

Posudek oponenta disertační práce

Student: Ing. Karel Petránek

Název práce: Detekce objektů ve videosekvenci

Školitel: Prof. RNDr. Eva Milková, Ph.D.

Oponent: Doc. Ing. Ladislav Beránek, CSc.

Práce Ing. Karla Petránka se zabývá oblastí detekce objektů v obraze a ve videosekvenci. Autor stručně popisuje některé algoritmy používané pro detekci objektů, přičemž se soustředí na popis konvolučních neuronových sítí, které posléze využívá pro detekci objektů. Autor dále navrhuje vlastní datovou sadu, na které testuje navržené metody – detektor umístění objektu ve videosekvenci založený na použití konvoluční neuronové sítě a částicovém filtru a modifikovaný operátor Local Binary Patterns.

Autor se zabývá oblastmi, které se v posledních letech velmi rychle rozvíjejí a kde bylo dosaženo mnoho významných výsledků, ať už v oblasti rozpoznávání tak i v oblasti hlubokých neuronových sítí. Celkově práce může sloužit jako určitý úvod do problematiky některých metod a přístupů v oblasti rozpoznávání objektů.

Budu nejprve stručně charakterizovat práci a následně uvedu dílčí připomínky odkazující na konkrétní místa práce.

Aktuálnost zvolené problematiky

Autor předložil práci, která se věnuje dnes velmi se rozvíjející problematice, oblasti rozpoznávání objektů v obraze a videosekvenci. Jedná se o téma s mnoha zajímavými výsledky a velkým množstvím vědeckých a odborných publikací, s mnoha aplikacemi, např. v oblasti rozpoznávání tváří, sledování pohybu zvířat nebo pohybu hráčů u různých sportů, autonomního řízení a v mnoha dalších oblastech.

Stanovené cíle doktorské disertační práce (splnění cílů)

Autor definuje hlavní cíl disertační práce následovně:

„Hlavním cílem disertační práce je navrhnout systém či systémy pro detekci objektů ve videosekvenci, které budou mít dostatečnou přesnost na to, aby výrazně usnadnily pro člověka únavné úlohy či úplně nahradily lidský dozor“. Na základě tohoto obecného cíle pak autor formuluje dílčí cíle:

1. Testování stávajících state-of-the-art metod na nestandardní datové sadě;
2. Analýza a zlepšení extrakce atributů;
3. Zlepšení detekce umístění objektu v obraze;
4. Propojení detekce objektu s jeho sledováním ve videosekvenci.

Z formulace hlavního cíle a dílčích cílů, zejména dílčích cílů 2,3 a 4, je vidět, že si autor stanovil cíle poměrně vysoko. V podstatě z každé dílčí oblasti by bylo možné zpracovat disertační práci. Možná i to je důvodem, proč autor popsal praktickou část své práce poměrně povrchně a není tak, alespoň pro mě, možné do detailu odhalit všechny přínosy autora ve výše uvedených oblastech. Nicméně se podle mého soudu autorovi v zásadě podařilo jím formulované cíle s připomínkami uvedenými v dalších odstavcích splnit.

Struktura a formální stránka práce

Struktura práce je víceméně obvyklá; autor nejprve uvádí teoretický rámec problematiky, v němž autor popisuje základní pojmy a současný stav řešené problematiky (kapitola 3), a to, architekturu detekčních algoritmů, algoritmy pro přezpracování obrazu, algoritmy pro extrakci atributů, klasifikační algoritmy, umělé neuronové sítě, detekci objektů ve videosekvenci a dále popisuje některé existující datové sady. Celkem tato teoretická část je na 38 stranách. Kapitola 4 Dosažené výsledky, kde autor představuje plnění cílů své práce, obsahuje popis plnění cílů, které si autor vytyčil. Autor na celkem 22 stranách představuje návrh nové datové sady, návrh vylepšení extrakce atributů pomocí Local Binary Patterns, návrh metody sledování objektů ve videosekvenci pomocí konvoluční sítě a částicového filtru, a segmentaci objektů ve videosekvenci. Kapitoly Diskuze výsledků (kapitola 5) a Závěr (kapitola 6) pak uzavírají celou práci (celkem 5 stránek).

Celkově je struktura práce standardní, nicméně se nemohu ubránit dojmu, že autorovi po sepsání teoretické části ubývaly síly. Očekával bych, že kapitoly s návrhem řešení a metod včetně popisu jejich konkrétního provedení budou zabírat většinu práce. Namísto toho se autor věnuje někdy zbytečnému popisu metod v oblasti detekce objektů, například obrázek 7 nebo popis klasifikačních algoritmů (kapitola 3.4), které jsou poměrně dobře známy. Autor se měl podle mého názoru věnovat více popisu vlastního přínosu, v kapitole 4 Dosažené výsledky uvést podrobné popisy navržených přístupů a modifikací použitých metod. Autor sice uvádí některé odkazy na svoje publikované práce, které obsahují podrobnější výsledky, nicméně ty měl buď vložit do příloh, nebo složit disertační práci z těchto publikovaných výsledků.

Po formální stránce je práce zpracována celkem pečlivě a přehledně. V práci se vyskytuje jen poměrně málo chyb, např. strana 11, první věta v podkapitole 3.3.2 „Algoritmus SIFT [Lowe, 1999] extrahuje ze vstupního snímku množinu význačných“ (chybí zde slovo bodů), použití sousloví „stávající state-of-the-art metody“ nebo zase anglické state-of-the-art místo českého stávající a další. Dále bych doporučoval číslování vzorců, a zvýraznění vektorových veličin ve vzorcích. Přece jenom to zlepšuje čitelnost textu, i když vzorce uvedené v práci nejsou příliš složité. Také formulace na několika místech nejsou příliš dobré, např. formulace na straně 29 7 ř. shora ... „ve špatném lokálním minimu“, nebo na straně 26, ř. 11 shora ... „rozdíl mezi ideálním výstupem a výstupem sítě“, na straně 45 někde je označení pro počet bitů LBP použito malé p , jinde velké P , a další formální nedostatky.

