

OPONENTSKÝ POSUDOK DIZERTAČNEJ PRÁCE

Meno doktoranda: Ing. Filip Kadlček

Názov práce: NÁVRH A OPTIMALIZACE OBRAZOVÝCH KLASIFIKÁTORŮ

1) Zameranie a aktuálnosť

Práca Filipa Kadlčeka je zameraná na zefektívnenie automatizovaného spracovania obrazu pomocou návrhu nových obrazových klasifikátorov a urýchlenia spracovania obrazových údajov návrhom novej vysoko priepustnej hardvérovej architektúry s konštantnou rýchlosťou spracovania obrazových údajov. Oblasť automatizovaného spracovania obrazu je vysoko aktuálnou oblasťou záujmu pre výskum a vedecké bádanie kde prevažujú dva hlavné trendy: zvýšenie presnosti detekcie a klasifikácie, a zvýšenie efektivity spracovania obrazových údajov s cieľom maximalizovať priepustnosť. Jednoduchý prístup k lacným kamerovým systémom, neustále narastajúci počet a rozlíšenie video kanálov v spojení s rastúcim výkonom výpočtových systémov záujem o výskum v tejto oblasti naďalej zvyšuje. Práca má stanovené štyri hlavné ciele: návrh nových aplikačne-špecifických príznakov, návrh novej aplikačne-špecifickej architektúry obrazových klasifikátorov, návrh architektúry s kompletne prúdovým pracovaním údajov a konštantnou rýchlosťou spracovania obrazových údajov, a návrh postupu pre automatizované generovanie aplikačne-špecifických klasifikátorov. Prvé dva ciele sú orientované hlavne na trend zvyšovania presnosti detekcie a klasifikácie. Druhý a hlavne tretí cieľ je orientovaný na trend zvyšovania efektivity spracovania obrazových údajov. Štvrtý cieľ je orientovaný na zjednodušenie vytvorenia nových obrazových klasifikátorov pomocou automatizovaného nástroja. Všetky tieto ciele sú vysoko aktuálne.

2) Prínos a originalita

Práca Filipa Kadlčeka vychádza z dosiahnutých výsledkov v jeho diplomovej práci s názvom „Implementace obrazových klasifikátorů v FPGA“. Autor používa viacero odkazov na túto prácu bez konkrétneho špecifikovania pôvodných výsledkov čo skomplikovalo rozlíšenie nových prínosov. Autor adresoval všetky jeho štyri pôvodné ciele a podarilo sa mu dosiahnuť pekné výsledky v každom z nich. Ako prvý originálny prínos považujem vytvorenie novej metódy návrhu aplikačne-špecifických tvarov príznakov pomocou genetického algoritmu pre konkrétnu cieľovú technológiu implementácie. Originalitu riešenia vidím v parametrizácii návrhu nových tvarov príznakov umožňujúci nepravidelné a nespojité tvary príznakov v spojení so súčasným vyhodnocovaním ich hardvérových nárokov v cieľovej FPGA technológii. Výsledkom je možnosť optimalizovať presnosť detekcie nových príznakov spolu s ich hardvérovými nárokmi. Druhý originálny prínos vidím v návrhu novej výpočtovej architektúry aplikačne-špecifického obrazového klasifikátora pre AdaBoost metódu. Originalita je vo vysoko-paralelnom prúdovom spracovaní údajov s konštantnou rýchlosťou výpočtu, ktorý je nezávislý od komplexnosti vstupného obrazu. Posledný originálny prínosom vidím v návrhu novej metódy pre automatizované vytváranie nových aplikačne-špecifických klasifikátorov. Originalita je v prepojení viacerých metód a postupov spolu s existujúcimi nástrojmi čo umožňuje vytvoriť nový klasifikátor automatizovaným postupom v niekoľkých krokoch od vytvárania nových príznakov cez návrh klasifikátora až po generovanie výsledného opisu v harvérovom opisnom jazyku.

