



# Posudek oponenta závěrečné práce

**Student:** Vojtěch Dalecký  
**Oponent práce:** Ing. Petr Pauš, Ph.D.  
**Název práce:** Automatické zpracování tištěné papírové účtenky za palivo  
**Obor:** Webové a softwarové inženýrství

**Datum vytvoření:** 27. 1. 2021

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – následující škálou 1 až 4:</i>
<b>1. Splnění zadání</b>	<b><u>1=zadání splněno,</u> 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno</b>
<i>Popis kritéria:</i> Posuďte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.	
<i>Komentář:</i> Předložená bakalářská práce se zabývá zpracováním obrazu, jehož cílem je z fotografie extrahovat užitečná textová data.  Práce splňuje bez výhrad všechny body zadání.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>2. Písemná část práce</b>	<b>80 (B)</b>
<i>Popis kritéria:</i> Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišený od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.	
<i>Komentář:</i> Textová část práce popisuje všechny důležité kroky, které bylo při analýze, návrhu a implementaci nutné provést.  Kapitola 1 definuje potřebné pojmy v oblasti zpracování obrazu včetně matematické formulace, což oceňuji. V kapitole 2 je rozebrána detekce textu ve fotografii a je zvolen detektor CRAFT. Zde by bylo vhodné citovat, odkud autor čerpá, že se jedná o jeden z nejlepších detektorů. Kapitola 3 pojednává o rozpoznání znaků pomocí knihovny Tesseract. Opět by bylo vhodné uvést, proč je zvolena právě tato. K následujícím kapitolám nemám výhrad. Jsou psané věcně a srozumitelně včetně dostatečné ilustrace pomocí obrazových výstupů jednotlivých částí algoritmu.  Kromě přechozích dvou výtek jsou zdroje dobře citovány.  Po grafické stránce práce odpovídá standardům ČVUT FIT. Výhrady bych měl ke gramatice, kde se často objevují chyby ve shodě podmětu s přísudkem (kap. 1.4, ř. 3; kap. 2.2.2, ř. 5; str. 21, ř. 3 a 4; kap. 6.2.1, předposl. řádek; kap. 7.1, ř. 1; kap. 7.2, ř. 1; kap. 7.5, ř. 3; Závěr, ř. 5,...), chybějící čárky ve větách, několik překlepů (diletace, lehvenstein, ...) a chybějící čárky za popisky obrázků.  Autor práce u využitých knihoven zmiňuje jejich licenci a řídí se jí.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>3. Nepísemná část, přílohy</b>	<b>95 (A)</b>
<i>Popis kritéria:</i> Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů	

**Komentář:**

Výsledkem práce je knihovna v jazyce Java sloužící pro rozpoznání údajů na účtenkách čerpacích stanic. Podle doložených výsledků v kapitole 7 je detekce ve většině případů úspěšná. Zde nemám výhrady.

*Hodnotící kritérium:*

*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*

**4. Hodnocení výsledků, jejich využitelnost**

95 (A)

*Popis kritéria:*

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

**Komentář:**

Knihovna vcelku úspěšně detekuje požadovaná data v účtenkách a bude dále využita společností Uniqway.

*Hodnotící kritérium:*

*Způsob hodnocení – nehodnotí se*

**5. Otázky k obhajobě**

*Popis kritéria:*

Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

**Otázky:**

- 1) Zvažoval jste i jiné metody pro zpracování textu, než jsou zmíněné CRAFT a Tesseract? Proč byly vybrány právě tyto?
- 2) Dokáže knihovna zpracovat účtenku real-time? Případně, jak dlouho zpracování jedné účtenky trvá?
- 3) Jak je řešeno, když rozpoznání dat selže?

*Hodnotící kritérium:*

*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*

**6. Celkové hodnocení**

90 (A)

*Popis kritéria:*

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.

**Text hodnocení:**

Práce splňuje ve všech bodech zadání a výsledkem je knihovna, která se solidní úspěšností detekuje požadovaná data. I přes výhrady ke gramatice a nezdůvodněné volbě knihoven CRAFT a Tesseract hodnotím stupněm A (90 bodů), jelikož ostatní části práce jsou provedeny dobře a knihovna je funkční a bude nadále využita.

Podpis oponenta práce: