

Návrh zadania diplomovej práce

Predbežná verzia (podpísaná) ¹

Študent:

Meno, priezvisko, tituly: Martin Schnürer, Bc.
Študijný program: Inteligentné softvérové systémy
Kontakt: xschnurer@is.stuba.sk

Výskumník:

Meno, priezvisko, tituly: Peter Malík, Ing. PhD.

Projekt:

Názov: Detekcia osôb hlbokými neurónovými sieťami
Názov v angličtine: Deep neural network for human detection
Miesto vypracovania: Ústav počítačového inžinierstva a aplikovanej informatiky, FIIT STU, Bratislava
Oblasť problematiky: Hlboké učenie a konvolučné neurónové siete

Text návrhu zadania²

Počas posledného desaťročia pritiahla detekcia osôb výraznú pozornosť v sfére počítačového videnia a rozoznávania vzorov najmä v dôsledku širokej škály úloh, na ktoré môžu byť aplikovateľné. Schopnosť spoľahlivo rozoznať polohu osoby na obrázku alebo videu je náročnou úlohou, s ktorou sa stretávame pri kamerovej detekcii osôb v priestore alebo u autopilota pri jazde autom. Detekcia osôb je jedným z najťažších problémov v kategórii detekcií objektov, nehovoriac o rozoznávaní polohy osoby na obrázku. Problém detekcie osôb môže byť jednoducho poňatý ako lokalizácia subjektov na danom obrázku, pričom subjekty sú konkrétne označené ohraničenou oblasťou na obrázku.

Analyzujte súčasný stav a navrhňte riešenie problematiky detekcie osôb a klasifikácie ich polohy na obrázku pomocou hlbokých konvolučných neurónových sietí. Navrhňte efektívny model hlbokéj neurónovej siete a vhodnú topológiu. Redukujte počet parametrov tohto modelu a maximalizujte jeho presnosť. Vyhodnoťte a porovnajte experimentálne výsledky detekcie osôb a klasifikácie ich polohy a zhodnoťte príspevok k aktuálnej problematike.

¹ Vytlačiť obojstranne na jeden list papiera

² 150-200 slov (1200-1700 znakov), ktoré opisujú výskumný problém v kontexte súčasného stavu vrátane motivácie a smerov riešenia

Literatúra³

- Szegedy, C. - Toshev, A. - Erhan, D.: Advances in Neural Information Processing Systems: Deep Neural Networks for Object Detection :NIPS, 2013: 9s:
- Nguyen, D. - Li, W. - Ogunbona P.: Pattern Recognition: Human detection from images and videos: Elsevier March 2016, Pages 148-175,

V Bratislave dňa 10.12.2017

Podpis študenta

³ 2 vedecké zdroje, každý v samostatnej rubrike a s údajmi zodpovedajúcimi bibliografickým odkazom podľa normy STN ISO 690, ktoré sa viažu k téme zadania a preukazujú výskumnú povahu problému a jeho aktuálnosť (uvedte všetky potrebné údaje na identifikáciu zdroja, pričom uprednostnite vedecké príspevky v časopisoch a medzinárodných konferenciách)

Výskumný zámer

Martin Schnurer

Detekcia osôb hlbokými neurónovými sieťami

Ing. Peter Malík PhD.

Riešenie problému: Detekcia osôb hlbokými neurónovými sieťami

Prečo je tento problém dôležité riešiť

- Rozoznávanie osôb má veľký význam pri interakcii človeka s počítačom. Častokrát sa vo svete používajú prístroje, ktoré sú riadené výlučne strojom. Do tohto prostredia môže byť zapojený aj človek a preto musíme dbať v prvom rade na bezpečnosť človeka a vyhnúť sa situácií, kde by mohol byť človek alebo jeho zdravie v ohrození. Včasné a správne rozoznávanie človeka strojom (pomocou kamerového systému) preto správne upravuje funkcionality stroja tak, aby dbal na bezpečnosť a zdravie človeka.
- Riešenie tohto problému má význam aj pri riešení verejnej bezpečnosti. K dispozícií sú čoraz lacnejšie a efektívnejšie riešenia kamerových systémov na verejných priestranstvách, avšak pri páchaní zločinu je často krát možné odhaľovať skutočnosti len pomocou ľudského experta. Automatická detekcia osôb, ich činnosti a odhaľovanie podozrivých by uľahčila ľuďom prácu a zvýšila tak bezpečnosť verejnosti a podporila znižovanie kriminality.
- Rozoznávanie osôb častokrát slúži ako medzičlánok k ďalšiemu spracovaniu obrazu. (napr. vystrihnutie časti obrazu na ktorých je človek)

Analýza oblasti

Rozoznávanie osôb a identifikácia ich polohy na obrázku je pre človeka jednoduchá a rýchlo vyriešiteľná úloha. Ak by sme sa však túto istú úlohu snažili vyriešiť pomocou počítača, narazili by sme na veľa problémov a nájdenie takého algoritmu, ktorý by tento problém vyriešil, by bolo viac než obtiažnou úlohou.

