

Potencial prostorskih informacij hiperspektralnega oslikovanja za prepoznavanje lesnih vrst

Luka Kopač¹, Miha Humar¹

¹ Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za lesarstvo

E-naslov: luka.kopac@bf.uni-lj.si, miha.humar@bf.uni-lj.si

POUDARKI:

- Z uporabo transformacije glavnih komponent je mogoče pretvoriti HSI slike v RGB format in jih vdelati s pomočjo konvolucijske nevronske mreže v nabor informativnih značilnic, ki kodirajo teksturo lesnih vrst.
- Nenadzorovani modeli zmanjšanja dimenzionalnosti, kot sta t-SNE in hierarhično grupiranje na osnovi kosinusne razdalje, ne zagotavljajo dovolj velike moči ločevanja za posamične lesne vrste.
- Glede na teksturo se posamične vrste razvrščajo v večje skupine na osnovi anatomije (iglavci, difuzno porozni listavci ter venčasto porozni listavci), kjer je razločevanje bolj enostavno.
- Logistična regresija omogoča klasifikacijo na osnovi prostorske informacije HSI z dokaj visoko točnostjo – 85,2 %, kjer so napačne razvrstitve predvsem med anatomsko podobnimi vrstami.

UVOD:

Les je kot široko razširjen material z relativno enostavno obdelavo skozi celotno človeško zgodovino služil različnim namenom. Ker nastaja kot naravno tkivo različnih drevesnih vrst, ima lahko zelo raznolike lastnosti, odvisne od izvora in rastnih pogojev (Rowell, 2012). Izbira ustreznih lesnih vrst je zato ključna glede na predvideno uporabo. Zaradi občutljivosti lesa na biotske in abiotske dejavnike razkroja se je ohranil le manjši delež zgodovinskih lesenih predmetov. Za pravilno umestitev teh predmetov v zgodovinski kontekst je pomembno določiti lesno vrsto, kar prispeva tudi k postopkom konserviranja in restavriranja (Lo Monaco idr., 2018). Pri tem so posebej zaželeni neporušni in neinvazivni postopki, zlasti kadar gre za edinstvene in dragocene predmete. Napredne slikovno-spektroskopske tehnike, kot je hiperspektralno oslikovanje (HSI), v kombinaciji z algoritmi strojnega učenja, lahko pomembno prispevajo k takšnim prizadevanjem (Ma idr., 2019).

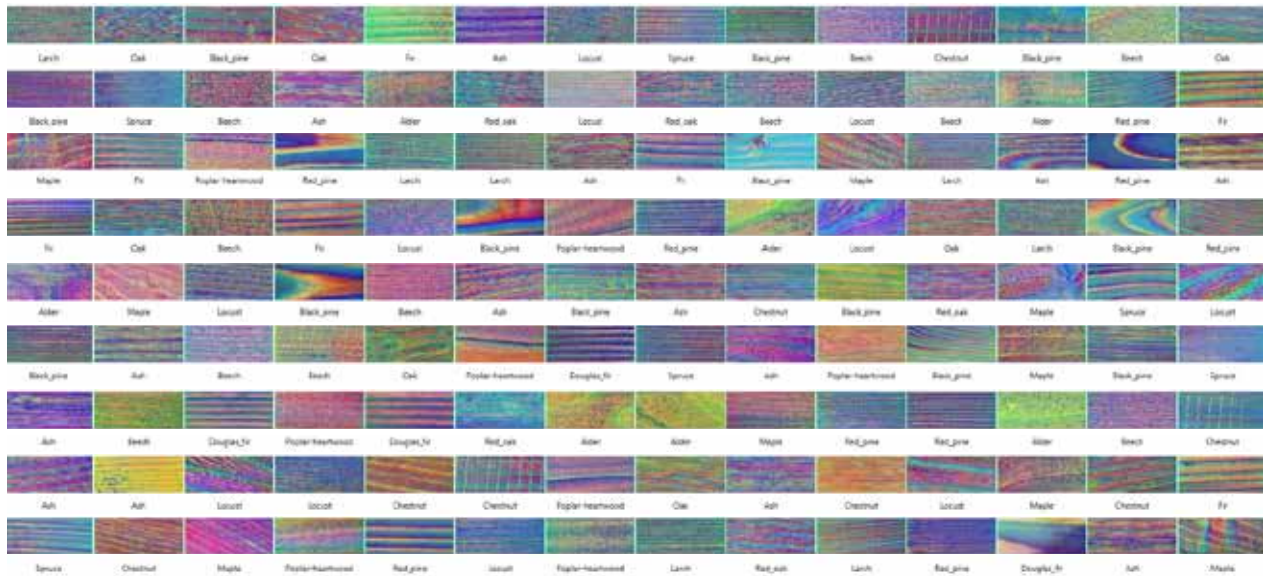
HSI združuje bližnje-infrardečo (NIR) spektroskopijo in slikanje, kar omogoča vpogled v prostorsko razločene kemijske informacije. Namesto običajne klasifikacije na osnovi spektrov je bil cilj te raziskave izkoristiti dopolnilne prostorske informacije, ki predstavljajo teksturo lesa.