3) Publikačná činnosť

Filip Kadlček publikoval svoje výsledky celkovo na 7 konferenciách. Tri konferenčné publikácie sú z menších Ph.D. seminárov (2 x PAD, 1 x Bata) a 4 z medzinárodných konferencií. Tri publikácie sú indexované vo vedeckej databáze Scopus. Najvýznamnejšia publikácia z medzinárodnej konferencie DDECS 2013 je indexovaná aj vo vedeckej databáze Web of Science. Najväčšou slabinou publikačnej činnosti je absencia publikovania v karentovanom časopise. Filip Kadlček v súčasnosti spolupracuje na novej časopiseckej publikácii ale zatiaľ táto publikácia nebola podaná a tak ju nemôžem započítať. Filip Kadlček spĺňa minimálne požiadavky na publikačnú činnosť.

4) Vedecká erudícia študenta

Dizertačná práca má jasne stanovené ciele, je dobre členená, text je zrozumiteľný, jednotlivé kroky sú dobre rozpísané a dosiahnuté výsledky sú kvantifikované a porovnané so súčasnými riešeniami. V úvodnom prehľade autor vôbec nespomenul gradientné príznaky a práve príznaky typu HOG sú vo všeobecnosti jedny z najpoužívanejších. Autor sa úplne vyhol aj novému trendu hlbokého učenia, ktoré v posledných rokoch dosahuje dobré výsledky v presnosti detekcie a aj v rýchlosti spracovania. Aj keď autor sa orientoval na metódy typu boost, tak ostatné metódy mali byť v krátkosti spracované v úvodnom prehľade. Autor použil celkovo 98 zdrojov literatúry, ale často chýba ich kompletný zápis. Viacero literárnych zdrojov nemá uvedený rok, čísla strán, verzie dokumentov, adresu zdroja a čas prístupu. Navyše, použitá literatúra nieje zoradená podľa poradia odkazovania a tak čitateľ musí listovať v zozname literatúry od prvých strán. Text práce obsahuje aj viacero chybných údajov. Autor pracujúci s technológiou FPGA od firmy Xilinx by mal vedieť, že FPGA obvody tejto firmy od rodiny virtex 5 obsahujú 6-vstupové LUT tabuľky, ktoré nahradili pôvodné 4-vstupové LUT tabuľky. Rovnako by mal vedieť že ak v Tabuľke 5 uvedie 161 DSP48e blokov, znamená to, že daná realizácia využíva 161 násobičiek a tento údaj by mal uviesť v Tabuľke 6. Nevhodné adresovanie tabuliek zhoršuje čitateľnosť. Príkladom je odkaz na Tabuľku 11 celých 7 strán pred jej umiestnením. Text obsahuje aj ďalšie technické nepresnosti. Autor pod FullHD rozlíšení rozumie 1920*1200, čo v skutočnosti je rozlíšenie s pomerom strán 8:5 označené ako WUXGA a FullHD je definované ako rozlíšenie 1920*1080 s pomerom strán 16:9. Ďalšie nedostatky je vidieť na obrázkoch. Obrázok 28 ukazuje dosť náhodné výsledky, pretože experimentu využívajúci genetický algoritmus bol vykonaný iba 3 krát. Obrázok 33 má zle zvolenú metriku porovnania, čo vedie k mylnému záveru, že suma správnej a chybné detekcie je väčšia ako 100%. Najväčším nedostatkom je nekorektné porovnanie rýchlostí spracovania obrazu autorovho riešenia so sekvenčným riešením. Autor pri porovnaní ignoroval obrovský nepomer použitých zdrojov v FPGA. Autor vôbec neuviedol, že sekvenčné riešenie môže byť jednoducho paralelizované zdvojením v kombinácii s rozdelením vstupného obrazu na polovicu a ich čiastočným prekrytím.

5) Ďalšie skutočnosti

Dizertačná práca má celkovo 116 strán a text je vhodne doplnený tabuľkami a obrázkami. V texte som našiel niekoľko gramatických chýb aj preklepov (príkladom je strana 25 „zbývali“ namiesto zabývali, strana 85 „takového“ namiesto také). Niektoré preklepy majú závažnejší charakter a príkladom je strana 17 kde je chybné uvedené „reálne čísla“ a správne majú byť racionálne čísla. Podobne je to aj na strane 22 kde autor uvádza počet možných funkcie $r(v,v)$ ako 10 a v predchádzajúcom texte jasne ukázal že daná funkcia môže mať iba 9 hodnôt. Rovnica 2.9 nemá definované všetky parametre. Viacnásobne sa v texte odkazuje na inú podkapitolu (odkaz na 5.2.6 pravdepodobne mal byť na 5.3). Ďalším príkladom nedôslednosti sú obrázky. Obrázok 30 ukazuje ukončenie behu algoritmu pri 155.000 vyhodnotení aj keď autor v texte uviedol že všetky experimenty mali nastavený strop 120.000 vyhodnotení. Obrázok 27 neobsahuje krivku pre 3