Rozoznávanie osôb bolo hlavnou témou v počítačovom videní už viac ako 20 rokov výskumu. Za tento dlhý čas sa vyvinulo mnoho metód, ktoré tento proces urýchlili a zlepšili jeho výsledok [2]. Častokrát sa pri rozoznávaní osôb tieto metódy zameriavajú na riešenie rôznych typov problémov. Niektoré sa zameriavajú na základné črty riešenia, iné sa zase zameriavajú na *učiace algoritmy*, alebo sú aj také, ktoré do riešenia začleňujú rôzne pokročilé algoritmy, napríklad algoritmus “deformovateľných častí modelu” [6]. Práca Viola a Jones navrhli riešenie problému pomocou kaskádových klasifikátorov, ktoré boli široko používané v realtime aplikáciách. Táto metóda bola rozšírená o radu čít a techník, avšak

základom tejto kaskády bolo skoré vyradenie väčšiny testovacích vzoriek, čím bola dosiahnutá vysoká výkonnosť a spracovanie v reálnom čase.

Pravdepodobne najpopulárnejšou pre rozoznávanie osôb je technika HOG [4]. Vykazuje ale nižšiu rýchlosť spracovania 1 FPS (Snímok za sekundu). Táto technika bola zárodkom pre nové metódy ako rozoznávanie pomocou SVM (angl. Support Vector Machine) a rozhodovacie stromy (angl. decision tree) [3][5].

Techniky hlbokého učenia boli taktiež použité na rozoznanie osôb a časom viedli k presnejším výsledkom [1]. Tento prístup je avšak stále pomalý - týmto spôsobom sa podarilo dosiahnuť výkonnosť spracovania obrazu 1 snímok za sekundu až 1 snímok za niekoľko minút. O zrýchlenie tohto prístupu sa pokúsil Benson et al. ktorý navrhol metódu pomocou hlbokých neurónových sietí s rýchlosťou spracovania 135 snímok za sekundu pre vstupný obraz s veľkosťou 480x640, avšak s omnoho menšou presnosťou.

V práci Angelova et al. sa prvý krát podarilo dosiahnuť pomocou hlbokých neurónových sietí spracovanie v reálnom čase a výsledky mali vysokú presnosť. V tejto práci boli použité "rýchle kaskády" spolu s hlbokými neurónovými sieťami. Avšak snaha bola redukovať architektonické črty, ako napríklad hĺbku neurónovej siete a veľkosť poľa vnemu (angl. receptive field). Aj napriek navrhovaným zmenám v architektúre je považované riešenie za pomalé a stále nedostatočné pre aplikácie v reálnom čase (angl. realtime applications). Táto architektúra je podľa autorov vylepšiteľná a ďalej skúmateľná zvýšením počtu vrstiev a regulovaním veľkostí menších konvolučných sietí.

Rozpracovanie problému

Na základe analýzy a predošlých prác venujúcim sa rozoznávaniu osôb som zachytil niekoľko problémov, resp. oblastí, v ktorých vidím možnosť pokroku a zlepšenia:

- Často krát sa práce snažili obetovať presnosť výsledkov na úkor zvýšenia rýchlosti spracovania snímok (za jednotku času). Mojm cieľom by preto bola snaha túto rýchlosť zachovať, pričom sa chcem zamerať na architektúru neurónových sietí a zlepšiť ich presnosť - na základe predošlých prác, ktoré navrhovali **zvýšiť počet skrytých vrstiev a upraviť zrnitosť a veľkosť** menších konvolučných neurónových sietí, z ktorých bola výsledná architektúra zostavená.

Referencie

- [1] A. Angelova, A. Krizhevsky, and V. Vanhoucke. Pedestrian detection with a large-field-of-view deep network. ICRA, 2015.
- [2] R. Benenson, M. Matthias, R. Tomofte, and L. Van Gool. Pedestrian detection at 100 frames per second. CVPR, 2012
- [3] R. Benenson, M. Matthias, T. Tuytelaars, and L. Van Gool. Seeking the strongest rigid detector. CVPR, 2013.
- [4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. CVPR, 2005.
- [5] P. Dollar, C. Wojek, B. Schiele, and P. Perona. Pedestrian detection: A benchmark. CVPR, 2009
- [6] D. Park, D. Ramanan, and C. Folwkes. Multiresolution models for object detection. ECCV, 2010