Takšen pristop bi lahko izboljšal identifikacijo lesnih vrst na osnovi HSI pri kompleksnejših problemih.

PODATKI:

Na voljo smo imeli 45 vzorcev za vsako izmed 15 ključnih evropskih lesnih vrst, ki so ekonomsko najbolj pomembne v Slovenskem prostoru. Dve prečni površini (50 mm × 25 mm) vsakega vzorca sta bili uporabljeni v analizi. Zajem podatkov je bil izveden s potisno-skenirnim HSI sistemom (ClydeHSI, Clydebank, Združeno kraljestvo), ki vključuje dva volfram-halogen reflektorja, kamero za kratkovalovno infrardeče območje (SWIR) s spektralnim razponom 1000–2500 nm in povprečno spektralno ločljivostjo 5 nm ter premično mizo za vzorce. Odstranili smo ozadje nastalih podatkovnih kock, nato pa uporabili analizo glavnih komponent (PCA) za zmanjšanje dimenzionalnosti podatkov z izvornih 288 spektralnih pasov na 3 kanale, ki so bili osnova za RGB sliko. Druga glavna komponenta je predstavljala rdeči kanal, tretja zeleni kanal in četrta modri kanal. Opisana priprava podatkov je bila izvedena v programskem jeziku Python. Rezultat tega procesa je bilo 90 RGB slik na posamezno lesno vrsto (Slika 1).

Slika 1: Naključen podzorec uporabljenih RGB slik



METODE:

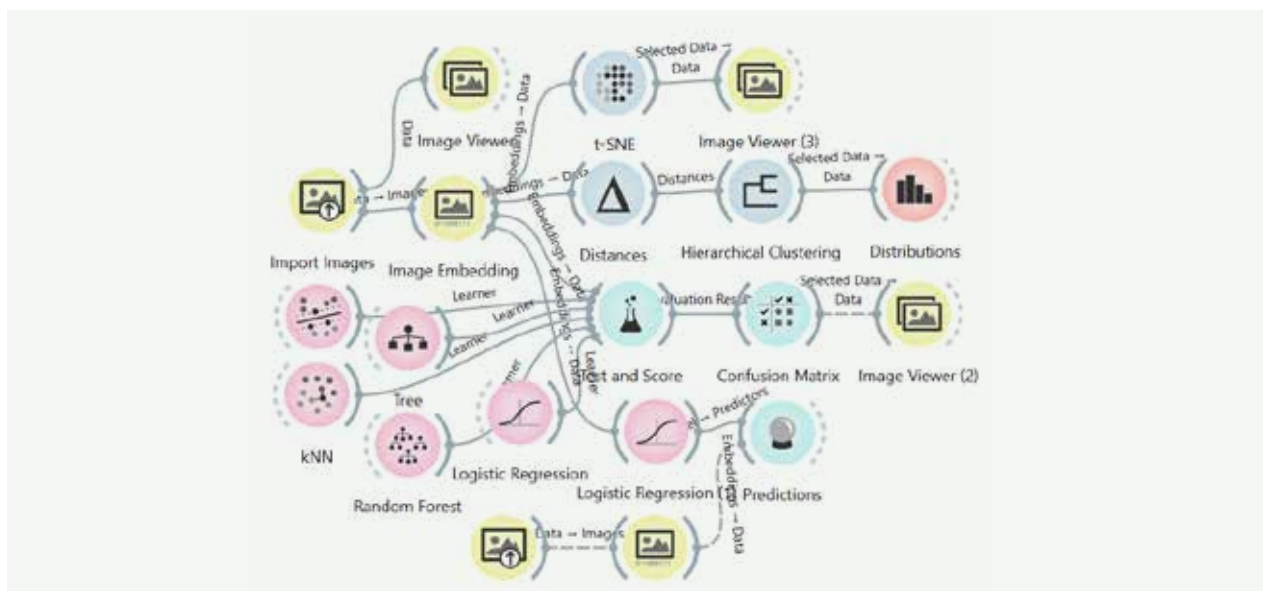
Za nadaljnjo obdelavo podatkov je bila uporabljena programska oprema Orange, namenjena podatkovnemu rudarjenju, strojnemu učenju in vizualizaciji podatkov. Slike so bile v Orange naložene z gradnikom »Import images« in preverjene z gradnikom »Image viewer«. Nato so bile posredovane skozi ključni sloj »Image embedding«, ki z uporabo konvolucijske nevronske mreže pretvori slike v nabor informativnih značilnic. V tej raziskavi je bil uporabljen model »InceptionV3«.

