POSUDEK OPONENTA ZÁVĚREČNÉ PRÁCE

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Moderní metody robustního strojového učení
Jméno autora:	Pavel Jakš
Typ práce:	diplomová práce
Fakulta:	Fakulta jaderná a fyzikálně inženýrská (FJFI)
Katedra:	Katedra matematiky
Oponent práce:	Doc. Ing. Tomáš Kroupa, Ph.D.
Pracoviště oponenta práce:	Katedra počítačů FEL ČVUT

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání **průměrně náročné**

Hodnocení náročnosti zadání závěrečné práce.

Pokyny 1. – 5. pro vypracování jsou velmi jasně specifikovány. Diplomant dostal za úkol nastudovat relevantní literaturu (metriky pro obrázky a základní metody adversariálního strojového učení), implementovat a porovnat hlavní metriky vizuální podobnosti a využít naprogramovaných metod ke tvorbě adversariálních vzorků. Protože se jedná vesměs o rutinní implementaci známých technik, hodnotím zadání jako průměrně náročné.

Splnění zadání **splněno**

Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.

Diplomová práce splňuje zadání bez výhrad. Autor uvedl práci přehledně sepsanou motivací. Metriky jsou jasně definovány, jejich implementace popsána a porovnání vyhodnoceno. Kód je veřejně dostupný.

Zvolený postup řešení **vhodný**

Posuďte, zda student zvolil správný postup nebo metody řešení.

Postup řešení odpovídá moderním technikám používaným v oblasti počítačového vidění ke tvorbě adversariálních obrázků.

Odborná úroveň **výborná**

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

Diplomant zužitkoval teoretické znalosti nabyté studium odborné literatury, dokázal využít navržené podklady k sepsání pěkného úvodu do adversariálního strojového učení a výpočetně efektivní implementace metrik.

Formální a jazyková úroveň **průměrná**

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Práce je psána matematicky srozumitelným stylem, s definicemi a větami se v textu zachází korektně, použitá notace je definována a předpoklady platnosti tvrzení vždy zmíněny. Samotný text práce však trpí některými nedostatky, kvůli kterým je práce místy hůře čitelná. Některé důležité detaily nejsou v práci jasně definovány (např. popis normování 1.-2. na str. 13, ve vzorci (3.7) chybí =C(y)) a důležitý odstavec 3.3 obsahuje nepřesné informace nebo chyby (např. v první větě odstavce má být „různou třídu“ místo „stejnou třídu“, souřadný popis dvojic přístupů dále je zmatečný, protože druhá dvojice je podřazena prvnímu kritériu). Práce také obsahuje velké množství gramatických chyb.

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

POSUDEK OPONENTA ZÁVĚREČNÉ PRÁCE

Výběr zdrojů, korektnost citací	**výborné**
Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.	
Práce cituje mnoho relevantních prací z oblasti computer vision a ML. Výběr zdrojů i korektnost citací hodnotím jako bezproblémové.	

Další komentáře a hodnocení
Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.
Diplomová práce je zajímavým příspěvkem k problematice výpočetní efektivity generování adv. obrázků na základě různých metrik. Oceňuji pečlivě sepsaný teoretický úvod do problematiky i s důkazy v kapitole 7, kde autor použije konceptuálně správnou definici Wassersteinovy metriky pomocí regulárních Borelovských (Radonových) měr. Implementační část práce je vhodně okomentována a vysvětlena, veřejně dostupný repozitář na github je plusem. Tabulka 5.1 a další dokládají šíři experimentů, které autor provedl a přesvědčivě vypovídají o výpočetní náročnosti výpočtu studovaných metrik nad obrázky.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.
Diplomovou práci hodnotím i přes uvedené formální nedostatky velmi pozitivně.

Návrh otázek pro obhajobu:
1. V úvodu kapitoly 3.4 se hovoří o robustním strojovém učení takto: „Na tento fenomén potom odpovídá robustní strojové učení, které se ve své podstatě snaží při učení neuronové sítě danou neuronovou síť naučit tak, aby, pokud možno, k existenci adversariálních vzorků nedocházelo.“ Podle mého názoru je to velmi nepřesný popis tématiky robustního učení. Jak si představujete, že zaručíte, aby k „existenci adversariálních vzorků nedocházelo“?
2. Jak byste řešil úlohu (3.10) pomocí známých numerický metod? Znáte nějaké metody pro hledání řešení min-max problémů?
3. Kterou metodu generování adv. obrázků byste si zvolil jako útočník s velmi omezenou výpočetní kapacitou a důrazem na rychlost generování?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře.**

Datum: 19.1.2025 Podpis: