



Rozpoznávanie tvárí pomocou doprednej neurónovej siete

Neurónové siete

Autor: Tomáš Kučečka
Študijný program: Softvérové inžinierstvo
Ročník: 1.
Rok štúdia: 2008/2009

Obsah

1	Zadanie	3
1.1	<i>Ciele.....</i>	3
2	Opis systému.....	4
2.1	<i>Opis filtrácie vstupov.....</i>	4
2.2	<i>Štruktúra neurónovej siete.....</i>	5
3	Opis aplikácie	6
4	Trénovanie.....	7
4.1	<i>Trénovacia množina</i>	7
4.2	<i>Trénovanie prístupom A</i>	7
4.3	<i>Testovanie prístupom B.....</i>	8
4.4	<i>Čo všetko bolo spravené.....</i>	9
5	Dosiahnuté výsledky	9
6	Možné vylepšenia systému	10
Príloha A	Výstup siete	12

1 Zadanie

Na tomto projekte som si za úlohu vybral problém hľadania tvárí na fotografiách. Mojou úlohou je pomocou neurónovej siete identifikovať na fotografiách tváre ľudí. Výsledkom by bola aplikácia, ktorej vstup by predstavovala fotografia a výstupom by bola tá istá fotografia, na ktorej by bola vyznačená nájdená pozícia tváre človeka.

1.1 Ciele

Z hľadiska zložitosti problému nie je mojim primárnym cieľom vytvoriť systém, ktorý by bol schopný rozpoznať tvár človeka na ľubovoľnej fotografii s vysokou pravdepodobnosťou. Hlavným cieľom je navrhnúť zaujímavé riešenie pre problematiku rozpoznávania tvárí, implementovať toto riešenie a overiť jeho správnosť a úspešnosť.

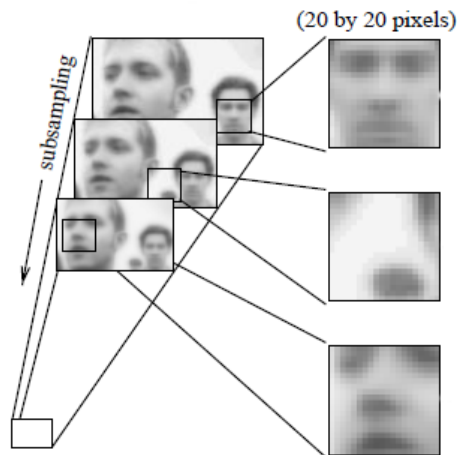
V prípade neúspechu vedieť zdôvodniť prečo je dané riešenie nevhodné, prípadne čo by bolo dobré zmeniť.

2 Opis systému

System pracuje na princípe pohybu snímacieho okna po zvolenej fotografii a zakresľovaním pozície do tejto fotografie v prípade nájdenia tváre týmto snímacím oknom.

V prvom kroku systém na pozícii snímacieho okna vyreže z fotografie malý obrazec veľkosti tohto snímacieho okna, teda 20 x 20 pixlov. Aplikuje na každý obrazec niekoľko filtrov. Na ohodnotenie každého obrazca použije už natrénovanú neurónovú sieť. Neurónová sieť vráti informáciu o tom, či daný obrázok tvárou je, alebo nie. V prípade ak daný obrázok tvárou je, vyznačí sa táto pozícia na obrázku nakreslením červeného štvorca na danom mieste. Systém sa následne presunie na ďalší obrazec zmenou pozície snímacieho okna o 2 pixely.

Po takomto prejde celú plochu vstupnej fotografie sa táto fotografia zmenší a opäť sa vykonajú vyššie uvedené kroky. Zmenšovanie fotografie prebieha, pokiaľ aspoň jedna z hrán fotografie nedosiahne veľkosť hrany vstupného obrazca. Postupné zmenšovanie fotografie znázorňuje obrázok číslo 1.



Obr. 1 Postupné zmenšovanie fotografie počas procesu hľadania tváre [1].

