

V4-2225 FD Gameplan

Delovni sklop 7

**Daljinsko zaznavanje z metodo
hiperspektralnega slikanja z letalom za zgodnje
odkrivanje okužb s FD**

Končno poročilo delovnega sklopa

Ljubljana, 24. 9. 2025

Poročilo pripravili:

Uroš Žibrat

Daša Donša

Andrej Vončina

Oddelek za varstvo rastlin

Kmetijski inštitut Slovenije

Kazalo vsebine

Uvod.....	3
Naloga 7.1 Hiperspektralno slikanje z letalom in brezpilotnim letalnikom.....	5
Izbor vinogradov.....	5
Letalsko slikanje.....	6
Obdelava hiperspektralnih slik.....	7
Naloga 7.2 Strojno učenje in validacija modelov.....	8
Metode dela.....	8
Rezultati in razprava.....	10
Predpriprava podatkov.....	10
Razvoj napovednih modelov.....	12
Vročinske karte	14
Zaključki.....	16

Uvod

Obstoječi pristopi za spremljanje zdravstvenega stanja vinske trte temeljijo predvsem na terenskih pregledih, ki so zaradi obsežnih preiskovalnih površin, omejenih virov ter časovnih in finančnih zahtev pogosto neučinkoviti. Glede na veliko površino vinogradov, kjer je prisotnih več sto rastlin, ki jih je treba v rastni sezoni večkrat pregledati, je tak način dela zamuden in naporen. Zato se pojavlja potreba po metodah, ki omogočajo hitrejša in bolj zanesljivo spremljanje zdravstvenega stanja rastlin, kot tudi možnost zgodnjega prepoznavanja okužb. Ena izmed takšnih metod je hiperspektralno slikanje, ki v primerjavi s klasičnimi pristopi omogoča določanje zdravstvenega stanja rastlin že pred razvojem vidnih znakov.

Ti sistemi uporabljajo vidni in bližnji infrardeči (400–1000 nm) ter kratkovalovni infrardeči (1000–2500 nm) spekter, kar omogoča zaznavanje sprememb v strukturi in kemiji rastlinskih listov, še preden te vplivajo na fenologijo rastlin. Hiperspektralni senzorji razdelijo svetlobni spekter na ozke pasove, praviloma širine manjše od 10 nm. Čeprav lahko uporabljamo široko- in ozkopasovne indekse, so za obdelavo velike količine spektralnih podatkov primernejše metode strojnega učenja, kot so npr. nevronske mreže. Pri določanju rastlinskih bolezni so bile uspešno uporabljene tako nadzorovane kot nenadzorovane metode, vendar se nobena ni izkazala kot absolutno boljša od ostalih, kar potrjuje potrebo po nadaljnjih raziskavah.

Hiperspektralno slikanje se vse pogosteje uporablja za odkrivanje stanja rastlin, saj omogoča zaznavanje biotskega in abiotskega stresa, kemometrične analize, določanje relevantnih valovnih dolžin in zgodnje odkrivanje bolezni, še pred pojavom vidnih znakov. Predhodne raziskave daljinskega zaznavanja bolezni vinske trte, kot sta zlata trsna rumenica in eska, so pokazale, da je bilo multispektralno slikanje z brezpilotnimi letalniki le delno uspešno. Podobne ugotovitve so bile dosežene tudi z uporabo multispektralnih satelitskih posnetkov, npr. v okviru projekta CRP V4-1103, kjer se je izkazalo, da slaba prostorska in spektralna ločljivost vodita do mešanih spektralnih podpisov ter s tem slabše napovedne moči klasifikacijskih modelov. V primerjavi s tem letalsko hiperspektralno slikanje teh pomanjkljivosti v veliki meri nima, dodatno pa jih je mogoče omiliti z zajemom podatkov v časovni vrsti, kar omogoča spremljanje fenološkega razvoja rastlin in bolezni ter zgodnje zaznavanje okužb v celotnih vinogradih.

