

Testování způsobů klasifikace pokrytí území vybraných evropských oblastí

pana Bc. Tomáše Boučka

Téma diplomové práce je částí projektu INEA „*Geo-harmonizer: EU-wide automated mapping system for harmonization of Open Data based on FOSS4G and Machine Learning*“, jehož nositelem je katedra geomatiky. Cílem diplomové práce bylo otestovat družicová data, která jsou volně dostupná všem uživatelům a která budou sloužit pro vytvoření bezešvých map pokrytí území Evropy za období od začátku 90. let do současnosti, analyzovat způsoby klasifikace, a připravit tedy pracovní postup pro automatizované postupy těchto klasifikací metodami strojového učení.

Pro analýzu celého území Evropy byla vybrána tři území reprezentující tři zeměpisně a klimaticky odlišné oblasti, které by mohly v souhrnu představovat přibližně rozsah možných typů povrchů, s nimiž se v Evropě setkáme. Jednalo se o území jižního Švédska, středních Čech a území Itálie v oblasti Říma. Pro třídy klasifikace byl konsorciem projektu zvolen systém využívaný v projektu CORINE Land Cover. Tento systém dělí území do tří hierarchických úrovní podrobnosti. Třídy CORINE nepředstavují v mnoha případech třídy pokrytí území, ale využití území, což činí práci s nimi složitější. První krok byl zaměřen na nejméně podrobnou klasifikaci do pěti základních tříd – zastavěné oblasti, zemědělská území, lesy, mokřady a vodní plochy. Pro diplomovou práci i projekt byla zvolena data družic Sentinel-2.

Diplomant pečlivě zkoumal iteračním způsobem nejvhodnější výběr trénovacích ploch na základě všech úrovní tříd CORINE. Byl to právě pečlivý výběr ploch z CORINE dat, zároveň analýza využívající hodnoty normovaného vegetačního indexu a úpravy vybraných trénovacích CORINE ploch na základě výsledků zobrazených kontingenční tabulkou ukazujících sporná území mezi dvojicemi tříd, které mu pomohly dosáhnout vysokých přesností klasifikace trénovacích ploch pro všechna tři zvolená území (Švédsko 97 %, Čechy 95 %, Itálie 96 %) pro první úroveň CORINE, tedy pro pět tříd. Pro druhou úroveň klasifikace – celkem 15 tříd) - po pečlivých úpravách vybraných trénovacích ploch získal diplomant přesnosti rovné 93 %, 94 % a 95 % pro jednotlivé oblasti (Švédsko, Čechy, Itálie). Výpočtem klasifikací třetí úrovně CORINE – celkem 44 tříd pro celou Evropu se autorovi podařilo získat opět vysoké přesnosti klasifikace pro Švédsko 91 %, Čechy 95 % a Itálii 97 %. Autor zároveň určil, které třídy na studovaných územích klasifikovat nelze.

Diplomant se v závěru nad rámec své práce s klasifikací maximum likelihood věnoval krátce i strojovému učení. Výsledky, jichž dosáhl pro všechny tři úrovně detailu CORINE Land Cover dat, nelze fakticky považovat za ekvivalentní, neboť software GRASS, s nímž pracoval, nebyl schopen použít celou trénovací množinu, ale pouze omezený počet pixelů. Přesto jsou výsledky zajímavé, neboť se v tomto způsobu výpočtu projevil např. menší dopad oblačnosti na výsledek klasifikace, zároveň se zlepšila klasifikace vodního toku, jehož kvalita vody na vodní hladině se podstatným způsobem odlišuje od kvality vody na hladině u většiny relativně málo znečištěných ploch.

Autor se účastnil pravidelně jednání o daném projektu a části věnované otázce přípravy klasifikace a posouzení vlastností jednotlivých typů povrchů vybraných území ve srovnání s třídami CORINE Land Cover. Svým přístupem pomohl připravit metodiku, která bude sloužit k dalšímu kroku, a tím je klasifikace vyššího počtu scén z evropského území. Jeho metodika bude posléze použita po verifikaci a případných dalších úpravách pro strojové učení jako obecný nástroj pro klasifikaci pokrytí území Evropy pro různé roky.

K práci mám následující dotazy:

1. Klasifikace zástavby je jedním z problémů daného řešení, neboť zastavěné oblasti představují radiometricky a spektrálně odlišná území, což pro 2. a 3. úroveň CORINE tříd přináší problémy. Jaké by bylo diplomantovo doporučení pro řešení těchto typů území, které pro 1. úroveň CORINE dat tvoří jedinou třídu.
2. Je autor toho názoru, že bude možno řešit zemědělská území celé Evropy se zkušenostmi získanými v této práci? Jinými slovy, byl přístup k zemědělským plochám stejný pro všechna tři vybraná území?

Autor splnil zadání diplomové práce, svým aktivním přístupem podstatným způsobem připravil první významnou část dříve zmíněného výzkumného projektu. Za kvalitu získaných výsledků, celkový přístup k práci, a textovou i obrazovou část diplomové práce mu uděluji známku

A - výborně.



V Praze 18. 6. 2020

Prof. Ing. Lena Halounová, CSc.
Vedoucí DP