pokyny_pre_vypracovanie_ZP.docx#Úvod)

# Rakovina kože

Rakovina kože sa môže vyskytnúť na hociktorom mieste na tele ale najčastejšie na koži, ktorá je často vystavená slnečnému svetlu ako napríklad tvár, krk a ruky. Hlavné typy rakoviny kože sú: skvamocelulárny karcinóm, bazocelulárny karcinóm a melanóm. Melanóm je menej častý typ rakoviny ako zvyšné dva (non-melanómové), ale má väčšiu pravdepodobnosť, že napadne okolité tkanivo a že sa rozšíri do ostatných častí tela. Najviac smrtí je spôsobených melanómom. Melanóm je choroba, pri ktorej sa v kožných bunkách vytvoria zhubné bunky.

## Štatistiky

Rakovina kože je jedna z najbežnejších typov rakoviny. Melanóm tvorí len 4% všetkých nádorov kože, avšak spôsobuje až 73% všetkých úmrtí na rakovinu kože. Približne 2,3 % mužov a žien majú diagnostikovaný melanóm kože počas života. Za rok 2019 v Spojených Štátoch Amerických pribudlo 96 480 nových prípadov s melanómom kože a 7 230 ľudí zomrelo.

|  | **Typy rakoviny** | **Odhad nových prípadov 2019** | **Odhad úmrtí 2019** |
| --- | --- | --- | --- |
| 1. | Rakovina prsníka (Ženy) | 268,600 | 41,760 |
| 2. | Rakovina pľúc a priedušiek | 228,150 | 142,670 |
| 3. | Rakovina prostaty | 174,650 | 31,620 |
| 4. | Rakovina hrubého čreva a konečníka | 145,600 | 51,020 |
| **5.** | **Melanóm kože** | **96,480** | **7,230** |
| 6. | Rakovina močového mechúra | 80,470 | 17,670 |
| 7. | Rakovina lymfatických uzlín | 74,200 | 19,970 |
| 8. | Rakovina obličiek | 73,820 | 14,770 |
| 9. | Rakovina krčka maternice | 61,880 | 12,160 |
| 10. | Leukémia | 61,780 | 22,840 |

Ročne pribudne 22,2 nových prípadov melanómu na 100 000 ľudí podľa štatistík z rokov 2012-2016. Melanóm kože sa viac vyskytuje u mužov ako u žien. U mužov je to 28.8 a u žien 17.5 nových prípadov ročne na 100 000 ľudí. Prípadov melanómu kože sa vyskytuje pri bielej rase viac ako pri ktorejkoľvek inej rase. Najčastejšie sa melanóm kože diagnostikuje osobám v rokoch 65-74 (24,1%). Percento úmrtnosti je najväčšie pri ľuďoch v rokoch 75-84 (23,9% z všetkých úmrtí).

## Prevencia

Pravdepodobnosť výskytu rakoviny kože je možné znížiť používaním ochrany proti slnečnému žiareniu a vyhýbaniu sa ultrafialovému žiareniu. Ultrafialové žiarenie je na základe aktuálnych vedeckých poznatkov klasifikované Medzinárodnou agentúrou pre výskum rakoviny (IARC) ako dokázaný karcinogén. Závery epidemiologických štúdií IARC potvrdili, že vystavenie sa UV žiareniu v mladosti najviac prispieva k poškodeniu kože a riziku rozvoja melanómu v ďalšej časti života. V súčasnosti je vo svete zaznamenávaný významný nárast melanómu kože najmä medzi mladými ženami do 30 rokov. Z tohto dôvodu bola v roku 2019 schválená novela zákona o ochrane, podpore a rozvoji verejného zdravia, ktorá sa zaoberala obmedzením osobám mladším ako 18 rokov navštevovať solária s umelým zdrojom UV žiarenia.

Ďalší spôsob ako predísť zlým následkom rakoviny kože je pravidelná kontrola buď u dermatológa, alebo samokontrolou využitím metódy ABCDE.