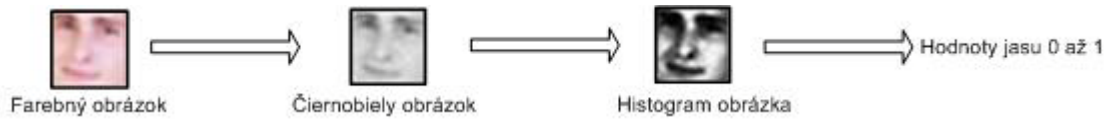
Bližšie o filtrovaní vstupných obrazcov, o tréningu neurónovej siete a o tréningovej množne pojednávajú nasledujúce kapitoly.

2.1 Opis filtrácie vstupov

Na každý obrazec, ktorý predstavuje vstup do neurónovej siete, sa aplikuje niekoľko filtrov v nasledujúcej postupnosti:

1. Zbavenie sa farieb – prevod na čiernobiely obrázok.
2. Vypočíta sa histogram obrázka
3. Výpočet jasů každého pixlu v obrázku

Tento postup filtrovania je zobrazený na obrázku číslo 2.



Obr. 2 Proces filtrovania vstupov

Dôvodom tejto filtrácie je zbavenie sa nadbytočných informácií o obrázku. Prvý krok predstavuje zbavenie sa farieb. V druhom kroku filtračného procesu sa zvýrazia obrisy na obrázku. To nám umožní lepšie sledovať špecifické črty tváří. Nakoniec sa pre každý jeden pixel vypočíta jeho jas, ktorý je v rozmedzí hodnôt 0 až 1. Táto vypočítaná hodnota predstavuje vstup do neurónovej siete.

2.2 Štruktúra neurónovej siete

Na rozpoznávanie tváří bola použitá dopredná neurónová sieť. Na jej implementáciu bolo použité kódovanie podľa Werbosa, teda zakódovanie siete do jednoduchšej štruktúry. Vybrané kódovanie bolo použité hlavne z dôvodu zvýšenia rýchlosti.

Prepojenie neurónov v tejto sieti nie je realizované spôsobom každý neurón s každým. Jej štruktúra je nasledovná [1]:

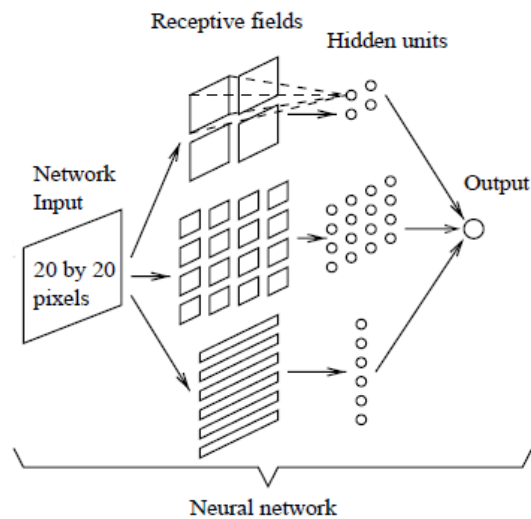
Vstupná vrstva pozostáva zo 400 vstupných neurónov (obrázok rozmeru 20 x 20 pixlov, každý jeden vstup predstavuje hodnotu jedného pixlu).

Skrytá vrstva je rozdelená do troch hlavných typov neurónov a to nasledujúco:

1. 4 skryté neuróny, ktoré snímajú obrazce vstupu rozmeru 10 x 10 pixlov
2. 16 skrytých neurónov, ktoré snímajú obrazce vstupu rozmeru 5 x 5 pixlov
3. 6 skrytých neurónov, ktoré snímajú prekrývajúce sa horizontálne obrazce vstupu rozmeru 20 x 5 pixlov

Spolu je to teda 26 neurónov v skrytej vrstve, ktoré však môžu byť replikované. Znamená to, že tieto tri typy skrytých neurónov sa zopakujú niekoľkokrát, v závislosti od nastavenia parametrov siete.

Výstupnú vrstvu tvorí jeden neurón, ktorý je prepojený so všetkými neurónmi skrytej vrstvy. Hodnota jeho výstupu môže nadobúdať hodnoty 0 až 1. V prípade, ak je hodnota výstupu v rozmedzí 0 až 0.5, vstupom je obrazec bez tváre. V opačnom prípade sa jedná o obrazec s tvárou. Štruktúru neurónovej siete znázorňuje obrázok číslo 3.



