

# Posudek oponenta diplomové práce

Autor práce: **Bc. Lukáš KÖLBL**

Název práce: **Automatic Image Classification**

## **Jazyková a grafická úprava**

Průměrné

## **Formální a obsahová stránka práce**

Podprůměrné

## **Vhodnost použitých metod**

Průměrné

## **Způsob zpracování a vyhodnocení**

Průměrné

## **Správnost získaných výsledků**

Průměrné

## **Vlastní přínos**

Průměrné

## **Doplnění hodnocení, připomínky:**

Práce se zabývá problematikou klasifikace obrázků z fotopastí do celkově 8 tříd. Data a způsob vyhodnocení jsou poskytnuty třetí stranou v podobě cvičné úlohy. Svou obtížností je vhodná pro středně pokročilé v oblasti strojového učení a počítačového vidění. Student použil obvyklé metody a dosáhl výsledků, které se v soutěži projeví 5. místem. Vzhledem k použité metrice vyhodnocení nelze přesně kvalitu tohoto výsledku hodnotit. Bylo by zajímavé kdyby student v práci interpretoval dosažené výsledky v podobě pravděpodobnosti určení správné třídy. Celkově je práce průměrná, místy až podprůměrná a působí dojmem, že nestihla projít finální kontrolou. Název práce je příliš obecný a měl by více reflektovat co se skutečně řešilo. To samé platí pro zásady pro vypracování. Angličtina je na slušné úrovni, jenom některé pojmy a fráze jsou nevhodné. Formální stránka pokulhává a je vidět, že student nemá zkušenosti s psaním odborných textů a během psaní DP je nenabyl. U referencí na obrázky, tabulky a citované práce chybí jejich typ, nebo jsou nevhodně uvedeny. Práce je rozdělena do kapitol nekonzistentně. Např. v praktické části jsou uvedeny teoretické pasáže. Některé pojmy nejsou vysvětleny (např. MDv4). Převzaté obrázky nejsou citovány. Tabulky a grafy někdy obsahují nedůležité informace a nebo naopak chybí důležité informace. Student použil vhodné architektury neuronových sítí pro daný problém. Ocenil bych u modelu Vision Transformer detailnější informace (popis self-attention a princip multi-head attention). Byly zvoleny vhodné metody augmentace ale opět mi chybí detaily jejich použití - např. v jakém pořadí se nasadily a jaký měly jednotlivé augmentace vliv na výsledek. U rovnice 4.8 chybí vysvětlení parametru  $p_t$ . Je zřejmě zásadní, protože v následující části je dle mého názoru chybně interpretován parametr  $\alpha_t$ , jako skalár, přičemž by se mělo jednat o vektor jehož prvky reprezentují významnost jednotlivých tříd. K práci patří zdrojový kód publikován na platformě github. Zde bych rovněž očekával vyšší kvalitu prezentace kódu a vyvětlení jeho použití a funkcionality. Student prezentuje dosažené výsledky pomocí reportovacího systému WandB, ovšem uvedené grafy poukazují na nestandardní chování trénování, kdy se hodnota ztrátové funkce pro validační data od počátku zvětšuje. Navzdory uvedeným nedostatkům si myslím, že práce má své kvality a student předvedl, že je schopen řešit problém klasifikace obrazových dat pomocí metod hlubokého učení.

## Dotazy

- Jak byly obrázky normalizovány? Jaké jsou standardní metody normalizace obrázků pro neuronové sítě?
- Proč jsou data z fotopastí tzv. single label? Neměly by od principu být multi-label? Vysvětlete rozdíl mezi těmito variantami a popište, jak se musí modifikovat architektura sítě, aby bylo možné řešit multi-label klasifikaci.
- Lze použít ztrátovou funkci kategoričké křížové entropie pro úlohu multi-label klasifikace?

## Splnění bodů zadání

úplně

## Doporučení k obhajobě

ANO

**Hodnocení: 2 - Velmi dobře**

V \_\_\_\_\_ dne \_\_\_\_\_

-----  
Ing. Marek Hruží, Ph.D.