### Metóda ABCDE

* **A** – Asymetry (Asymetria) - Dermatológa by sme mali navštíviť v prípade, že je naše znamienko asymetrické. To spoznáme tak, že si ho pomyselne rozdelíme na dve polovice a zameriame sa na jeho tvar. Pokiaľ sú obe polovice súmerné, tak nie je treba nič riešiť. Pokiaľ je ale každá polovica iná, mali by sme zbystriť pozornosť.
* **B** – Borders (Okraje) - Okrem plasticity by sme sa mali u znamienok sústrediť aj na ich okraje. Ohraničenie melanómu totiž býva často nepravidelné či rozpité, pričom zdravé znamienko má okraje pravidelné a ostré.
* **C** – Colour (Farba) - O mnohom môže vypovedať aj farba znamienka. Pokiaľ pozorujeme, že niektorá z jeho častí zrazu zmenila farbu, mali by sme si ho nechať skontrolovať, nepravidelne škvrnité alebo čierne znamienka môžu byť nebezpečné.
* **D** – Diameter (Priemer) - Všímajte si aj veľké pehy a znamienka. Melanómy bývajú zvyčajne väčšie než zdravé znamienka a väčšinou sa ich veľkosť pohybuje nad pol centimetra v priemere.
* **E** – Evolution (Vývoj) - Môžeme povedať, že je potrebné venovať väčšiu pozornosť znamienku vždy, keď  náhle zmení tvar, veľkosť, farbu či štruktúru. V takom prípade je na mieste vyhľadať pomoc kožného lekára, môže sa totiž jednať o nebezpečný melanóm. Dôležitá je pri samovyšetrení aj pravidelnosť. Keď budeme kontrolovať znamienka pravidelne, všimneme si aj drobné odchýlky, ktoré sa na ňom mohli vytvoriť priebehom času.

# Strojové učenie

Strojové učenie je vedecký odbor algoritmov a štatistických modelov, ktorý využívajú počítačové systémy na vykonanie špecifických úloh bez použitia explicitne definovaného postupu a namiesto neho sa spolieha na vzorce a vyhodnotenia. Považuje sa za podmnožinu umelej inteligencie. Algoritmy strojového učenia tvoria matematický model na základe nejakej vzory dát, ktorú označujeme ako „trénovacie dáta“, za účelom robiť predikcie alebo rozhodnutia bez explicitného naprogramovania vykonať danú úlohu.  Mitchell, T. (1997). Machine Learning. McGraw Hill. p. 2. [*ISBN*](https://en.wikipedia.org/wiki/International_Standard_Book_Number) [*978-0-07-042807-2*](https://en.wikipedia.org/wiki/Special:BookSources/978-0-07-042807-2). Mitchell (1997) definuje algoritmy strojového učenia “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.” Teda strojové učenie nám umožňuje vyriešiť úlohy, ktoré sú príliš náročné na vyriešenie fixnými programami napísanými a navrhnutými ľuďmi. Z vedeckého a filozofického pohľadu je strojové učenie zaujímavé pretože rozvoj nášho chápania strojového učenia si vyžaduje rozvíjanie nášho chápania zásad, ktoré sú základom inteligencie.

## Typy algoritmov

Typy algoritmov strojového učenia sa líšia v ich prístupoch, typoch dát na vstupe a výstupe a type úloh alebo problému, ktoré majú riešiť.

### Supervizované učenie

Supervizované učenie začína s cieľom predpovedať známy výstup alebo cieľ. V tomto type algoritmov sa často porovnávajú výkonnosti rôznych modelov na základe výsledkov na rovnakých datasetoch. Medzi typické problémy supervizovaného učenia patrí rozpoznávanie ručne písaného textu, klasifikovanie obrázkov rôznych objektov (napr. je to mačka alebo pes?, je to lietadlo alebo auto?) alebo klasifikovanie dokumentov (napr. je to klinické skúšanie srdcového zlyhania alebo finančná správa?). Ide predovšetkým o úlohy, ktoré môže školený človek urobiť dobre, a preto sa počítač často snaží priblížiť ľudskému výkonu. Supervizované učenie sa zameriava na klasifikáciu, ktorá zahŕňa výber z podskupín, ktorá najlepšie opisuje novú skupinu údajov, alebo predikciu, teda odhad neznámeho parametra (napr. teplota v Žiline zajtra poobede).

