

HANDWRITTEN DIGIT RECOGNITION

Miroslav Štrba

Master Degree Programme (2), FIT BUT

E-mail: xstrba02@stud.fit.vutbr.cz

Supervised by: Adam Herout

E-mail: herout@fit.vutbr.cz

ABSTRACT

This paper describes training and testing parts for two handwritten digit recognition algorithms. Recognition accuracy of each algorithm is about 90 %. The first recognition method is based on artificial neural networks; the classifier is trained using unsupervised learning. The second algorithm is based on randomized tree deciding based on tags which are pixels with local neighborhood. Both methods are improved by using strong classifier principle.

1 ÚVOD

Existuje veľké množstvo rôznych implementácií rozpoznávacích algoritmov založených na rôznych spôsoboch učenia (support vector machine, neurónové siete, adaBoost ...). Tento článok popisuje dva nezávislé systémy. Prvá metóda, popísaná v časti 2, rozpoznáva číslice pomocou samoorganizujúcich sa máp (ďalej len SOM), ktoré svojou podstatou pripomínajú pamäť. Druhá metóda, popísaná v časti 3, využíva stromové klasifikátory a využíva tzv. tagy. Obe metódy sú trénované na štandardnej dátovej sade pre ručne písané číslice zostavenej skupinou NIST [2].

2 SAMOORGANIZUJÚCE SA MAPY - SOM

Táto kapitola popisuje systém založený na neurónových sieťach. Ako spôsob usporiadania neurónov boli zvolené samoorganizujúce sa mapy (SOM). Základom tohto rozhodnutia je fakt, že víťazný neurón SOM ovplyvňuje aj okolité neuróny čím výrazne urýchľuje proces učenia neurónovej siete. Viac o SOM je popísané v článku [3].

2.1 TRÉNOVANIE

Úlohou programu na trénovanie je vytvorenie neurónovej siete s lepšou schopnosťou rozpoznávania ako tá predchádzajúca. Počas trénovacieho procesu sa ku každému neurónu ukladá okrem váh vstupov aj informácia o tom, pre aký typ číslice je daný neurón víťazný t.j. jeho váhy najlepšie reprezentovali vstup. Naučený neurón reprezentuje typ číslice, ktorý sa počas trénovania najčastejšie stal vstupom pre daný víťazný neurón. Optimálna veľkosť SOM sa ukázala byť 10×10 . Vzhľadom na jej rýchle pretrénovanie bola trénovacia sada rozdelená na menšie časti. To spolu z náhodnou inicializáciou váh neurónov spôsobuje ich variabilitu pri zostavovaní silného klasifikátora.

2.2 ROZPOZNÁVANIE

Proces rozpoznávania je oproti tréновaniu výrazne jednoduchší. Najskôr sa načíta SOM. K načítanej sieti sa potom prikladá položka s dátovej sady. Následne sa určí víťazný neurón a podľa jeho typu je vstup prehlásený za danú číslicu. Počas rozpoznávania sa zbierajú rôzne štatistické informácie. Najdôležitejšou informáciou je úspešnosť rozpoznávania. Číslica je úspešne rozpoznaná ak sa typ víťazného neurónu zhoduje s hodnotou získanou s referenčného súboru.

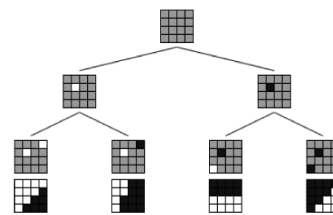
3 STROMOVÉ KLASIFIKÁTORY

Práca vychádza z článku [1]. Základom algoritmu je nájdenie vzťahu medzi významnými oblasťami, ktoré by boli unikátne pre jednotlivé druhy číslic. Významnou oblasťou je okolie pixelu, v ktorej je aspoň jeden pixel inej farby ako všetky ostatné pixely. Tie sa vyskytujú na hranách číslice a poskytujú informáciu o tvare číslice. Každé takéto okolie sa potom zaklasifikuje do určitej triedy. Zaklasifikovaná oblasť sa nazýva *tag*. Smer medzi dvoma tagmi sa nazýva pravidlo. Pomocou sady pravidiel rozhodovacieho stromu je potom možné určiť aký typ číslice sa na obrázku nachádza.

3.1 TRÉNOVANIE

Tréovanie je možné rozdeliť do niekoľkých fáz. V prvej fáze sa vyhľadajú významné oblasti. V druhej sa oblasti klasifikujú. Klasifikácia sa uskutočňuje pomocou stromu. Tento strom sa nazýva *tagovací strom*. Kritérium prechádzania tagovacím stromom je hodnota pixela na danej pozícii. Ukážka tagovacieho stromu je na obrázku 1.

V poslednej fáze tréovania sa pomocou tagovacieho stromu zostavuje *rozhodovací strom*. Podstatou rozhodovacieho stromu je využívanie pravidiel na triedenie číslic do podstromov až listov. Aby obrázok číslice obsahoval pravidlo musí byť pravidlo v spojení s pravidlami v nadradených uzloch. Keďže nie je možné počas tréovania preskúmať všetky pravidlá vo všetkých uzloch vyberá sa iba časť pravidiel pre každý uzol, ktoré sú ohodnotené. Vzniká tak variabilita medzi stromami nad rovnakou tréovacou sadou. Tá je opäť dôležitá pri vytváraní silného klasifikátora. Optimálna hĺbka rozhodovacieho stromu je 15.