Za združevanje slik na podlagi novih značilnic je bila uporabljena metoda t-porazdeljenega stohastičnega vlaganja sosedov (t-distributed stochastic neighbour embedding – t-SNE). Vzporedno so bile izračunane kosinusne razdalje, ki

so bile nato uporabljene v gradniku »Hierarchical clustering«, rezultati pa so bili zaradi velikega števila slik vizualizirani s porazdelitvami po gručah dendrograma.

Za klasifikacijo je bilo uporabljenih več modelov strojnega učenja: podporni vektorski stroj (SVM), enostavno odločitveno drevo, naključni gozd, k-najbližjih sosedov (kNN) in logistična regresija. Njihova uspešnost je bila ocenjena z gradnikom »Test and score« z navzkrižno validacijo z 10 pregibi, za najboljši model pa je bila izdelana tudi matrika zmede. Logistična regresija je bila dodatno uporabljena za napoved vzorcev beljave topola, ki zaradi svoje teksturne podobnosti jedrovini iste vrste niso bili vključeni v učni nabor podatkov. Celoten potek dela v okolju Orange je prikazan na Sliki 2.

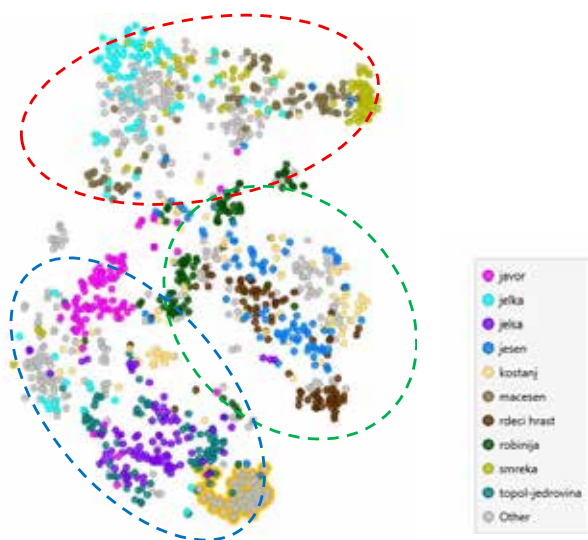
Slika 2: Proces dela v okolju Orange



REZULTATI:

Gruče večine lesnih vrst pri transformaciji v dvo-dimenzionalen t-SNE prostor, niso bile jasno ločene (Slika 3). Pri nekaterih vrstah, kot so bukev (siva, označena na sliki), javor (roza) in smreka (svetlo zelena), je bila pretvorjena teksturna informacija bolj edinstvena kot pri drugih vrstah. Predvidevali smo da te vrste omogočajo bolj enostavno razločitev od drugih. Kljub temu je ločevanje gruč sledilo anatomskim razredom lesnih vrst. Razvidni so bili trije glavni razredi: iglavci, difuzno porozni listavci in venčasto porozni listavci.