### Nesupervizované učenie

Na rozdiel, pri nesupervizovanom učení neexistujú žiadne výstupy, ktoré by sme mali predpovedať. Namiesto toho sa snaží nájsť prirodzene sa vyskytujúce vzorce alebo zoskupenia v rámci daného datasetu. Toto hodnotiť je už zo svojej podstaty náročnejšia úloha a často sa takéto vzorce alebo zoskupenia, ktoré sa získali prostredníctvom nesupervizovaného učenia hodnotia podľa výkonnosti v následných úlohách supervizovaného učenia (napr. sú tieto nové vzorce nejakým spôsobom užitočné?).

### Posilňujúce učenie

Posilňujúce učenie je oblasť strojového učenia zaoberajúca sa tým, ako by softvéroví agenti mali konať v prostredí s cieľom maximalizovať nejaký prínos alebo odmenu. Tento typ učenia sa odlišuje od supervizovaného učenia tým, že nepotrebuje vopred označené vstupy prislúchajúcim výstupom. Namiesto toho sa zameriava na nájdenie rovnováhy medzi skúmaním (nepreskúmaného územia) a využívaním (súčasných poznatkov).

## Modely

Vykonanie strojového učenia zahŕňa vytvorenie modelu, ktorý je natrénovaný na nejakých trénovacích dátach a potom spracúva ďalšie dáta aby robil predikcie. Existuje niekoľko rôznych typov modelov, ktoré sa používajú v strojovom učení.

### Regresná analýza

Regresná analýza zahŕňa veľké množstvo štatistických metód na odhad vzťahu medzi vstupnými premennými a ich vlastnosťami. Najbežnejšia forma regresnej analýzy je lineárna regresia, ktorá vytvára funkciu, ktorá najlepšie vyhovuje daným dátam na základe matematických kritérií, ako napríklad suma najmenších štvorcových chýb (SSE). Toto kritérium sa často rozširuje pomocou regularizačných (matematických) metód na zmiernenie pretrénovania. Pri riešení nelineárnych problémoch sa používa napríklad logistická regresia alebo polynomiálna regresia (používa sa v MS Excel na určenie trendu).

### Rozhodovacie stromy

Učenie rozhodovacieho stromu používa rozhodovací strom ako prediktívny model na prechod z pozorovaní o nejakej položke (v strome predstavuje vetvu) k záverom o cieľovej hodnote položky (v strome predstavuje list). Stromy, v ktorých cieľová premenná môže mať diskrétnu množinu hodnôt, sa nazývajú klasifikačné stromy. V týchto stromových štruktúrach listy reprezentujú označenia triedy a vetvy predstavujú spojenia prvkov, ktoré vedú k týmto označeniam triedy. Stromy, v ktorých cieľová premenná môže mať spojité hodnoty (zvyčajne reálne čísla), sa nazývajú regresné stromy. Pri rozhodovaní sa môže použiť rozhodovací strom na vizualizáciu postupnosti rozhodnutí vedúcich k danému výsledku.

### Genetický algoritmus

Genetický algoritmus je heuristika, ktorá napodobňuje proces prirodzenej selekcie pomocou metód ako mutácia a kríženie, aby sa vytvorili nové genotypy v nádeji, že sa nájdu dobré riešenia daného problému. V 80. a 90. rokoch 20. storočia sa používali genetické algoritmy v strojovom učení. Dnes sa však často kombinujú genetické algoritmy s inými modelmi strojového učenia na dosiahnutie najlepšieho výsledku.

### Umelé neurónové siete

Umelé neurónové siete boli pôvodne vyvinuté na modelovanie činnosti nervového systému mozgu. Dnes už sa však môžu použiť na riešenie veľkého množstva úloh. Umelá neurónová sieť sa rovnako ako mozgová neurónová sieť skladá z neurónov a axónov, teda prepojení týchto neurónov. Každý neurón môže prijať signál a ten následne spracovať a preposlať do ďalších neurónov. Neurónovú sieť teda tvorí veľké množstvo navzájom prepojených neurónov, ktoré si medzi sebou posielajú signál od vstupu až po výstup. Spracovanie signálu v neuróne je reprezentované nelineárnou funkciou súčtu všetkých vstupov neurónu a výsledok je jeden výstup. Učenie umelej neurónovej siete prebieha v nastavovaní váh každého neurónu pre konkrétne vstupy. V dnešnej dobe sa umelé neurónové siete používajú na veľké množstvo úloh ako napríklad rozpoznanie jazyka, prekladanie, filtrovanie na sociálnych sieťach, pri videohrách a taktiež ich veľké využitie je v medicínskej diagnostike.