Obrázok 1: Ukážka tagovacieho stromu s hĺbkou 2.

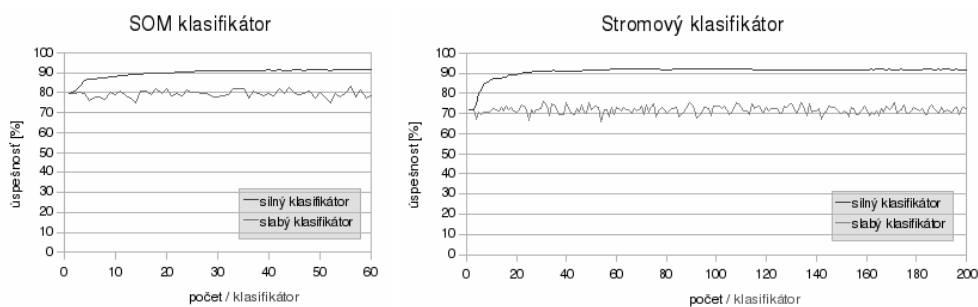
3.2 ROZPOZNÁVANIE

Proces rozpoznávania je časovo oveľa menej náročný. Na začiatku je nutné načítať testovaciu sadu a tagovací strom. S týchto súborov sa vytvorí testovacia množina. Množina už obsahuje vyhľadané tagy. Číslice sú teda pripravené na rozpoznávanie. Načíta sa rozhodovací strom a pomocou pravidiel sa jednotlivé vzorky zadeľujú do listov. Podľa toho do akého listu číslica prišla sa na základe štatistiky nazbieranej počas tréovanie rozhodne o jej type.

4 VÝSLEDKY

V grafoch na obrázku 2 je vidieť úspešnosť klasifikátorov. U SOM sa pohybuje na hranici 80 % a u stromových klasifikátorov je to 73 %. Táto hodnota je nedostačujúca a klasifikátory sú

označené ako *slabé*. Pre eliminovanie chybovosti vzniknutej unikániami vlastnosťami konkrétnej číslice je možné klasifikátory kombinovať. Kombináciou vznikne jeden *silný* klasifikátor, ktorého výsledky sú výrazne lepšie. Momentálne silný klasifikátor využíva princíp hlasovania slabých klasifikátorov a typ číslice určí na základe počtu hlasov pre daný typ. Už aj u takto jednoduchého spôsobu dosahuje SOM úspešnosť 91,4 % a u stromového silného klasifikátora 91,9 %.



Obrázok 2: Výsledky slabých a silných klasifikátorov. Vľavo sú zobrazené výsledky SOM klasifikátorov. Vpravo sú výsledky stromových klasifikátorov.

5 ZÁVER

V tejto práci sú popísané dva postupy ako rozpoznávať ručne písané číslice, ktoré boli implementované ako súčasť mojej diplomovej práce. Prvý postup využíva neurónovú sieť a jeho úspešnosť sa pohybuje na okolo 80 %. Zlúčením viacerých rovnakých SOM klasifikátorov úspešnosť vzrastie na 91,4 %. Druhý postup využíva stromový klasifikátor, ktorého úspešnosť sa pohybuje na úrovni 73 %. Zlúčením stromových klasifikátorov sa úspešnosť posunie na hranicu 92 %. Pri kombinácii silných klasifikátorov je možné dosiahnuť úspešnosť 95 %. Tieto hodnoty sú primerané k času vzniku odborných článkov, z ktorých som vychádzal, ale algoritmy nedosahujú také výsledky ako v súčasnosti, kde chybovosť najúspešnejšieho postupu, využíva kombináciu štyroch state-of-the-art techník a autorom je Daniel Keysers je 0,35 %. Ďalším cieľom diplomovej práce bude vytvorenie ešte aspoň jedného systému a následné zlúčenie klasifikátorov naučených rôznymi postupmi do jedného, ktorý by mal popisovať číslice komplexnejšie a tým aj účinnejšie.

Pod'akovanie: Táto práca vznikla čiastočne za podpory grantu VUT FIT, FIT-S-10-2 a špecifického výskumu MSM0021630528.

REFERENCIE

- [1] Geman, D., Amit, Y., Wilder, K.: Joint Induction of Shape Features and Tree Classifiers. In: IEEE Trans. PAMI. Volume 19. 1997.
- [2] LeCun, Y., Cortes, C.: THE MNIST DATABASE of handwritten digits. 2007. Dokument dostupný na URL <http://yann.lecun.com/exdb/mnist/> (December 2009).
- [3] Kohonen, T.: The self-organizing map. In: Neurocomputing. Volume 21. Issues 1-3. 6 November 1998. Pages 1-6. ISSN 0925-2312.