

Posudek oponenta diplomové práce

Diplomant: Bc. Michal Wertheim

Téma předložené práce je „Zpracování obrazu v systému Android - odečet hodnoty plynoměru“. Jako první věc je potřeba podotknout, že elektronická a tištěná verze se minimálně formálně rozcházejí. Za další je také nutné zmínit, že práce je znovu přepracována z minulého ročníku. Z práce vyplývá, že zadání diplomant formálně splnil v bodech 1,2,4. Avšak 5 bod zadání (Ověření funkcionality a spolehlivosti) dostatečně nesplnil (vysvětlení v závěru posudku). Splnění části bodu 3 zadání nebylo možné ověřit, jelikož se nepodařilo zprovoznit Neuroph část aplikace. Část využívající knihovnu Tesseract OCR nevykazovala výstupní číslo s desetinnou částí a v práci není korektní příklad výstupu uveden.

Diplomant projevil jisté znalosti v problematice vytváření programu pro zařízení Android a využití dalších knihoven (Tesseract OCR, Neuroph). Kladem jsou autorem zprovozněné knihovny Tesseract OCR a pravděpodobně i knihovna Neuroph v programu pro zařízení Android. Tento klad je ale snížen příliš jednoduchým a nešťastně rozvrženým GUI a nestabilitou aplikace (viz sekce Aplikace).

Formální stránka práce:

Občasná pravděpodobně neintencionální změna velikosti písma, časté překlapy či nevhodné koncovky (1-2 na stránce) spolu s častým netechnickým či zavádějícím vyjadřováním a občasným vynecháním slova ve větě značí nedostatečnou korekci formální stránky práce.

Některá vyjádření autora by si zasloužila revidovat, jelikož nejsou technické či vypovídající. Za všechny uvádím "Metody, které se při zpracování obrazu používají, jsou většinou zpracovány jako dvourozměrný signál, který je kalkulován známými algoritmy ze signal-processingu." (str29) a "Metoda segmentace, jež je jednou z nejbližších ke vnímání lidského oka, je barevný model HSV." (str30).

Všechny rovnice jsou nestandardně referencovány. Rovnice jsou označeny popiskem stejně jako obrázků, namísto standardního čísla rovnice na okraji stránky (Např. str.28, vztah 1). Uvedená rovnice ovšem není rovnicí v pravém slova smyslu, protože neobsahuje znaménko rovná se. Jedná se pouze o "vztah pro výpočet checksumu", jak je popsáno v předchozím odstavci práce. V této rovnici je taktéž označena suma skrze iterační proměnnou n v rozsahu od 1 do n , což není z hlediska dvojího použití proměnné n matematicky korektní zápis, i vzhledem k textové vysvětlivce u tohoto vztahu.

Samostatnou kapitolou jsou obrázky. Obrázek 14 má scestný a netechnický popis grafu Gaussovy funkce s nepopsanými parametry. Je možné diskutovat, zda u projekcí binárních obrazů tedy sloupcových grafů mají být popisky os, ale dle mého názoru ano (obr. 14,15,25,41,42,44,45). Necentrování obrázky (14,15,50). Obrázky, jež mají zobrazovat určitou podoblast ohraničenou čarou, trpí vždy volbou příliš úzké šířky této linie, nebo nevhodně zvolenou barvou, či velikostí obrázku (23,29,38,42). U obrázku 18 nejsou vyznačeny popisky referencované v textu. Převzatý obrázek 21 je nevhodně zvolen vzhledem k velkému počtu obsažených versus popsáných informací.

Citace uvedené v Bibliografii jsou číslovány bez použití čísel v závorce, což je v rozporu s vyhláškou.

Celkově práce vykazuje přespříliš formálních chyb, ať už po textové nebo grafické stránce.

Obsah a členění práce:

První kapitola se zabývá rozбором systému Android a fungováním jednotlivých aplikací poměrně do hloubky, tato část byla mírně přepracována z minulého roku. Další kapitola stručně popisuje knihovnu OpenCV, použitý jazyk a IDE. Z hlediska logického členění by do této kapitoly měli patřit i popisy použitých knihoven Tesseract OCR a popřípadě frameworku Neuroph.

Ve 3 kapitole se autor rozepsal o samotné aplikaci, vytkl bych nepřesné popisy v sekci Aktivita Zpracování, kde taktéž obrázek 9 neodpovídá zpracovávajícímu řetězci v příložené aplikaci. V aplikaci jsou jednotlivé názvy přepsány a jsou doplněny o tlačítka Lokalizace BC a Čtení BC.