# Konvolučné neurónové siete

Konvolučná neurónová sieť (KNS) je špeciálny typ neurónovej siete, ktorej hlavné využitie je klasifikovanie obrázkov. Napríklad rozhodovanie či je na obrázku mačka alebo pes, alebo rozpoznanie ručne písaného textu.

## KNS a NS

Konvolučná neurónová sieť je veľmi podobná klasickej neurónovej sieti. Pozostávajú z neurónov, ktoré majú váhy, ktoré sa sieť učí. Každý neurón dostane vstup, vykoná súčet vynásobení váh so vstupmi a prípadne sa aplikuje aktivačná funkcia. Celá sieť stále vyjadruje jednu diferencovateľnú funkciu. Z obrázku na vstupe vyprodukuje skóre jednotlivých tried na výstupe sieti. Avšak konvolučná neurónová sieť predpokladá že má na vstupe obrázok.

Hlavným dôvodom prečo sa na obrázky používajú konvolučné neurónové siete je, že obrázky sú veľké. Väčšina obrázkov, ktoré sa v dnešnej dobe používajú pri rozpoznávaní majú veľkosť 224x224 alebo väčšie. Farebné obrázky navyše bývajú reprezentované 3 farebnými kanálmi (RGB) a teda takýto obrázok má 224 \* 224 \* 3 = 150 528 vstupných bodov/pixelov. Typická skrytá vrstva v takejto sieti má 1024 uzlov, takže by sa neurónová sieť musela učiť 150 528 \* 1024 = 154 140 672 váh. Viac ako 150 miliónov váh len na jednu vrstvu. Takáto sieť by bola obrovská a bolo by takmer nemožné učiť takúto sieť. Navyše pri obrázkoch ani toľko váh nepotrebujeme, pretože ako vieme tak pixely jedného obrázka sú najviac popisujúce v rámci kontextu ich susedov. Všetky objekty obrázkov sú tvorené malými, presne umiestnenými vlastnosťami, ako napríklad kruhová dúhovka oka, alebo štvorcový hranatý roh papiera. Potom učiť sa váhy pre každý jeden pixel môže byť veľmi zbytočné.

Ďalším veľmi podstatným dôvodom nepoužívania klasickej neurónovej siete je fakt, že pozícia objektu sa môže meniť. Ak by bola sieť naučená rozpoznávať psa na obrázku, tak v prípade, že sa pes na obrázku posunie, sieť ho vôbec nemusí nájsť, pretože by sa neaktivovali potrebné neuróny a takáto sieť by bola zbytočná.

Konvolučná neurónová sieť rieši oba tieto problémy .

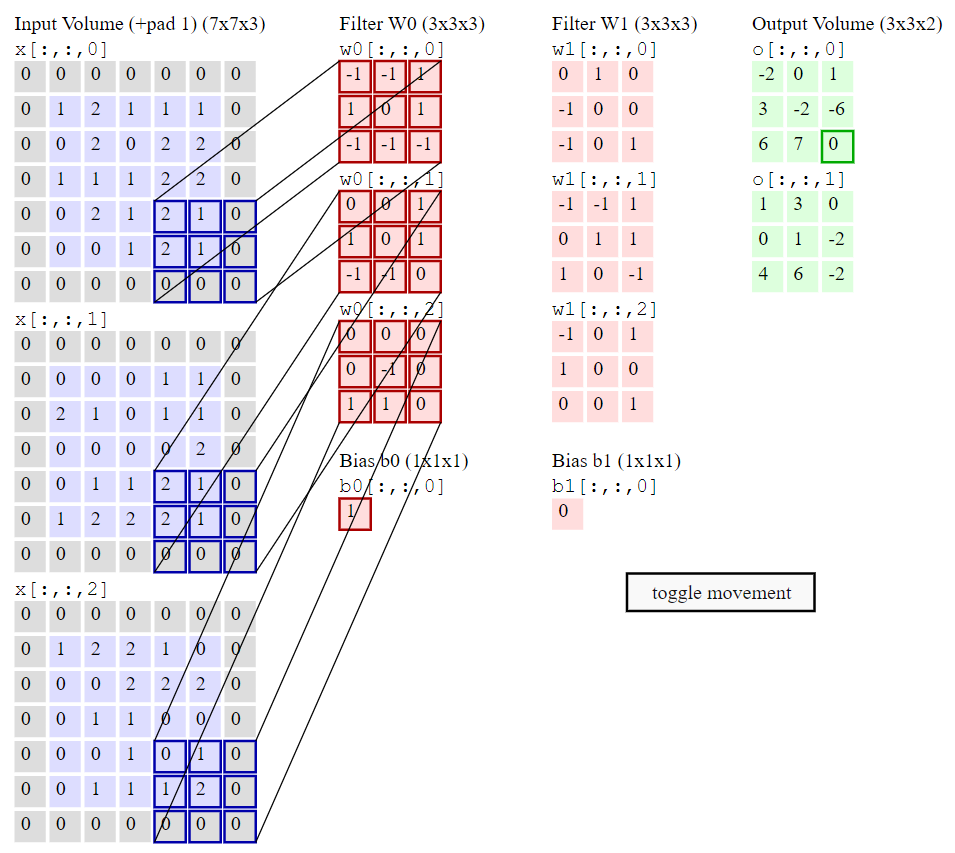
## Čo je KNS?