Slika 3: t-SNE transformacija v 2D prostor



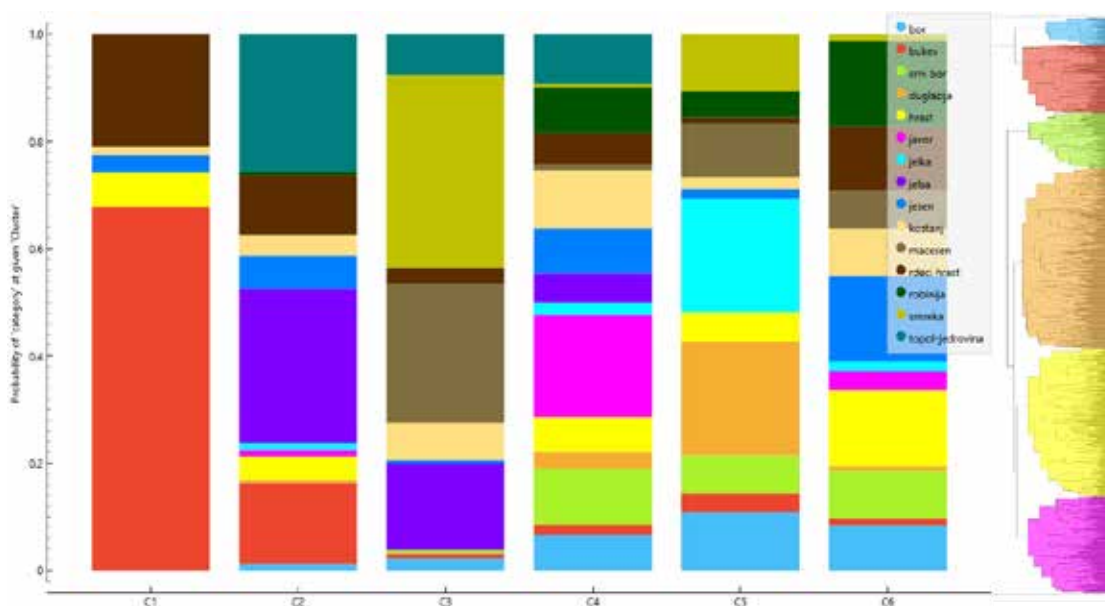
Iglavci (rdeča elipsa), kot so smreka, jelka in macesen, so se združevali bližje skupaj pri višjih vrednostih osi y na grafu t-SNE. Venčasto porozni listavci (zelena elipsa), kot so jesen, kostanj in rdeči hrast, so bili umeščeni v osrednji del, med iglavce in difuzno porozne listavce (modra elipsa), kot so javor, jelša in topol, ki so imeli nižje vrednosti osi y. Nekatere vrste so bile očitno teksturno bolj podobne vrstam iz drugih razredov. Tekstura javorja se je na primer izkazala za bolj podobno nekaterim venčasto poroznim vrstam kot difuzno poroznim vrstam.

Z uporabo dendrograma na podlagi kosinusnih razdalj med podatkovnimi vektorji in ustrezne mejne vrednosti smo razdelili podatke v šest gruč. Vsaka gruča je vsebovala več različnih vrst, ki so bile ponovno grupirane predvsem na podlagi anatomskih podobnosti v glavne razrede (Slika 4). V tem primeru so bile difuzno in venčasto porozne vrste listavcev pogosto v istih gručah in je razlikovanje izpostavilo predvsem razlike med iglavci in listavci.

Prva gruča (C1) je vsebovala predvsem bukev, ki ima dokaj izrazite teksturne značilnosti, kot na primer zrcala. Ta gruča je vključevala tudi hrast, rdeči hrast in nekaj primerkov jesena ter kostanja. Druga gruča (C2) je vključevala predvsem topol, jelšo in bukev. Tretja gruča (C3) je vsebovala iglavce, kot sta smreka in macesen, vključene pa so bile tudi nekatere vrste listavcev. Preostale tri gruče (C4, C5 in C6) so bile bistveno bolj raznolike po sestavi in so vsebovale tudi večje število vzorcev, kar je razvidno tudi na dendrogramu (skupaj 997 vzorcev v primerjavi s 353 v prvih treh gručah).

Za klasifikacijo na podlagi vdela ve slik je bilo preizkušenih več modelov. Najboljšo klasifikacijsko natančnost, 85,2 %, smo dosegli z uporabo logistične regresije, kar je bilo pričakovano, saj vdela va z nevronske mreže pripravi podatke posebej

Slika 4: Porazdelitev lesnih vrst po gručah (levo) in dendrogram (desno)



Slika 5: Matrika zmede - logistična regresija

	bor	bukev	crni bor	duglazija	hrast	javor	jelka	jelsa	jesen	kostanj	macesen	rdeci ...	robinija	smreka	topol...	Σ
bor	58	0	10	2	0	0	6	1	0	1	6	0	0	6	0	90
bukev	1	83	0	0	0	0	0	2	0	0	0	0	3	0	1	90
crni bor	9	0	71	3	0	0	2	0	0	2	3	0	0	0	0	90
duglazija	3	0	4	80	0	0	2	0	0	0	0	0	0	1	0	90
hrast	0	1	0	0	82	0	0	0	1	3	0	2	1	0	0	90
javor	0	1	2	0	0	83	0	2	2	0	0	0	0	0	0	90
jelka	8	0	2	2	0	1	60	0	1	0	3	0	0	5	0	90
jelsa	0	0	0	0	0	2	0	81	0	1	0	0	0	0	6	90
jesen	0	0	0	0	1	0	1	0	77	5	1	2	2	0	1	90
kostanj	1	0	1	0	2	6	0	0	4	69	0	0	6	0	1	90
macesen	4	0	2	3	0	0	1	0	0	1	77	0	0	2	0	90
rdeci ...	0	0	0	0	5	0	0	1	1	1	0	81	1	0	0	90
robinija	0	0	0	0	1	0	0	0	1	4	0	1	83	0	0	90
smreka	5	0	2	0	0	0	4	0	1	1	4	0	0	73	0	90
topol...	0	1	0	0	0	0	1	2	0	1	0	0	0	85	0	90
Σ	89	86	94	90	91	92	85	89	88	89	94	86	96	87	94	1350