Obr. 3 Štruktúra neurónovej siete [1].

Skrytú vrstvu neurónovej siete je možné replikovať. Táto replikácia nám umožňuje zvýšiť správnosť výstupu. Jej hodnota je nastaviteľná, pričom odporúčaná hodnota je 3. Pri tejto hodnote dosahovala sieť najlepšie výsledky, ale aj jej naučenie trvalo najdlhšie.

3 Opis aplikácie

Aplikácia s používateľom komunikuje prostredníctvom GUI (Graphical User Interface). Je implementovaná v programovacom jazyku *Java*. Nasleduje v krátkosti zoznam toho, čo aplikácia ponúka a aké nastavenia umožňuje.

Aplikácia ponúka:

- Trénovanie neurónovej siete dvoma rôznymi prístupmi
- Nastavenie parametrov neurónovej siete
- Rôzne kontrolné výpisy počas testovania neurónovej siete
- Otestovanie siete na ľubovoľnej fotografii
- Rozloženie fotografie na menšie obrazce(*subimaging*)
- Zmenšovanie fotografie(*scaling*)

Aplikácia umožňuje tieto nastavenia parametrov neurónovej siete:

- Hodnota počtu replikácií skrytej vrstvy
- Rýchlosť učenia sa
- Momentum
- Hranica akceptácie
- Rozptyl

4 Trénovanie

Trénovanie siete prebiehalo dvoma rôznymi prístupmi. Trénovacia množina a jednotlivé prístupy trénovania siete, prístup A a prístup B, sú popísané nižšie v tejto kapitole. Ich označenie v aplikácii je zhodné s označením v tomto dokumente.

4.1 Trénovacia množina

Testovaciu množinu tvorili obrazce tváre rôzneho natočenia s uhlom natočenia hlavy do 10° . Táto množina bola vytvorená z voľne dostupnej množiny fotografií tvárí z internetu, určenej práve na takéto, prípadne podobné, účely. Následne bola z fotografie vystrihnutá tvár a zmenšená na rozmer 20 x 20 pixlov. Príklad trénovacích vstupov s tvárou znázorňuje obrázok číslo 3.



Obr. 3 Tváre fotiek s rôznym natočením hlavy

Dôvodom rôzneho natočenia tvárí je vedieť identifikovať na fotografiách rôzne polohy týchto tvárí s uhlom natočenia do 10° .

Ďalej bola trénovacia množina tvorená obrazcami bez tváre. Jednalo sa o obrazce rozmerov 20 x 20 pixlov. Tieto obrazce vznikli vystrihnutím obrazca z fotografií rôznych prostredí. Pri výbere množiny fotografií bola snaha vyberať fotografie s premenlivým prostredím.

4.2 Trénovanie prístupom A

Trénovanie prístupom A spočívalo vo vytvorení si niekoľkých trénovacích množín. Každá táto trénovacia množina obsahovala iba fotografie rozmeru 20 x 20 pixlov. Jednalo sa o fotografie tvárí a o fotografie prostredia (bez tvárí). Celkový počet týchto vstupov je 518, z čoho 174 vstupov predstavuje obrazce s tvárou.

Z oboch týchto testovacích množín dostávala neurónová sieť vstupy náhodne v jednotlivých epochách. Výsledky testovania sú popísané v tabuľke 1.

Tabuľka 1. Trénovanie prístupom A.

Replikácia skrytej vrstvy	Rýchlosť učenia	Momentum	Epocha	Percentuálna úspešnosť
1	1.0	0.7	100	100%
2	1.0	0.7	300	100%
3	1.0	0.7	3500	99,80%

Z tejto tabuľky je vidieť, že tréningovanie týmto prístupom je veľmi úspešné. Narastanie počtu epoch natréningovania siete pre danú množinu dát, je priamo úmerný počtu replikácií skrytej vrstvy.

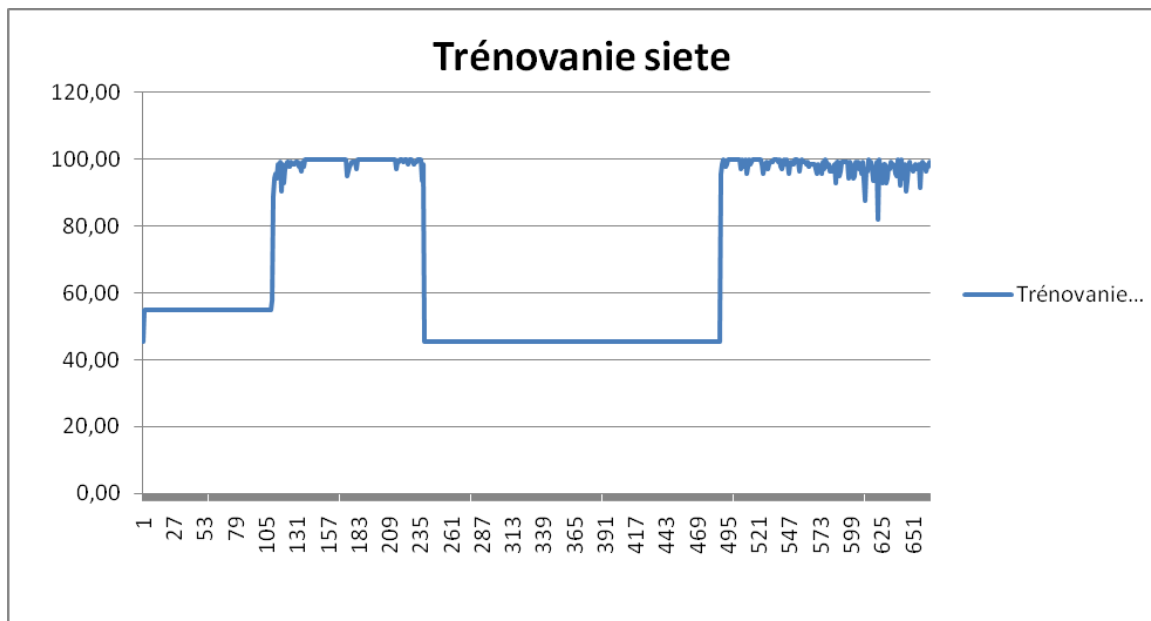
Aj napriek tomu, že testovaním týmto prístupom bola sieť schopná rozoznať tváre na jednotlivých tréningových obrazcoch s vysokou pravdepodobnosťou, sa tento postup ukázal ako nevhodný. Je to z dôvodu rozmanitosti okolitého prostredia. Týmto spôsobom bolo totiž otestované len zanedbateľné množstvo obrazcov bez tvárí. Sieť týmto pádom nevedela rozoznať iné pozadie na rôznych fotografiách. **Z tohto dôvodu sa pristúpilo k testovaniu prístupom B.**

4.3 Testovanie prístupom B

Tréningovanie prístupom B spočívalo v použití tréningovej množiny tvárí z prístupu A, kombináciou s fotografiami z internetu ľubovoľného rozmeru.

Hlavný rozdiel, oproti prístupu A, je použitie omnoho robustnejšej množiny obrazcov bez tváre. Tento prístup počas testovania rozkladá fotografie stiahnuté z internetu (bez tváre) na obrazce rozmerov 20 x 20 pixlov. Každý z týchto obrazcov následne posiela do neurónovej siete, striedajúc vstupy z množiny s tvármi.

Overovanie úspešnosti počas testovania bolo realizované vytvorením tretej množiny - testovacej, ktorá obsahovala obrazce s tvármi a taktiež obrazce prostredia, všetky veľkosti 20 x 20 pixlov. Treba však zdôrazniť, že **ani jeden z týchto testovacích obrazcov sa nenachádzal v tréningovej množine**. Pribeh testovania týmto prístupom je znázornený v grafe číslo 1.



Graf č. 1 Tréningovanie prístupom B. Na vodorovnej osi je zobrazený počet obrazcov v tisícoch.