Konvolučné neurónové siete sú v podstate neurónové siete, ktoré používajú konvolučné vrstvy, ktoré sú založené na matematickej operácií, ktorá sa nazýva konvolúcia.

### Konvolúcia a filter

Hlavným rozdielom konvolučnej vrstvy oproti normálnej vrstve je zdieľanie váh. Každý neurón konvolučnej vrstvy sa pozerá na štvorcový výrez obrázka. Neuróny zoberú na vstup každý pixel z tohto výrezu a zdieľa váhy bez ohľadu na to, kde sa tento výrez nachádza na obrázku, takže sú rovnocenné. Ďalšia interpretácia konvolučnej vrstvy namiesto množiny rovnocenných neurónov, je jeden neurón, ktorý nazývame filter. Tento filter je štvorcová matica, ktorá obsahuje váhy a konvolúciou prechádza celý obrázok výrez po výreze. Konvolúcia pozostáva zo 4 hlavných bodov:

1. Prekrytie filtra cez jeden výrez obrázka.
2. Vykonanie násobenia po prvkoch medzi hodnotami vo filtri a ich zodpovedajúcimi hodnotami na obrázku.
3. Sčítanie všetkých výsledkov násobení po prvkoch a prípadné prevedenie cez aktivačnú funkciu ako pri klasickej vrstve. Výsledné číslo bude jeden pixel vo výstupnom obrázku.
4. Kroky sa opakujú cez všetky výrezy celého obrázka.



Na obrázku je vidieť ako postupuje konvolúcia. Na vstupe je obrázok 7x7x3 s paddingom 1 a je prechádzaný dvoma filtrami o veľkosti 3x3x3 a stride je 2, teda výstup má rozmer 3x3x2.

Filter teda môžeme interpretovať aj ako nejaký detektor vlastností, ktorý vo výsledku vráti veľké číslo, ak detekuje danú vlastnosť.



Na obrázku je 96 filtrov o veľkosti [11x11x3], ktoré sú naučené Krizhevskym. Na niektorých filtroch je jasne vidieť, ktoré vlastnosti obrázka aktivujú (rôzne okraje, tvary, ...).

### Stride

Stride je parameter konvolučnej vrstvy, ktorý určuje veľkosť kroku medzi jednotlivými výrezmi, cez ktoré prechádza filter pri konvolúcií. Takmer vždy sa tento parameter nastavuje na 1, tak aby ďalší výrez bol hneď nasledujúci. Výhoda vyššej hodnoty tohto parametra je, že zníži výšku a šírku výstupu. Avšak v praxi stride 2 alebo viac môže byť považované za príliš agresívny prístup a teda zníženie veľkosti výstupu je za cenu straty niektorých informácií.

### Padding

Pri konvolúcií pixely, ktoré sú viac pri strede sú prejdené filtrom viac krát, ako tie pri kraji obrázka, takže tam môže dôjsť k strate nejakej informácie. Aby sa predišlo takejto strate informácií, je možné nastaviť parameter pooling, ktorý pridá 0 na okraje obrázka aby nebola pridaná žiadna informácia navyše, keďže násobením 0 sa výsledná hodnota nezvýši.