mutácie ale pri opise obrázku v texte spomína 3 mutácie, čo pravdepodobne mali byť 2 mutácie. Obrázky 27, 28 a 29 sú ťažko čitateľné, pretože autor použil nevhodné typy a farby jednotlivých kriviek. Príkladom je obrázok 28 kde autor použil plnú žltú čiaru, bodkovanú zelenú farbu, a plnú čiernu farbu dva krát pre rôzne krivky aj keď mohol použiť iné vhodnejšie kombinácie. Tieto mnohé chyby znižujú čitateľnosť a celkovú úroveň dizertačnej práce.

Otázky do rozpravy k dizertačnej práci:

- 1) V texte na strane 75 bolo uvedené, že presnosť výpočtu pri použití pevnej rádovej čiarky je nižšia v porovnaní výpočtu s pohyblivou rádovou čiarkou. Toto tvrdenie nieje úplne korektné. Vysvetlite v akých konkrétnych príkladoch je tvrdenie korektné.
- 2) Na obrázku 44 je ukázaná architektúra s paralelným vyhodnotením AdaBoost častí. Ako je riešená vstupná pamäť v tomto prípade? Obrázok 44 naznačuje použitie vstupnej pamäte v každej časti čo naznačuje nízku efektivitu alokácie pamäte.
- 3) V texte odznelo opísanie bloku ASUM ako jeden z možných miest kde môže nastať saturovanie priepustnosti pri veľkom počte slabých klasifikátorov. Zaujímalo by ma aké sú limity pre danú architektúru. Koľko slabých klasifikátorov môžem použiť pri zachovaní rýchlosti a bez nutnosti rozdeľovať klasifikátor na časti?
- 4) V texte som nepostrehol informáciu o limitoch bloku ASUM. Zaujímalo by ma či je možné použiť ten istý blok ASUM pri výpočte vstupného obrazu s rôznym nižším rozlíšením. Ak nie, je možné tento blok rozšíriť pre takúto úlohu? Ako by sa to mohlo prejaviť na spotrebe hardvérových zdrojov a priepustnosti?
- 5) Porovnanie novej architektúry RT-MPOD a sekvenčnej architektúry od Zemčík a kol. 2013 bolo urobené bez zarátania výrazne rozdielnych hardvérových zdrojov. Sekvenčné riešenie môže byť jednoducho paralelizované zdvojením využívajúc rozdelenie vstupného obrazu na polovicu a ich čiastočné prekrytie. Skúste porovnať priepustnosť keď využijete túto možnosť pri blízkej spotrebe hardwarových zdrojov.
- 6) V texte bol opísaný automatizovaný nástroj na generovanie jadra AdaBoost klasifikátora. Je tento nástroj verejne prístupný? Ak nie, plánujete tak urobiť?
- 7) Nová architektúra klasifikátora využíva 8 bitový vstup. Aký dopad na presnosť by mohlo mať rozšírenie na 3 farebné kanály s vhodnou voľbou farebného štandardu? Skúšali ste experimentovať s farebným vstupom?

Záverom konštatujem, že predložená písomná práca Ing. Filipa Kadlčeka spĺňa obecné uznávané podmienky na dizertačnú prácu a predstavuje prínos k rozvoju vedy a techniky v danej oblasti. Tézou dizertačnej práce považujem za dizertabilnú. Predloženú prácu doporučujem k obhajobe a navrhujem udelenie akademickej hodnosti philosophiae doctor (PhD.).

V Bratislave, 3. 3. 2017

Ing. Peter Malík, PhD.
Ústav informatiky SAV
Dúbravská cesta 9
845 07 Bratislava