Čeprav je bilo hiperspektralno slikanje že uspešno uporabljeno za določanje okužb vinske trte z različnimi patogeni, so bile dosedanje raziskave večinoma omejene na poskuse v rastlinjakih ali na manjših površinah in pogosto niso vključevale podatkov v časovni seriji. Uporaba posebnih terenskih platform za fenotipizacijo sicer omogoča standardizirane pogoje slikanja, vendar je omejena zaradi počasnega zajema in majhne uporabnosti na večjih vinogradniških površinah. Zato obstoječe metode, čeprav

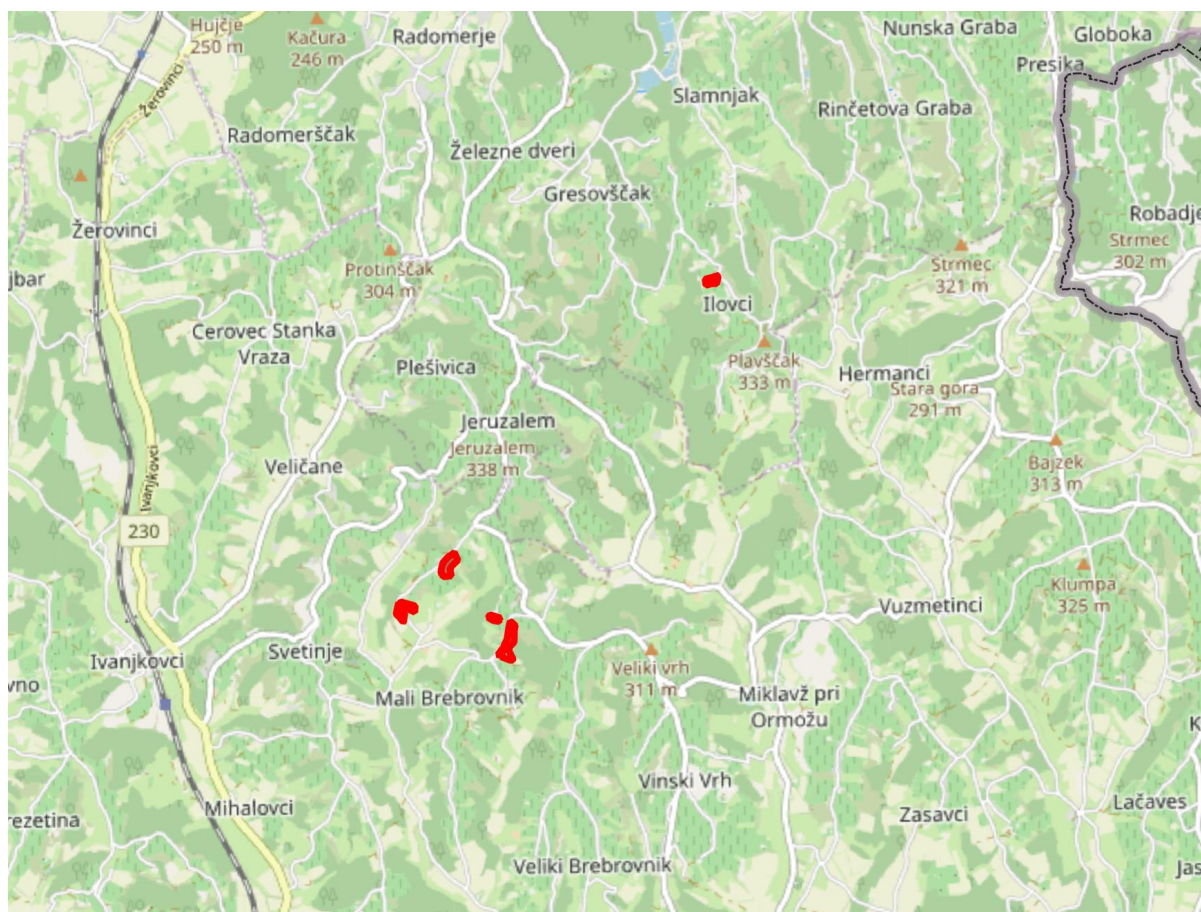
razmeroma uspešne, še ne omogočajo zanesljivega zgodnjega odkrivanja bolezni v dejanskih vinogradniških razmerah.

Namen dela v okviru DS7 je razvoj klasifikacijskih modelov za prepoznavanje zlate trsne rumenice, ki jo povzroča fitoplazma *Flavescence dorée* (FD), z uporabo letalskega hiperspektralnega slikanja v časovni seriji. S tem smo preverili možnost zgodnje zaznave okužb in omogočili pravočasno ter usmerjeno ukrepanje na večjih vinogradniških površinah, kar pomembno prispeva k učinkovitejšemu obvladovanju bolezni in zmanjšanju tveganja njenega širjenja.

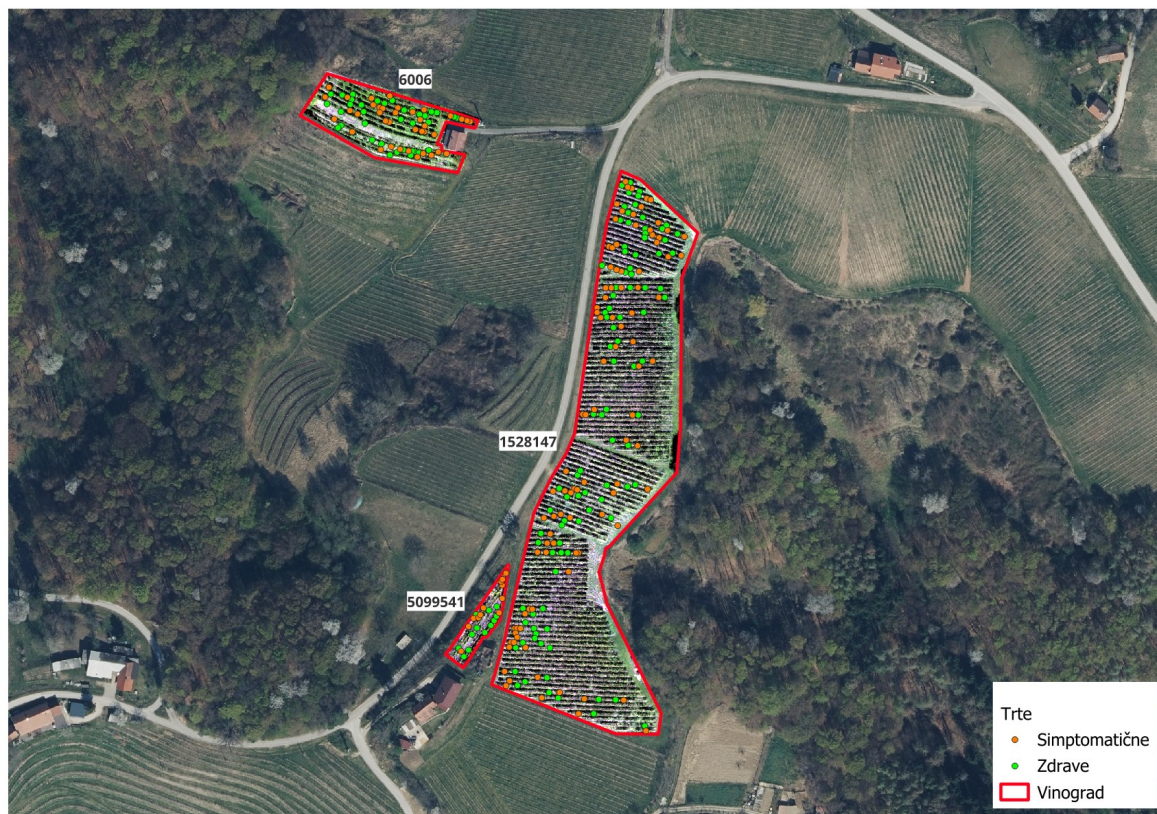
Naloga 7.1 Hiperspektralno slikanje z letalom in brezpilotnim letalnikom

Izbor vinogradov

Izbor vinogradov je bil opravljen na območju Jeruzalemsko-Ormoških goric, kjer je bil v zadnjih letih zaznan porast okužb z zlato trsno rumenico (Slika 1). Na podlagi strokovnih priporočil in pregleda terena smo izbrali sedem vinogradov z razmeroma visokim deležem simptomatičnih trt. V posameznem vinogradu smo vsako jesen izvedli vizualno oceno stanja in z visokonatančnim GNSS sistemom (Stonex S9i, Stonex, Italija) evidentirali koordinate trt z značilnimi simptomi okužbe (zvijanje in rumenenje oz. rdečenje listov, krhkost listne ploskve ipd.), hkrati pa določili tudi koordinate primerljivega števila vizualno zdravih trt (Slika 2).



Slika 1: Izbrani vinogradi.



Slika 2: Primer označenih simptomatičnih in zdravih rastlin v treh izbranih vinogradih v letu 2024.

Letalsko slikanje

Z namenom zgodnjega zaznavanja okužb s *Flavescence dorée* (FD) smo v dveh rastnih sezonah izvedli snemanje z letalom, opremljenim s hiperspektralnim snemalnim sistemom (Hypex VNIR-1600 in SWIR-384, Norsk Elektro Optikk AS, Norveška). Tak pristop omogoča zajem večjih površin v krajšem času v primerjavi z brezpilotnimi letalniki ali zgolj vizualnim ocenjevanjem.