### Pooling

Susedné pixely v obrázkoch zvyčajne majú podobné hodnoty, takže konvolučné vrstvy takisto dajú rovnaké hodnoty pre susedné pixely vo výstupe. Vo výsledku väčšina informácií, ktoré sa nachádzajú vo výstupe konvolučnej vrstve sú redundantné. Napríklad ak použijeme filtre na detekciu hrán a na niektorom mieste nájdeme hranu, je veľmi pravdepodobné, že hranu nájdeme aj o 1 pixel vedľa a teda je to stále tá istá hrana, takže sme nenašli nič nové.

Pooling vrstva rieši tento problém. Táto vrstva redukuje veľkosť vstupu združovaním hodnôt. Obvykle sa na pooling využívajú jednoduché matematické operácie ako min, max alebo priemer. Zvyčajne sa na združujú pixely po výrezoch veľkých 2x2 pixely. Teda šírka a výška výstupného obrázka bude o polovicu menšia ako vstupného obrázka. Napríklad pri max-pooling vyberieme z výrezu 2x2 len jedno najväčšie číslo.

### Aktivačné vrstvy

Pri konvolučných neurónových sieťach sa veľmi často hneď za konvolučnú vrstvu pridáva aktivačná vrstva, ktorá aktivuje výstup konvolučnej vrstvy niektorou z aktivačných funkcií. Zvyčajne sa používa ReLU alebo Tanh. Často pri uvádzaní konvolučnej vrstve v architektúre sa automaticky predpokladá, že súčasťou tejto konvolučnej vrstvy je aj aktivačná vrstva (najmä ReLU).

### Plne prepojená vrstva

Plne prepojená vrstva sa pridáva za konvolučné vrstvy, kde sú všetky neuróny prepojené so všetkými výstupmi predošlej vrstvy. Výstup z poslednej konvolučnej vrstvy sa roztiahne do jedného vektora, ktorý sa použije ako vstup pre túto vrstvu. Týchto plne prepojených vrstiev sa môže v KNS nachádzať aj viacej. Tieto vrstvy pracujú ako klasická neurónová sieť. Prenásobia vstupy s váhami, sčítajú ich a aplikujú na tento súčet aktivačnú funkciu.

### Softmax

Ako posledná, výstupná, vrstva KNS sa takmer vždy používa Softmax vrstva. Táto vrstva je plne prepojená vrstva s softmax aktivačnou funkciou. Táto vrstva obsahuje toľko neurónov, koľko rôznych tried je potrebné klasifikovať.

Softmax aktivačná funkcia je zovšeobecnenie sigmoidnej aktivačnej funkcie pretože tiež prideľuje pravdepodobnosť každému z neurónov (jeden neurón pre každú triedu). Pre neurón i a počet tried j, sa softmax funkcia vypočíta nasledovne:

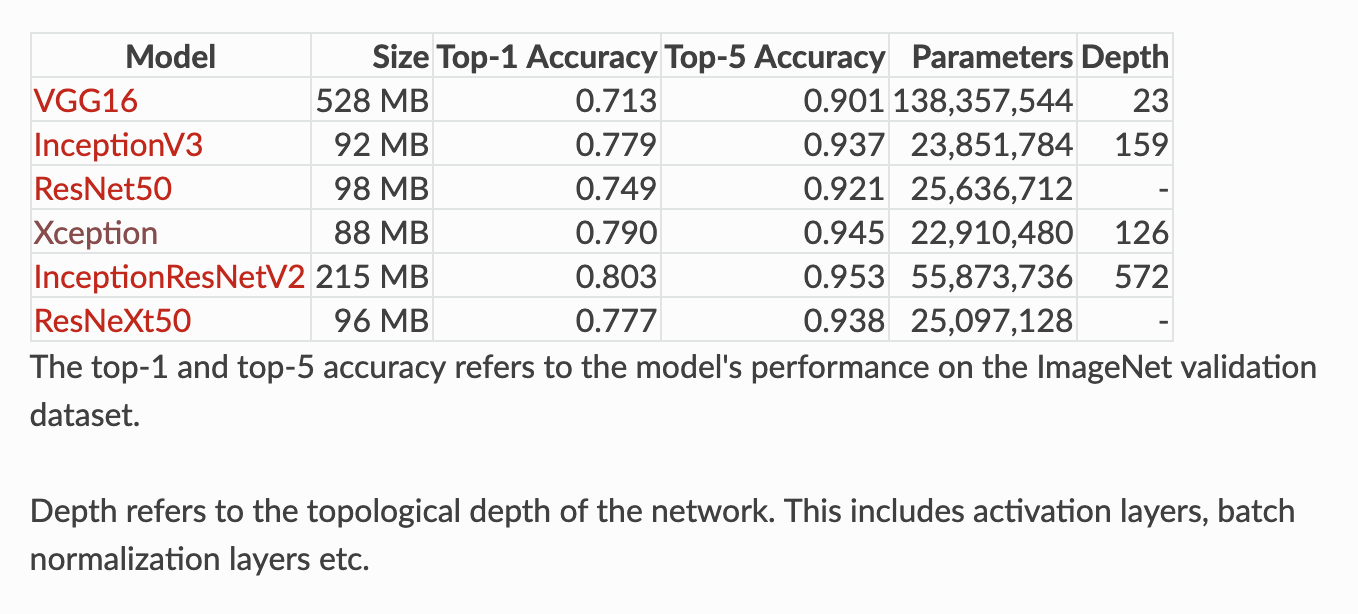
Najskôr sa vypočíta exponent váhami prenásobené vstupy neurónu i a následne sa vydelí súčtom exponentov všetkých váhami prenásobených vstupov všetkých j neurónov, aby sa normalizoval výstup a teda aby pravdepodobnosti pre všetky triedy po sčítaní dali 1.