za to vrsto klasifikacije. Matrika zmede za klasifikacijo z logistično regresijo je prikazana na Sliki 5.

Največ napačnih razvrstitev je bilo pri rdečem boru, jelki in kostanju, kjer je model nepravilno klasificiral več kot 20 vzorcev. Za najbolj teksturno razločene vrste so se izkazale bukev, duglazija, hrast, javor, jelša, rdeči hrast, robinija in topol, pri katerih je bilo pravilno razvrščenih 80 ali več od skupno 90 vzorcev. Večina napačnih razvrstitev se je pojavila med pričakovano podobnimi vrstami; na primer, nekateri vzorci rdečega bora so bili napovedani kot drugi iglavci (črni bor, duglazija, jelka, macesen ali smreka).

SKLEPI:

Raziskava je pokazala, da je za klasifikacijo lesnih vrst mogoče uporabiti dopolnilne prostorske oziroma teksturne informacije, ki jih zagotavlja hiperspektralno oslikovanje. Razvrščanje z uporabo metode t-SNE ali hierarhičnega združevanja na podlagi kosinusnih razdalj ni imelo dovolj ločilne moči za natančno razlikovanje posameznih vrst, čeprav je bilo ločevanje med iglavci, difuzno poroznimi listavci in venčasto poroznimi listavci bolj izrazito. Logistična regresija se je izkazala za uspešnejšo, saj je pri 10-kratni navzkrižni validaciji dosegla klasifikacijsko natančnost 85,2 %. Kot kaže matrika zmede, je bila večina vrst pravilno napovedana, pri čemer so se napačne razvrstitve večinoma nanašale na anatomsko podobne vrste. Kljub temu se je potrebno zavedati, da je tovrstna identifikacija lesnih vrst občutljiva na stanje površine, ki vključuje obdelavo, biotski ter abiotski razkroj in onesnaženost površine. V nadaljevanju bi bilo smiselno združiti modele klasifikacije na osnovi spektralnih in prostorskih podatkov hiperspektralne podatkovne kocke ter tako doseči višje točnosti identifikacije lesnih vrst v kompleksnih predmetih kulturne dediščine.

LITERATURA IN VIRI:

Lo Monaco, A., Balletti, F., & Pelosi, C. (2018). WOOD IN CULTURAL HERITAGE PROPERTIES AND CONSERVATION OF HISTORICAL WOODEN ARTEFACTS. In *European Journal of Science and Theology* (Vol. 14, Issue 2).

Ma, T., Inagaki, T., Ban, M., & Tsuchikawa, S. (2019). Rapid identification of wood species by near-infrared spatially resolved spectroscopy (NIR-SRS) based on hyperspectral imaging (HSI). *Holzforschung*, 73(4), 323–330. <https://doi.org/10.1515/hf-2018-0128>

Rowell, R. M. (Ed.). (2012). *Handbook of Wood Chemistry and Wood Composites* (2nd ed.). CRC Press. <https://doi.org/10.1201/b12487>

KLJUČNE BESEDE:

Hiperspektralno oslikovanje, identifikacija, tekstura lesa, lesne vrste

ZAHVALE:

Avtorji se zahvaljujejo za finančno podporo Slovenski raziskovalno inovacijski agenciji (ARIS) v okviru raziskovalnih projektov J7-50231 (GROWTH) in J7-50226 (NextGen HS), raziskovalnega programa P4-0015 (Les in lignocelulozni kompoziti) ter infrastrukturnih centrov (IC LES PST 0481-09 in IO-E012 E-RIHS). Prav tako se zahvaljujejo za podporo projekta Slo-Ita Interreg WoodInnovate, projektu MATRES (VIP Univerza v Ljubljani, ARIS, RSF-A, SN-ZRD/22-27/510) ter projektu Teaming CoE GreenHer.