Prelete smo izvedli dvakrat v rastni sezoni:

- prvič v juniju oziroma juliju, ko vizualni simptomi okužbe še niso izraženi,
- drugič v septembru, ko so simptomi že dobro vidni.

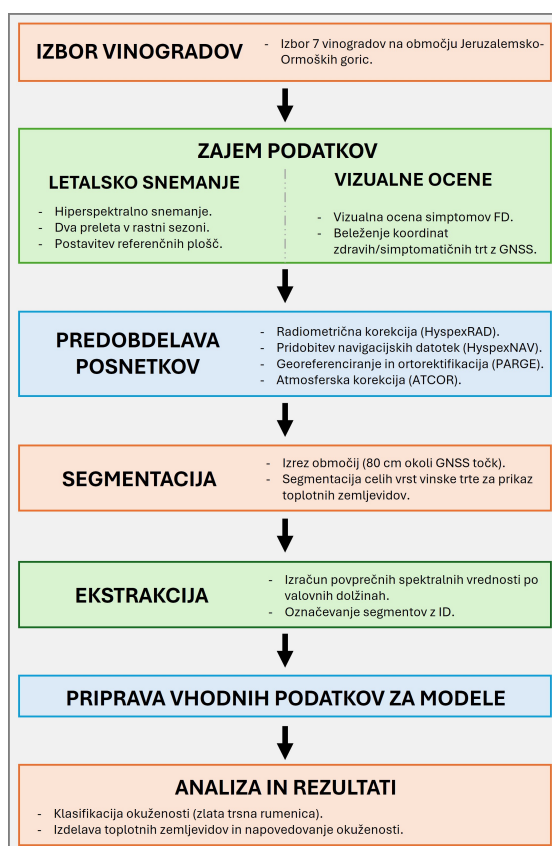
Ob snemanju smo na tla postavili referenčne plošče za pretvorbo signala v svetlobni odboj. Drugi prelet v sezoni je bil časovno usklajen z vizualnim ocenjevanjem na terenu, kar je zmanjšalo variabilnost podatkov in izboljšalo razmerje med signalom in šumom.

Obdelava hiperspektralnih slik

Posnetke smo najprej radiometrično korigirali s programom HypspxRAD in pridobili navigacijske datoteke (HypspxNAV). Georeferenciranje z znanimi točkami v bližini vinogradov ter ortorektifikacijo smo izvedli v programu PARGE (ReSe Applications LLC, Švica). Za izračun natančnejšega svetlobnega odboja na dnu atmosfere smo posnetke atmosfersko korigirali v programu ATCOR (ReSe Applications LLC, Švica).

Segmentacijo smo izvedli tako, da smo okoli vsake GNSS-točke, ki je predstavljala simptomatično ali zdravo rastlino, ustvarili območja (bufferje) s polmerom 40 cm. Tako smo za kasnejšo uporabo zajeli le območje trt, izločili pa dele, ki predstavljajo medvrstni prostor (npr. travo, zemljo). Za vsako območje smo nato izračunali povprečne vrednosti po valovnih dolžinah in jih opremili z identifikacijsko številko. Obdelavo podatkov trt in segmentacijo posnetkov smo opravili s programoma R in QGIS. Pripravljeni tabelarni podatki so bili uporabljeni v klasifikacijskih modelih (Slika 3).

Za prikaz toplotnih zemljevidov smo pri istih vinogradih dodatno označili celotne vrste vinske trte in jih segmentirali, s čimer smo ustvarili podlago za nadaljnje napovedovanje okuženosti z zlato trsno rumenico.



Slika 3: Diagram poteka obdelave hiperspektralnih slik.

Naloga 7.2 Strojno učenje in validacija modelov

Metode dela

Uporabili smo standardiziran cevovod za obdelavo hiperspektralnih podatkov in učenje modelov, pri čemer smo posebno pozornost namenili dosledni preprečitvi podatkovnega pronicanja med razvojnimi koraki.

Surove spektre (svetlobni odboj merjen kot odstotek odbite vpadne svetlobe za vsak spektralni pas) smo najprej korigirali z metodo multiplicative scatter correction (MSC), da smo zmanjšali vpliv sipanja svetlobe v atmosferi. Na tako korigiranih signalih smo izvedli detrending, s čimer smo odpravili počasno spreminjajočo se osnovnico in mehanske drifte, nato pa smo izračunali Savitzky–Golayeve (SG) derivate 2. reda, ki poudarijo informativne pasove in stisnejo nizkofrekvenčni šum. Vsi parametri predobdelave so bili določeni izključno na podatkih za učenje v okviru validacijskih zank, transformacije pa so bile v vsakem razcepu prilegane le na učni del in šele nato uporabljene na pripadajočem validacijskem delu.