Seznam literatury je dostatečně široký a spolu s teoretickou částí prokazuje dobrou schopnost autora v práci s literárními zdroji.

Nové poznatky, které byly dosaženy v doktorské disertační práci

Autor provedl analýzu současných metod a přístupů v oblasti detekce objektů. Ta může sloužit jako určitý úvod pro zájemce o tuto oblast. Dále autor navrhl modifikaci metody LBP (Local Binary Pattern), která slouží k popisu vlastností obrázků pomocí příznaku, zpravidla histogramem. Autor měl ale popsat tuto modifikaci podrobněji. Totéž je možno konstatovat u autorem navržené modifikace SharpMask a modifikace LBP. Celkově je škoda, že autor popsal jím navržené modifikace shora uvedených metod velmi povrchně. Postrádal jsem zde alespoň trochu matematický popis jednotlivých modifikací, pseudokód navržených algoritmů, ilustrativní obrázky a další. Například autor pro segmentaci objektů (kapitola 4.4) využívá modifikovanou síť SharpMask, kterou v původní podobě zobrazuje v teoretické části na obrázku 13. Nicméně v kapitole 4.4 mohl autor také znázornit obrázkem, jaké úpravy původní architektury sítě provedl. Zlepšilo by to čitelnost a bylo by hned zřejmé, jaké úpravy autor provedl.

Některé popisy včetně datové sady mohl autor dát do příloh, případně na CD. Takto se obtížně posuzují jednotlivé detailní přínosy autora.

Připomínky

Dílejší připomínky k některým místům disertace mohou být vnímány také jako náměty do diskuse, případně jako náměty na doplnění dalšího výzkumu.

- Kapitola 4.1 Návrh datové sady. Popis datové sady je příliš stručný. Pro účely dalšího zpracování by bylo potřeba uvést další údaje, např. použitý software pro záznam, vzorkovací frekvence, způsob záznamu – např. záznam byl proveden statickou kamerou umístěnou nad experimentálním boxem určitým způsobem (včetně např. obrázku umístění kamery).
- Autor na straně 14 uvádí (druhý odstavec) „Nevýhodou uniformních LBP je ztráta informace z 20 % vzorů, které jsou zařazeny do kategorie ostatních vzorů a náchylnost na šum v oblastech s konstantní světelností. V rámci disertačního výzkumu bylo navrženo rozšíření uniformních LBP, které zmíněné problémy řeší, viz kapitola 4.2 ...“ Nicméně kapitola 4.2 uvádí pouze stručný popis modifikace metody LBP, zmíněné problémy autor v této kapitole ani jinde v disertační práci neřeší.
- V tabulce 6 na straně 59 se mi zdá hodnota Precision u metody R-CNN nízká. Nemohlo zde dojít k chybě?
- Výsledky testování v kapitole 4.4.2 by mohly být provedeny důkladněji, například by bylo vhodné znázornit křivku precision-recall pro různé objekty použité datové sady. Také zde mohlo být provedeno porovnání s jinými metodami popisovanými v práci.
- V práci bohužel nejsou zmínky o použitých programových nástrojích, postupech, případně použitém frameworku, knihovnách, implementaci a případné optimalizaci apod. Využil autor prostředí Matlab nebo nějaký open source framework?
- V některých případech by bylo vhodné okomentovat, proč autor volil určitý přístup nebo metodu nebo jejich kombinaci.

Publikační činnost autora ve vztahu k práci

Publikační činnost autora (uvedená od str. 80) zahrnuje celkem 24 publikačních aktivit autora práce (z toho 3 v časopisech registrovaných v databázi SCOPUS a 4 z konferencí registrovaných ve WoS). Publikační činnost potvrzuje nejen autorovu pracovitost a systematičnost, ale i dlouhodobý zájem o oblast, která je tématem disertační práce.

Otázky k obhajobě

- Provedl jste modifikaci metody LBP. Mohl byste podrobně teoreticky vysvětlit, proč došlo ke zvýšení diskriminačních schopností vámi modifikované LBP, jak uvádíte v kapitole 4.2.4?
- Použil jste u částicového filtru konturový model nebo vzor objektu? Co považujete za výhodnější? Nebo jste použil jiný přístup?
- Jak by se vyrovnala metoda Lukas-Kanade s rychlým pohybem myši? Testoval jste to?
- Co je myšleno označením $O(2^p)$, např. na straně 14 1 ř. shora „... exponenciální počet $O(2^p)$, ...“?

Závěr

Závěrem mohu konstatovat, že autor disertační práce má odpovídající teoretické znalosti předmětného tématu, orientuje se v problematice předmětného výzkumu. Problematika, kterou se autor zabývá, je velmi aktuální. Bohužel povrchní popis dosažených výsledků podle

mého názoru snižuje celkově dobré pracovní výsledky autora. Nicméně doufám, že autor u obhajoby vysvětlí podrobně své postupy. Závěrem konstatuji, že disertační práci Ing. Karla Petránka „Detekce objektů ve videosekvenci” doporučuji přijmout k obhajobě před komisí studijního oboru Aplikovaná informatika a po úspěšné obhajobě udělit Ing. Karlu Petránkovi titul Ph.D.

České Budějovice, 15.1.2018

Podpis: