

## Posudek oponenta diplomové práce

**Student:** Jiruška Adam, Bc.  
**Téma:** Mobilní aplikace pro automatický záznam šachové partie (id 22538)  
**Oponent:** Bobák Petr, Ing., UPGM FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**  
Zadání je obtížnější, vypracované řešení je průměrné.
- 2. Splnění požadavků zadání** **zadání splněno**  
Všechny body zadání jsou splněny.
- 3. Rozsah technické zprávy** **splňuje pouze minimální požadavky**  
Rozsahem se práce nachází u spodní hranice - závěr je uveden na straně 39 (cca 54 normostran).
- 4. Prezentací úroveň předložené práce** **65 b. (D)**

Abstrakt bohužel jen stroze opakuje zadání práce a neposkytuje žádné informace o dosaženém výsledku. První kapitola je nepřilíš vhodně pojmenována jako "Teorie". **Jednotlivé sekce na sebe hladce navazují**, avšak často se v sekcích vyskytuje úvod "co bude následovat, jak je kapitola dále členěna". **Text je členěn přehledně až do 4. úrovně nadpisů.**

Problematické jsou dle mého názoru sekce **popisující návrh a implementační detaily neuronových sítí**, ostatní sekce jsou bez větších připomínek.

Sekce 2.2 - Neuronové sítě pro klasifikaci obrazu - poskytuje **vágní a nepřesná tvrzení** týkající se popisu existujících datasetů a neuronových sítí, např.:

- popis datasetu ImageNet (str. 11): *"Každý obraz obsahuje tzv. bounding boxy a ty jsou označeny názvem objektu, který ohraničují. Celkový počet fotografií je přibližně 1 500 000."* - ImageNet je ve skutečnosti mnohem větší (přes 14 mil. snímků) a jen zanedbatelné procento tvoří snímky s bounding boxy.
- vágní popis architektury MobileNetV2 (str. 12): *"Oproti první verzi má tato verze mnohem větší úspěšnost předpovídání. Také je efektivnější, což by se u druhé verze dalo předpokládat."*, *"Práce se neuronovými sítěmi není vůbec jednoduchá."* - O kolik procent je úspěšnost lepší? V jakém smyslu je efektivnější? V čem tkví složitost?

Sekce 4.3.2 - Tvorba a trénování neuronové sítě - popis použitých/navržených architektur je nedostatečný.

- Není uvedeno zdali a případně jak jsou před vstupem do sítě vzorky **předzpracovány**,
- u žádné popisované sítě není uvedena **velikost vstupu**,
- chybí popis způsobu **inicializace vah**, použitý **optimalizátor**, learning rate, batch size a další důležité **hyperparametry**,
- popis navržených architektur dle textu **nelze replikovat** - např. u konvolučních vrstev není uvedena velikost, počet filtrů a stride (navíc skripty související s popisem a trénováním neuronových sítí **nejsou součástí elektronické přílohy práce**),
- obrázek 4.8. nenesou žádnou hodnotu - zobrazuje jen změň čar, kde nejsou popsány osy a k průběhům chybí legenda.

- 5. Formální úprava technické zprávy** **75 b. (C)**  
Technická zpráva je **dobře srozumitelná**. Po pravopisné stránce se v textu vyskytuje **menší množství chyb a překlepů**.

Po typografické stránce je formátování textu **v pořádku** až na drobné výjimky:

- v elektronicky odevzdané formě je na titulním listu chybně uveden semestrální projekt místo diplomové práce,
- dvojitě uvozkovy jsou nesprávně vysázeny obě nahoře a za uvozeným textem není vysázena mezera (např. "rozbitých"pixelů, str. 6),
- místo spojovníků jsou použity pomlčky (např. popis obrázků 2.2),
- rovnice nejsou číslovány a často za nimi chybí interpunkce pokud se jedná o konec věty (např. rovnice na straně 9),
- špatně vysázená mocnina na straně 37,

- obrázek 3.1 má malé rozlišení a příliš velké fonty ve srovnání s textem technické zprávy.

### 6. Práce s literaturou

70 b. (C)

Autor cituje **relevantní zdroje** související s řešeným problémem. V sekci 4.3.2 autor popisuje, že při průzkumu sítě narazil na "tipy", jak vytvořit "sofistikovanější" architektury, ale neodkazuje na žádný zdroj, kde poznatky načerpal. U výčtu datasetů v sekci 2.2.1 je využit **pouze jeden nepříliš kvalitní zdroj**, což se v důsledku negativně projevilo na faktické přesnosti textu této sekce.

### 7. Realizační výstup

65 b. (D)

Výstupem práce je mobilní aplikace pro automatický přepis tahů šachové partie, která kombinuje **množství nástrojů a technologií**. V prvním kroku je rozpoznána a rektifikována šachovnice za pomoci standardních funkcí z knihovny OpenCV. **Úspěšnost rozpoznání šachovnice je 93 %**, velmi však závisí na správném zarovnání šachovnice dle obrysu na obrazovce zařízení. Následně jsou extrahována jednotlivá políčka a je provedena klasifikace typu figurky pomocí neuronové sítě (Keras, TensorflowLite). **Úspěšnost klasifikace typu figurky je 79 % pro jedno políčko šachovnice, úspěšnost rozpoznání provedeného tahu na celé šachovnici je však minimální.**

Uživatelské rozhraní je **jednoduché ale funkční**. Osobně mi přijde v rozhraní matoucí a nešťastné, že po stisknutí tlačítka "Začít nahrávat" se nahrávat **nezačne** a je třeba nahrávání **iniciovat dalším tlačítkem** na další obrazovce. Autor dále popisuje, že provedl testování uživatelského rozhraní na vzorku **17 osob**. Prvním úkolem v testovacím scénáři bylo **zobrazit nápovědu**. Toto bohužel, není velmi reprezentativní zahájení scénáře. Uživatel mobilní aplikace po prvním spuštění určitě nemá a nechce hledat nápovědu. (Pokud tomu tak ve skutečnosti je, tak to s největší pravděpodobností svědčí o špatně navrženém UI/UX.)

### 8. Využitelnost výsledků

Vzhledem k minimální úspěšnosti klasifikace všech políček aplikace není schopna spolehlivě přepisovat provedené tahy. Velkým přínosem se tak stává především vzniklý **dataset o velikosti 8500 vzorků**, bohužel však není veřejně dostupný a není ani součástí elektronické přílohy diplomové práce.

### 9. Otázky k obhajobě

1. Byla data před vstupem do sítě nějak předzpracována? Jaká je předpokládaná velikost vzorku na vstupu sítě?
2. Experimentoval jste u zkoumaných architektur s augmentací dat, optimalizátorem, případně learning rate a dalšími hyperparametry? V textu zmiňujete, že jste vytvořil přibližně 100 různých typů sítí a každou trénoval třemi různými způsoby. Čím se architektury lišily a o jaké tři způsoby trénování se jednalo?
3. Je ve Vašem řešení u predikce tahů využita znalost o pravidlech šachu jak zmiňujete v návrhu systému? Pokud ano, jak a kde pravidla aplikujete?
4. V textu letmo zmiňujete, že jste vytvořil model pro klasifikaci prázdného a obsazeného políčka. Jak tento model využíváte v navrženém systému?

### 10. Souhrnné hodnocení

68 b. uspokojivě (D)

Realizačním výstupem práce je **komplexní inženýrské dílo na průměrné úrovni**, nicméně vzhledem k výše zmíněným prezentačním a formálním nedostatkům navrhuji hodnocení stupněm **D**.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 18. června 2020

Bobák Petr, Ing.  
oponent