Po predobdelavi (MSC, detrending in SG-derivati) smo spektre projicirali v frekvenčno domeno z diskretno Fourierjevo transformacijo, da bi zajeli periodične vzorce, ki v časovni/dolžinskovalovni domeni niso jasno razločljivi. Iz nastalih koeficientov smo glede na naraščajočo frekvenco obdržali prvih 50 komponent, saj te povzemajo večino informativne energije signala ob razmeroma majhnem prispevku šuma. Tako ustvarjen frekvenčni opis smo nato podali metodi delnih najmanjših kvadratov (PLS) v obliki diskriminantne analize za nadaljnje zmanjšanje dimenzionalnosti. PLS je izluščil latentne spremenljivke (LV), ki maksimirajo kovarianco med značilnicami in ciljno spremenljivko, s čimer smo zmanjšali kolinearnost, omejili prenaučeno (angl. overfitting) ter izboljšali stabilnost in razločljivost modela. Optimalno število LV smo določili z 10-krat ponovljeno navzkrižno validacijo in vrednotenjem natančnosti diskriminantne analize s primerjavo maksimalne razdalje, razdalje med centri in Mahalanobisovo razdaljo. Ta postopek smo izvedli za vsak model.

Za vizualizacijo morebitnih nelinearnih struktur v spektralnih podatkih smo uporabili t-SNE (t-distributed Stochastic Neighbor Embedding). Namen analize je bil preveriti, ali se zdravi in bolni primerki naravno razporedijo v ločene grozde v nizkorazsežnem prostoru, kar bi nakazovalo na izrazite lokalne razlike v spektrih. Za robustnost smo preizkusili več nastavitvenih kombinacij hiperparametrov ter dovolj veliko število iteracij (> 1000), da se je stroškovna funkcija stabilizirala. t-SNE je primeren za naše podatke, ker ohranja lokalno sosesko (bližine) in zato dobro razkrije drobne nelinearne strukture ter deluje neodvisno od razrednih oznak, zato omogoča nepristranski vpogled v "naravno" geometrijo podatkov.

Podatkovni niz smo naključno razdelili na učni in evalvacijski del v razmerju 70:30, pri čemer smo uporabili stratificirano delitev. Evalvacijski del je ostal v celoti neodvisen in nedotaknjen do končnega preverjanja uspešnosti. Na učnem delu smo za ocenjevanje posplošitvene sposobnosti in za izbiro nastavitev izvedli 5-krat ponovljeno 5-kratno navzkrižno validacijo, kar je zagotovilo stabilne ocene variance in zmanjšalo odvisnost od posamezne delitve. V vsakem ponovitvenem sklopu smo vse korake predobdelave in prileganja modela znova izvedli iz nič, skladno s postopkovnimi dobrimi praksami.

Za modeliranje smo uporabili podporne vektorje (SVM), kjer smo hiperparametre prilagodljivo optimizirali po vnaprej definiranem iskalnem prostoru z 81 kombinacijami. Optimizacija je potekala znotraj validacijskih zank na učnem delu, pri čemer smo za vsako kombinacijo hiperparametrov izračunali povprečno uspešnost čez vse gube in ponovitve ter izbrali nastavitve, ki je dosegla najboljši kompromis med prileganjem in posplošitvijo. Ko so bili hiperparametri določeni, smo končni SVM model ponovno prilegli na celoten učni del, z enakim zaporedjem predobdelave (MSC, detrending, SG derivati) naučenim zgolj na tem delu ter nato enkratno ocenili delovanje na neodvisnem 30-odstotnem evalvacijskem naboru. Metrike uspešnosti smo izračunali na učnih validacijah za primerjavo modelov in na evalvacijskem delu kot neodvisno oceno končnega modela, s čimer smo zagotovili realističen vpogled v zanesljivost napovedi. Celoten potek omogoča ponovljivost in metodološko pravilnost, saj strogo loči faze nastavljanja, izbire modela in končne ocene.

