

## **Optimalizace tvorby trénovacího a validačního datasetu pro zvýšení přesnosti klasifikace v DPZ**

Autorka: Bc. Barbora Chytilová

Vedoucí práce: RNDr. Lucie Kupková, Ph.D.

Předložená diplomová práce řeší často diskutované a stále aktuální téma ověření přesnosti výsledků klasifikace obrazových dat dálkového průzkumu. Navazuje zejména na práce Foodyho a Mathura publikované v letech 2004 až 2009 zabývající se počtem validačních bodů, poměrem trénovacích a validačních bodů a vlivu velikosti těchto množin na přesnost vybraných typů klasifikace. V souladu s literaturou si diplomatka vytyčila tři cíle, a to i) zjistit vliv podílu trénovacích a validačních pixelů na přesnost klasifikace maximum likelihood (ML), ii) sledovat vliv velikosti validační množiny na přesnost klasifikace při daném počtu trénovacích dat, iii) ověřit, zda pro dosažení téže přesnosti klasifikace postačí pro metodu support vector machine (SVM) nižší počet trénovacích pixelů než v případě metody ML. Pro klasifikaci diplomatka zvolila snímek pořízený senzorem MSI družice Sentinel-2. Validační data pořídila vizuální interpretací a vektorizací ortofota s rozlišením 0,25 m. Pracovala s 1 000 pixely, které v různém poměru, popř. počtu řadila mezi trénovací a validační množinu. Každý výpočet byl prováděn 1 000 krát stratifikovaným výběrem daného počtu bodů. Tím mohla být pro každou kombinaci trénovacích a validačních dat vypočtena i směrodatná odchylka použitých charakteristik přesnosti, konkrétně kappa indexu, celkové, uživatelské a zpracovatelské přesnosti. Cíle práce i design celého experimentu považuji za vhodně zvolené a svojí náročností odpovídající diplomové práci.

Diplomová práce má osm kapitol a standardní členění. V úvodní části autorka seznamuje s motivací a cíli práce. Kapitola „Literární rešerše a úvod do problematiky“ v kostce objasňuje principy a metody dálkového průzkumu včetně hodnocení přesnosti klasifikace. Vzhledem k zaměření práce by bylo možné některé pasáže výrazně zkrátit (např. objektová a neřízená klasifikace). Následující kapitola, která je rešerší prací zabývajících se metodami výběru trénovacích a validačních dat, by z formálního hlediska mohla být zařazena do kapitoly předcházející. Z obecného hlediska je literatura vhodně zvolena a řádně citována a počet zdrojů odpovídá diplomové práci. Co však nelze přehlédnout a co výrazně snižuje kvalitu předloženého textu, jsou nevhodná nebo nelogická vyjádření, nevhodné překlady, záměna slov. V kapitole popisující statisticky komplikovanější metody hodnocení přesnosti klasifikace není vždy text pro čtenáře přehledný. Samostatným fenoménem je přítomnost jednotlivých z kontextu zcela vytržených slov evokujících otázku, zda se jedná o test pozornosti hodnotitele. Příklady výše uvedených výték:

- Str. 16 poslední řádek: „Rovnoměrnost prostoru ...“.
- Str. 21 předposlední odstavce: „Průměrový vektor v tomto shluku popisuje míru úrovně.“ Co je míra úrovně?
- Str. 23 nahoře: „Přesnost výsledků je nutné hodnotit vždy s ohledem na polohu.“ ... z textu nevyplývá logika věci, možná by pomohl příklad.
- Str. 23 úvod třetího odstavce opakuje odstavce druhý. Co jsou měřítka přesnosti? Je celková přesnost „opatřením“?
- Str. 25 v prvním odstavci jsou pomíchány pojmy hodnocení přesnosti a klasifikace.
- Str. 28 poslední odstavce sekce 2.1 – slovo „literatura“ bez kontextu.
- Str. 28 Co je střední matice?
- Str. 29 Jaký je vztah přesnosti klasifikace a veličiny „Power“ v obr. 3?

Následující kapitola stručně popisuje zájmové území. Jedná se o krajinu v podhůří Krkonoš s mozaikou lesních, zemědělsky obdělávaných ploch a zástavby. Vzhledem k heterogenitě krajinného pokryvu jde z hlediska klasifikace a hodnocení její přesnosti spíše o výzvu. Bylo by zajímavé provést stejný test v území, které je z hlediska krajinného pokryvu a jeho plošné roztržitosti homogennější. Další kapitola seznamuje s použitými daty a definuje třídy klasifikace. Analýzou trénovacích dat pro klasifikátor ML se ukázalo potřebné sloučit některé kategorie a počet tříd se zredukoval z původních osmi na šest. V kapitole 4.5.1 je správně uvedeno, že klasifikátor ML předpokládá normální rozdělení trénovacích dat. Jedná se ale o

rozdělení dat uvnitř trénovacích tříd a nikoli o zastoupení tříd v obraze, jak je chybně smícháno v grafu 1 na str. 41! Pravděpodobnost výskytu dané třídy v obraze může být parametrem klasifikace ML. Vliv tohoto parametru mohl být v prováděných pokusech, na rozdíl o mnohých reálných situacích, také testován, avšak zřejmě nebyl. Obecná nastavení klasifikátoru ML nejsou v práci uvedena. Vlastní metodika a způsob implementace navržených výpočtů jsou popsány srozumitelně.