Na tomto grafe je vidieť priebeh ako sa úspešnosť učenia siete pohybuje najskôr okolo 50%, následne sa sieť naučí rozpoznať testovaciu množinu úplne (100%), ale jej úspešnosť po istom čase znova klesne. Tento jav spôsobilo do siete vloženie novej sady vstupných obrazcov. Tieto nové vstupné obrazce mali úplne iný charakter od predošlých.

Celkovo sa tento prístup ukázal ako omnoho efektívnejší, hlavne vďaka omnoho pestrejšej množine vstupov.

4.4 Čo všetko bolo spravené

Pre zlepšenie celkového úspechu siete rozpoznávania tvárí na fotografiách boli použité rôzne spôsoby filtrácie, prístupy k trénovaniu a zmeny nastavení parametrov siete. Vyššie uvedené filtračné metódy a prístupy k trénovaniu sú už výsledkom testovania a hľadania možností vylepšenia trénovania neurónovej siete.

Vstup siete napríklad predstavovali hodnoty RGB, kombinácie jej zložiek, až napokon už samotný jas fotografie. Sieť bola testovaná pre rôzne hodnoty replikačných vrstiev, nastavení hodnoty rozptylu a podobne.

Na základe výsledkov trénovania bol navrhnutý prístup B, ako vylepšenie prístupu A. Týmto vylepšením sa nakoniec podarilo rapídne zlepšiť rozpoznávanie tvárí na fotografiách.

Pre zlepšenie sledovania chýb siete a hodnotenie výstupov bolo implementovaných viacero vylepšení, akými sú napríklad:

- exportovanie výstupov do súborov (označené fotografie s tvármi)
- exportovanie úspešnosti siete do textových súborov (vykresľovanie grafov)

5 Dosiiahnuté výsledky

Ako je vidieť z výsledkov testovania uvedených pri prístupoch A a B, úspešnosť siete pri učení sa na trénovacích vstupoch bola slušná. Napriek tomuto, úspešnosť siete na rozoznávaní tvárí na fotografiách je veľmi slabá. Tento fakt môže byť zapríčinený nasledujúcimi dôvodmi:

- Slabá filtrácia vstupných dát. Je potrebná lepšia filtrácia vstupov, aby sieť dostavala na vstupy menej redundantných dát
- Sieť bola testovaná iba na malom počte testovacích vstupov. Spolu bolo testovaných iba niečo nad 4 miliónov obrázkov. Boli by potrebné rozšíriť testovaciu vzorku. Znamená to vyskúšať aspoň 90 miliónov rozmanitých vzoriek.
- Trénovacia množina treba rozšíriť o identifikáciu tých obrázkov, pri ktorých sa sieť pomýlila a neskôr použiť už iba túto množinu na dotrénovanie neurónovej siete.
- Pri zmenšovaní obrázka sa veľmi rýchlo stráca kvalita tohto obrázka. Treba zlepšiť implementáciu zmenšovania rozmeru fotografií.

Testovanie siete preukázalo jej úspešnosť hlavne na menej pestrých fotografiách, ktoré sa nevyznačovali veľkým množstvom zaoblení a prechodov.

Celkovo ale vytvorený systém nepovažujem za nesprávny. Po zohľadnení vyššie uvedených bodov a vylepšení, je podľa mňa možné predpokladať mnohonásobne vyššiu efektivitu od implementovaného systému.

V prílohe A sú uvedené niektoré príklady rozpoznania tvárí na fotografiách.

6 Možné vylepšenia systému

Okrem už spomenutých vylepšení systému v 5 kapitole, ktoré si myslím výrazne prispeli k zlepšeniu úspešnosti systému, je ešte niekoľko ďalších vylepšení, ktoré som nestihol v projekte implementovať.

- Zahadzovanie svetlých pixlov. Pri každom vstupe by bolo zaujímavé spraviť priemer jasů pixlov a následne zmeniť farbu všetkých pixlov s podpriemernou hodnotou jasů na bielu. Od tohto prístupu očakávam najmä odbremenenie vstupov a tým aj siete od ďalších nadbytočných údajov, čo by malo za následok zvýšenie jej správnosti.
- Zaznamenávanie si vzoriek trérovacích vstupov počas trérovania, aby sme neskôr mohli siete opäť podať už prejdené obrazce. Dôvodom je, aby sa sieť nezabúdala, čo sa už naučilo dôsledkom stáleho

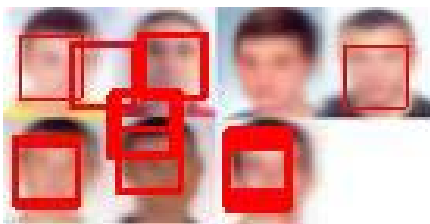
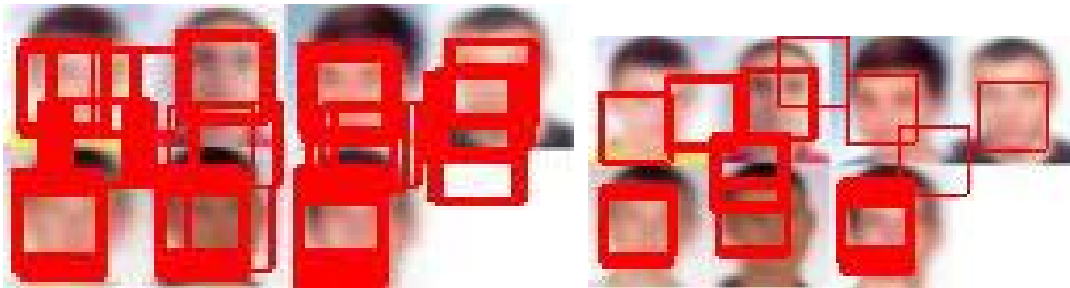
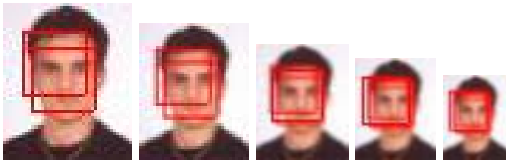
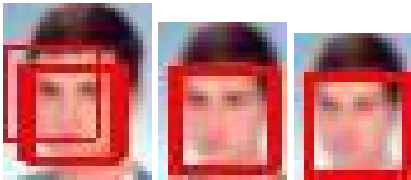
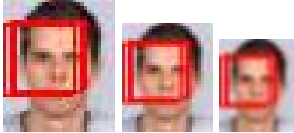
Použitá literatura

1. Rowley, H., Baluja, S., Kanade T.: Neural Network-Based Face Detection, In: IEEE Transactions On Pattern Analysis and Machine intelligence, 1998

Príloha A Výstup siete

V tejto prílohe sa nachádzajú ukážky výstupu neurónovej siete, implementovanej v rámci tohto projektu. Trénovanie siete bolo realizované prístupom B.

Príklad pomerne presného identifikovania tváre na jednoduchých fotografiách (fotografie študentov z akademického informačného systému):



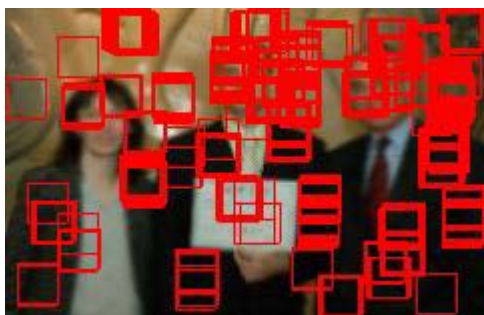
Príklad postupného zmešovania vstupnej fotografie. Najskôr vysoký počet zlých výstupov siete, ktorý sa postupne zmešoval so zmešováním obrázka:



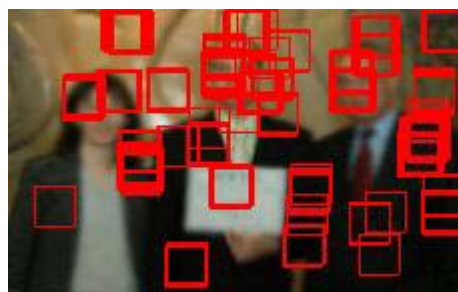
1 zmešenie



5 zmešenie



9 zmešenie



15 zmešenie



20 zmešenie



30 zmešenie