Podatkovni nabor sestavljajo hiperspektralne meritve iz sedmih vinogradov, pri čemer je bil zajem izveden dvakrat v rastni sezoni skozi dve zaporedni leti, kar za vsak vinograd predstavlja štiri kampanje in skupno osemindvajset zajemov. Iz tega nabora smo sestavili podatkovne nabore, ločene po letih, sezoni zajema podatkov, po vinogradih ter v kombinacijah (npr. obe leti v eni sezoni za vsak vinograd). Takšna zasnova nudi celovit pregled nad poskusom, hkrati pa omogoča oceno prenosljivosti med leti, sezonske občutljivosti ter lokalne specifičnosti posameznih vinogradov.

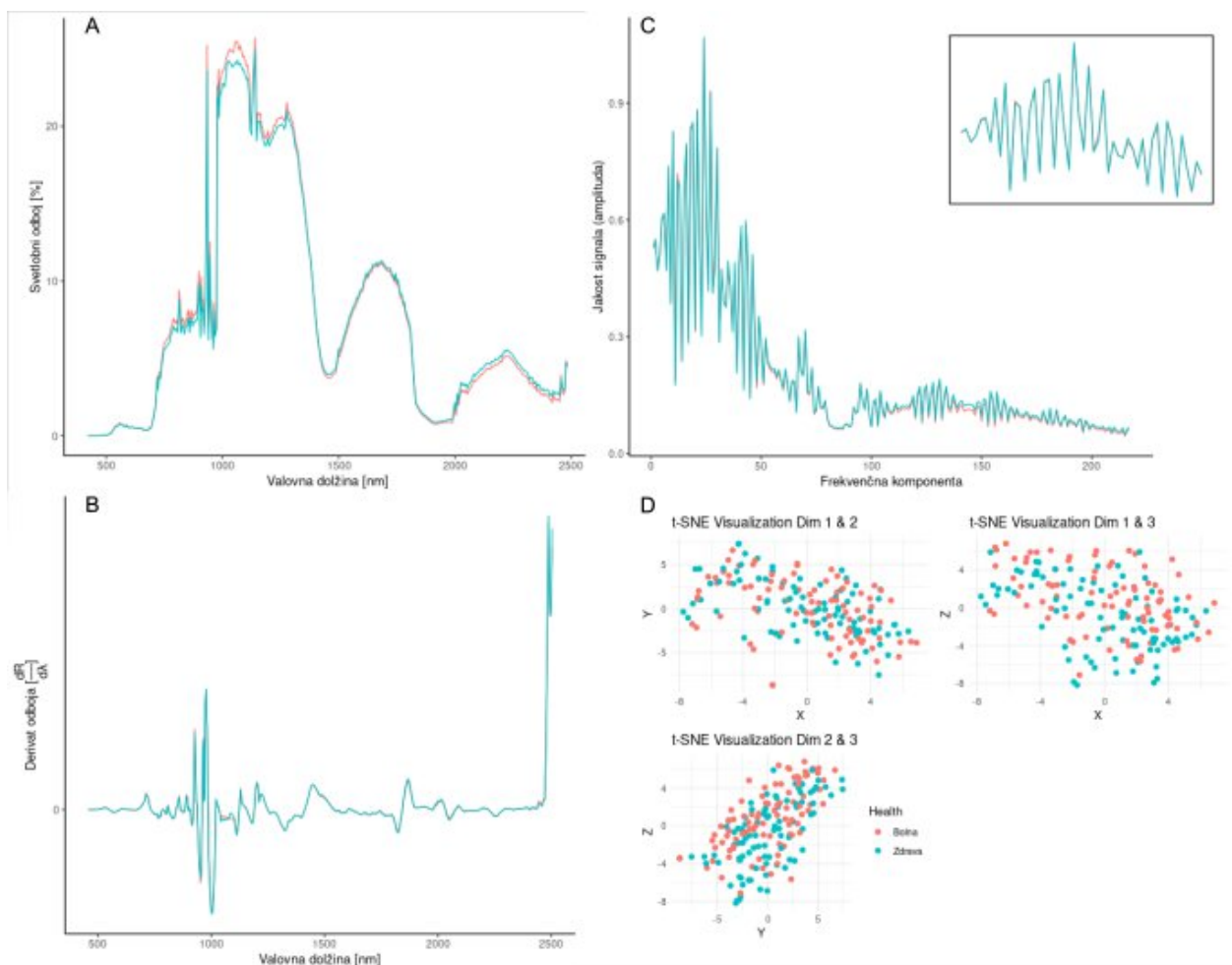
Rezultati in razprava

Predpriprava podatkov

Povprečni spektralni podpisi zdravih in bolnih rastlin so izkazali le manjše razlike, ki so bile izrazite predvsem v bližnjem infrardečem (NIR) in kratkovalovnem infrardečem (SWIR) delu spektra (Slika 4A). V vidnem območju (VIS) med skupinama nismo zaznali večjih odstopanj, kar kaže, da so spremembe, povezane z bolezenskimi procesi, bolj prisotne v strukturalno in vodno občutljivih valovnih območjih ter manj v pigmentno občutljivem VIS delu. Uporaba 2. derivatov s Savitzky–Golay glajenjem je dodatno poudarila razlike, zlasti v območju 700–1000 nm (t. i. “red-edge” in zgodnji NIR), kjer so se izkazale manjše spremembe naklonov in infleksijskih točk, ki v izvirnih (0. reda) spektrih niso bile jasno razvidne (Slika 4B).