Kapitola popisující Zpracování obrazu rozebírá teorii čtení čárových kódů a jednoduše popisuje využití algoritmy a postupy pro zpracování obrazu plynoměru. Chybějící reference: pojem vycentrování v se v kapitole: 4.3.6 nenachází (str38). Kapitola Lokalizace čárového kódu je nešťastně situována mezi

kapitolami Detekce zobrazovače a Segmentace číslic plynoměru (str.40). Sekce Detekce hran příliš nekoresponduje s umístěním v kapitole Barevné prostory (str.35).

Teoretická kapitola Neuronové sítě začínající velmi obecně, shrnuje základní teorii neuronové sítě s jednou vnitřní vrstvou je nešťastně umístěna, až před samotný Závěr práce.

Celkově není členění práce v uvedených příkladech v logickém sledu.

Faktické chyby:

V sekci věnované lokalizaci čárového kódu je detekce ilustrována na obrázku obsahujícím 2 čárové kódy, avšak algoritmus detekuje pouze jeden z nich. Autor neuvádí proč (obr.32). Dále se diplomant při rovnání snímku podle dominantních linií v obraze odkazuje na jen ve speciálních případech splněné tvrzení, že jsou tyto dominantní linie „rovnoběžné“ s rovinou zobrazovače, aniž by charakterizoval podmínky pro pořízení fotografií. Jeho tvrzení pravděpodobně může platit pro zvolenou galerii obrázků, ne však už obecně například pro plynoměry zachycené s perspektivním zkreslením.

Aplikace:

Samotná aplikace neodpovídá kvalitativně očekávanému výstupu diplomové práce. Volba Nastavení naprosto nesouvisející s prací - dva textové vstupy nijak nepřispívají k řešení, autor uvádí, že jsou přítomny pro budoucí potřebu. Informace o aplikaci – Obsahuje název aplikace. Aplikace je celkově méně stabilní, při nestandardním vstupu danou akci android ukončí s nedefinovanou chybovou hláškou a aplikace se navrátí do menu.

Manuál ke zprovoznění aplikace v příloze splnil svůj účel, popisuje však instalaci nefunkčního software a popisovanou neuronovou síť se podle něj zprovoznit nepodařilo a způsobovala pád aplikace (Android 4.2.1). Rozložení GUI není příliš přehledné, ani výpovědní hodnota výstupů jednotlivých kroků zpracování obrazu (příliš malé obrázky).

Samotná lokalizace čárového kódu i lokalizace zobrazovače funguje nad přiloženou galerií obstojně, avšak následná segmentace a rozpoznání číslic pomocí Tesseract OCR nevykazuje už tak nízkou úspěšnost 60% uváděnou autorem ani z daleka. Čtení čárového kupodivu nevypisuje žádný výsledný kód, ačkoliv aplikace po zmáčknutí tlačítka chvíli pracuje a můžeme jenom spekulovat, zda autor výsledek pouze nezobrazil, tak či tak jej nelze ověřit.

Závěr:

Poslední odstavce v kapitole Neuronové sítě stručně popisuje realizaci použité sítě ve frameworku Neuroph, kde byla tato síť učena na 700 generovaných a 300 segmentovaných obrazech číslic. Autor dále nespecifikuje, zda uváděná přesnost 83% byla vyhodnocena nad trénovacími daty, nebo nad daty mimo trénovací množinu číslic segmentovaných z fotogalerie.

V závěru uváděná úspěšnost rozpoznání 85% není v práci dostatečně podložena žádnými daty a není dostatečně popsáno, jak a z čeho ji autor vypočetl. Jelikož se nepodařilo zpracování skrze neuronové sítě v aplikaci zprovoznit, nebylo možné autorem uváděnou úspěšnost potvrdit.

Diplomant sám navrhl segmentační část zpracování obrazu, pro klasifikaci čísel využil knihoven Tesseract OCR a neuronové sítě skrze Framework Neuroph. V konečném důsledku si však uživatel v aplikaci musí vybrat vizuálně lepší výsledky z nabízených kroků segmentace.

Celkově práce působí hektickým a nedopracovaným dojmem. Forma ani obsah předložené práce nesplňuje obecné požadavky na VŠKP a minimálně bod 5 zadání, tedy ověření funkcionality a spolehlivosti detekce a rozpoznání na reálných snímcích nebyl splněn. Nesplnění bodu 5 je jen logickým důsledkem nesplnění předchozího bodu 3, jelikož vypisovaná rozpoznaná čísla (pomocí Tesseract OCR) neobsahovala desetinou čárku a celková odborná úroveň řešení je velmi nízká.

Musím uzavřít, že předložená diplomová práce nesplňuje příslušné požadavky na VŠKP.

Klasifikace :

F/35

V dne :

.....

oponent

(jméno + podpis)