Kapitola Výsledky shrnuje výstupy prováděných výpočtů ve formě grafů a tabulek a slovního popisu. Obecně jsou obsáhlé a detailní, ale opět obsahují nepřesnosti a formální nedostatky, které jsou zřejmě důsledkem kvapného zpracování, např.:

- Str. 61 – „... a pro 25 trénovacích bodů byla již hodnota koeficientu nulová.“ Počet trénovacích bodů se ale v tomto testu přece neměnil, měl konstantní hodnotu 375, jak je uvedeno na str. 60.
- Str. 61 – věta „V grafu je vidět zlom ...“ z pohledu oponentky nevystihuje situaci znázorněnou v grafu 14. Hodnota  $\kappa$  koeficientu se zde pohybuje kolem hodnoty 0,6 a v závislosti na počtu validačních bodů se jeho hodnota mění v řádu setin. Nelze tedy mluvit o zlomu a následné velice nízké hodnotě, jak je uvedeno dále.
- Str. 62 – opět je zmíněn zlom v grafu 15, který ale není zásadní (bylo myšleno minimum pro hodnotu 400?), a dále je uvedeno, že pro počet validačních bodů menších než 150 je celková přesnost na velice nízké hodnotě. V grafu 15 je ale pro 50 a 100 validačních bodů celková přesnost klasifikace vyšší než pro 150 bodů! Stejně tak další tvrzení o růstu přesnosti klasifikace do maxima s narůstajícím počtem validačních bodů nemohlo vzniknout na základě grafu 15.
- Str. 75 – Poněkud klopotným způsobem je vysvětleno, že porovnání mezi klasifikátory bylo provedeno v intervalu 275 až 475 bodů s krokem 50. Proč jsou tedy v grafech 26 a 27 zobrazeny a v textu komentovány výsledky i pro 50 a 100 trénovacích bodů? Můžete vysvětlit, co rozumíte „zlomy“ v grafech 26 a 27?

Formální nedostatky:

- Počet desetinných míst uváděných v grafech.
- Pro všechny grafy byla použita jednotná šablona s popisem svislé osy „Hodnota přesnosti klasifikace“. Někdy se jedná o celkovou přesnost, jindy o uživatelskou či zpracovatelskou přesnost. Bylo by vhodné toto v grafech odlišovat (zejména jsou-li dvě charakteristiky přesnosti na téže stránce).

V kapitole Diskuze autorka hodnotí výsledky v kontextu již publikovaných prací. Některá tvrzení si protirečí. Z prvního testu např. vyplynulo, že optimální se jeví rozdělení dat v poměru 1/3 trénovacích a 2/3 validačních, což je v souladu s uvedenou literaturou (str. 79). Na str. 81 je ale uvedeno, že tento poměr nemusí být „univerzálně nevhodnější“. Dále se uvádí, že „svou roli bude hrát i počet trénovacích bodů, který v této práci testován nebyl.“ V prvním testu se ale pracovalo s různým počtem trénovacích bodů. Jednalo by se tedy o jiný test? Závěrem lze říci, že autorka podala řešení v úvodu vytyčených cílů. Výsledky prvního a třetího jsou v souladu s literaturou, komentář k výsledkům druhého testu je rozporuplný.

Diplomová práce Bc. Barbory Chytilové je zaměřena na praktické, potřebné a užitečné téma dálkového průzkumu. Praktické výstupy práce by po úpravě mohly být přetvořeny v zajímavou publikaci. Kvalitu práce snižuje textová část, která by si zasloužila nemalé množství úprav, jak naznačují připomínky a příklady uvedené výše. Předloženou diplomovou práci proto sice **doporučuji** k obhajobě, ale hodnotím stupněm **dobře**.

Otázka do diskuse: Jak mohl výsledky klasifikace ovlivnit fakt, že vizuální interpretace byla provedena nad ortofotem pořízeným zpravidla v jarních měsících a družicový snímek pochází ze srpna, resp. jak hodnotíte přesnost a spolehlivost výsledku vizuální interpretace?