Za nadaljnjo analizo smo uporabili hitro Fourierjevo transformacijo (FFT), ker pretvorba spektralnih krivulj v frekvenčno domeno omogoča kompaktno predstavitev glavnih trendov in periodičnih komponent spektra, obenem pa deloma loči koristne informacije od visokofrekvenčnega šuma. FFT olajša tudi zmanjšanje dimenzionalnosti in zmanjša medsebojno koreliranost sosednjih valovnih pasov, kar je ugodno za razvoj modelov strojnega učenja. Izbrali smo prvih 50 frekvenčnih komponent, ker nizkofrekvenčni koeficienti nosijo večino energije signala (opisujejo gladke, fiziološko relevantne trende), medtem ko višjefrekvenčni koeficienti večinoma predstavljajo šum in drobne lokalne variacije brez jasne biološke interpretacije. Takšen prirez ohrani ključne informacije, hkrati pa stabilizira modele in zmanjša nevarnost prenaučenosti (Slika 4C).

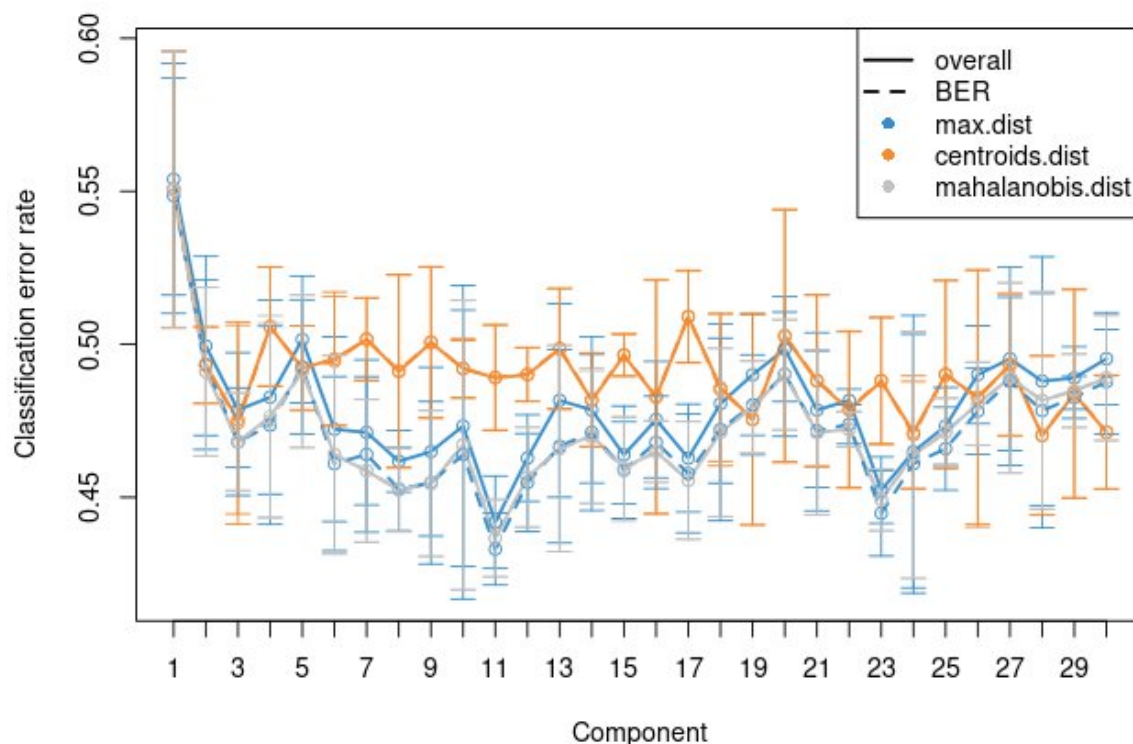
Podatki zdravih in bolnih rastlin se v t-SNE prostoru niso ločili v jasne, stabilne skupine, ne glede na preizkušene nastavitve in kombinacije t-SNE komponent. Tudi morebitni lokalni agregati niso bili stabilni po različnih naključnih začetkih, kar dodatno kaže na odsotnost robustne nelinearne strukture, ki bi ločevala skupine. Razlike med skupinama so majhne in verjetno porazdeljene po več valovnih območjih, zato zahtevajo nadzorovano ekstrakcijo značilk (npr. PLS) in nadaljnje klasifikacijske modele, medtem ko nenadzorovana t-SNE projekcija sama po sebi ne razkrije izrazite ločnice. Tak vzorec je skladen s prejšnjimi rezultati: razlike med skupinama so majhne in večinoma zajete v specifičnih valovnih območjih (NIR/SWIR), ne pa v celotni globalni strukturi spektra, ki jo t-SNE poskuša lokalno ohraniti. Poleg tega t-SNE kot nenadzorovana metoda ne uporablja informacij o razredu, zato lahko ostanejo razločevalni, a šibki signali skriti v znotraj-skupinski variabilnosti.