## Architektúry KNS

Ako aj normálne neurónové siete, aj pri konvolučné sa skladajú z rôznych vrstiev. Avšak pri KNS existuje viacero rôznych typov vrstiev. Vo všeobecnosti sa konvolučné neurónové siete skladajú z troch hlavných typov vrstiev: konvolučné (vrátane aktivačnej vrstvy), pool vrstvy (najčastejšie max) a plne prepojených vrstiev (často zakončených softmax vrstvou). Najbežnejší spôsob tvorby konvolučných neurónových sietí je nakopiť niekoľko konvolučných vrstiev nasledovaných pool vrstvami a toto vrstvenie opakovať až dokým sa obrázok postupne nezmenší na veľmi malý (niekedy až na veľkosť 1 pixela). Následne sa aplikuje niekoľko plne prepojených vrstiev zakončených softmax vrstvou. Inak povedané, najbežnejší vzorec pre tvorbu architektúry KNS je nasledovný:

Zvyčajne 0 < N <= 3, M >= 0, 0 < K < 3. Samozrejme počet jednotlivých vrstiev a ich usporiadanie nie je jediná vec, ktorá ovplyvňuje výkonnosť konvolučnej neurónovej siete. Záleží aj na nastavení rôznych hyperparametrov, ako je veľkosť a počet filtrov v jednotlivých konvolučných vrstvách, stride, padding. Veľkosť výrezu v pool vrstvách, počet neurónov v plne prepojených vrstvách a mnoho ďalšieho.

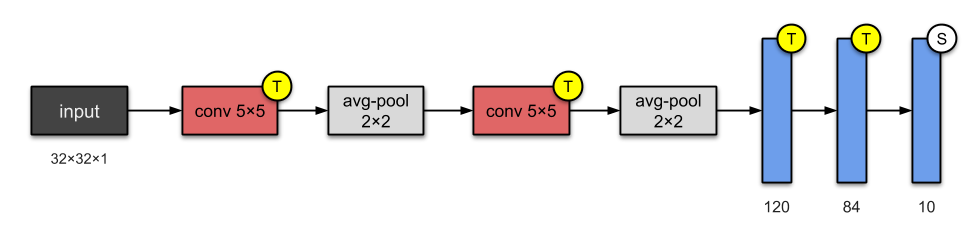
V posledných rokoch sa ukazuje vo svete veľké množstvo rôznych architektúr KNS. Mnohé z nich sú tak veľké a komplikované, že je veľmi náročné si predstaviť a zobraziť ako vyzerajú. V oblasti rozpoznávania obrazu bol zaznamenaný obrovský pokrok, avšak ako poznamenal výskumník z Googlu Christian Szegedy „most of this progress is not just the result of more powerful hardware, larger datasets and bigger models, but mainly a consequence of new ideas, algorithms and improved network architectures“. Avšak z toho veľkého množstva rôznych architektúr, existuje niekoľko veľmi často používaných sa dobrými výsledkami a je vidno že aj keď nie sú veľmi veľké, tak ich výsledky sú skvelé.



