

Oponentní posudek diplomové práce Bc. Lucie Hromádkové „Klasifikace lučních porostů v Krkonoších s využitím leteckých hyperspektrálních dat a s pomocí vector machines klasifikace“

Předložená diplomová práce rozvíjí téma využití algoritmů strojového učení pro klasifikaci hyperspektrálních dat za účelem mapování lučních společenstev v Krkonošském národním parku (KRNAP). Navazuje tak na předcházející výzkumné aktivity katedry aplikované geoinformatiky a kartografie PřF UK v Praze. Práce má dva hlavní cíle – testování vlivu vstupních parametrů metody podpůrných vektorů (SVM) na přesnost klasifikace a testování různých schémat výběru trénovacích pixelů na přesnost klasifikace. Dílčím cílem je pak porovnání výsledků klasifikace algoritmy neuronových sítí (NN) a SVM.

Práce má logickou strukturu, cíle jsou jasně definovány. Autorka věcně a přesně popisuje východiska jednotlivých problémů, v dostatečném rozsahu cituje odbornou literaturu, podrobně popisuje vlastní řešení a výsledky, které diskutuje s literárními zdroji. Výběr literatury je vhodný. Výjimkou jsou některé pasáže objasňující např. transformace PCA a MNF, kde bych doporučila odkázat na dostupné knižní zdroje a ne na dokumentaci softwaru ENVI. Po formální, jazykové i grafické stránce se jedná o velmi kvalitní dílo. I přes svůj rozsah, který je dán množstvím prováděných experimentů, je text čtivý a srozumitelný.

Všechny testy jsou prováděny systematicky a jejich design i výsledky jsou přehledně dokumentovány. Za nejpřínosnější považuji část týkající se velikosti a uspořádání trénovacích množin. Může posloužit jako praktický návod pro plánování budoucích pozemních kampaní nejen v KRNAP. Otázkou do diskuse je metoda výběru náhodných pixelů na hranici trénovacího polygonu. Myšlenka vychází z publikací Foody a Mathur (2004, 2006). Co v práci postrádám je ověření vazby mezi hraničními pixely ve smyslu geografických hranic a hraničními pixely v příznakovém prostoru. Oceňuji, že výběr trénovacích pixelů pro jednotlivá schémata byl proveden dávkově na základě autorkou vytvořených skriptů v jazyce Python.

Porovnání vlivu hodnot parametrů algoritmu SVM na přesnost klasifikace vedlo k závěru, že jako nejvhodnější se jeví použití transformační (jádrové) funkce vyjádřené polynomem 6. stupně a povolení maximální přípustné chyby („penalty parameter“). To evokuje otázku, kam by vedlo další zvyšování těchto parametrů?

Výsledky klasifikací SVM a NN jsou hodnoceny na základě chybové matice vůči kontrolním datům a lze je porovnat vizuálně z přiložených map. Bylo by zajímavé doplnit i kontingenční tabulku číselně vyjadřující shodu (a rozdíly) v obou klasifikacích.

Jako nedostatek práce, který ale dle popsaných skutečností není zcela zaviněn její autorkou, je nedostatečná geometrická kvalita výsledku terénního průzkumu. Je těžko pochopitelné, proč nebylo k určení hranic polygonů jednotlivých tříd použito měření GNSS. Z tohoto pohledu je třeba ocenit autorčinu snahu tuto skutečnost alespoň částečně napravit dodatečným průzkumem, který se kladně projevil ve výsledcích navštívených oblastí Přední Rennerovky a Friesovy boudy.

Diplomová práce Bc. Lucie Hromádkové z hlediska formy a rozsahu zpracování vysoko přesahuje požadavky kladené na tento typ prací. Proto ji **doporučuji** k obhajobě a navrhuji hodnocení stupněm „výborně“.