V tabuľke na obrázku je porovnaných 6 architektúr. Ich veľkosti, hĺbka a počet parametrov sa môžu veľmi líšiť a pritom výsledky sú veľmi podobné.

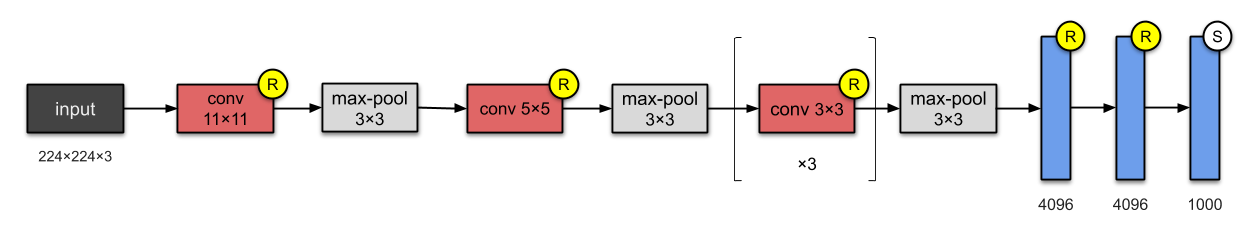
### LeNet-5

Jedna z najjednoduchších architektúr, ktorá má len 2 konvolučné a 3 plne prepojené vrstvy a len okolo 60 000 parametrov je LeNet-5 architektúra z roku 1998.



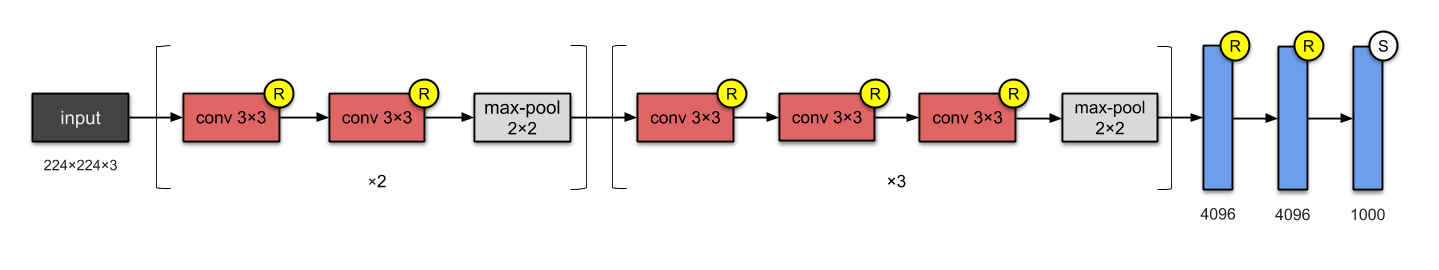
### AlexNet

Novšia architektúra z roku 2012 AlexNet ako prvá prišla z myšlienkou vymeniť Tanh aktivačnú funkciu za ReLU. Má 8 vrstiev a okolo 60 miliónov parametrov. V tej dobe to bola najväčšia konvolučná neurónová sieť použitá na ImageNet dataset.



### VGG-16

Je vidieť, že v poslednej dobe sa začali konvolučné neurónové siete zväčšovať, pretože najpriamejší a najjednoduchší spôsob ako zlepšiť KNS je zväčšiť jej veľkosť. Vo Visual Geometry Group (VGG) v roku 2014 vytvorili sieť zvanú VGG-16, ktorá má 13 konvolučných a 3 plne prepojené vrstvy. Táto architektúra pridala niekoľko vrstiev na AlexNet architektúru a použila menšie filtre (2x2 a 3x3). Pozostáva z približne 138 miliónov parametrov a zaberá až 528MB miesta. Okrem VGG-16 vytvorili aj väčšiu VGG-19.



# [Zoznam použitej literatúry](Pokyny_pre_vypracovanie_ZP.docx#Zoznam_použitej_literatúry)

# [Zoznam príloh](Pokyny_pre_vypracovanie_ZP.docx" \l "Prílohy)

# 

# Prílohy