Slika 4: Združeni prikazi: (A) povprečni spektralni podpisi zdravih in bolnih rastlin – manjše razlike predvsem v NIR/SWIR, v VIS brez večjih odstopanj; (B) 2. derivati s Savitzky–Golay glajenjem, ki poudarijo razlike v območju 700–1000 nm; (C) amplitudni spekter po FFT z izrezom prvih 50 frekvenčnih komponent (okence) uporabljenih za nadaljnje analize; (D) t-SNE projekcija brez vidnih, stabilnih grupacij med skupinama.

Za redukcijo dimenzionalnosti in izračun značilk smo uporabili metodo delnih najmanjših kvadratov (Slika 5). Število latentnih spremenljivk (LV) v PLS neposredno vpliva na natančnost kasnejšega modela strojnega učenja, ker določa razmerje med zajeto informacijo in uvedenim šumom. Z majhnim številom LV-jev model pogosto podovrednoti (underfitting), saj del spektralne variance, ki je pomembna za razločevanje zdravih in bolnih rastlin, ostane nepojasnen, zato so meje razreda plitve in napovedna točnost nižja. Z naraščanjem števila LV-jev se praviloma poveča natančnost in stabilizira meje med razredi. Po “komolčni” točki (najnižja napaka diskriminatne analize PLS) pa dodatni LV večinoma predstavljajo redundanco ali šum. Algoritem strojnega učenja se začne prilagajati naključnim nihajem (prenapriranje), poslabša se generalizacija in narašča variabilnost rezultatov med ponovitvami validacije. Gre za klasični kompromis, t. i. “prekletstvo dimenzionalnosti” pri omejenem številu vzorcev. Tak postopek daje

SVM-ju kompakten, informativen vhod, ki maksimira natančnost in robustnost na novih (testni niz) podatkih.



Slika 5: Potek napake diskriminantne analize PLS skozi število LV-jev. Na osi X je število komponent, na osi Y pa povprečna stopnja napake s standardnim odklonom. Slednjo smo merili z tremi metrikami razdalj (največja razdalja, razdalja med centroidi, in Mahalanobisova razdalja). V prikazanem primeru smo izbrali 11 komponent. Pri 10 LV je povprečna natančnost slabša, standardni odklon pa večji, kar nakazuje slabšo stabilnost modela. Pri 23 LV je podoben rezultat, kot pri 10, vendar večje število komponent poveča verjetnost prevelikega prilaganja modela podatkom (overfit).

Razvoj napovednih modelov

Rezultati kažejo dosledno boljšo ločljivost simptomov jeseni kot spomladi in zmeren do izrazit generalizacijski padec med učno (združeno obe leti in vsi vinogradi $F1 = 0,721$) in testno množico (združeno obe leti in vsi vinogradi $F1 = 0,600$). Padec je največji pri vinogradih z majhnim številom potrjenih primerov, kar nakazuje prenaučenosť in visoko varianco. Vzorec po sezonah in znotraj posameznih vinogradov je podoben, jesenske vrednosti $F1$ so praviloma višje, relativni rang terminov se ohranja, absolutne vrednosti pa na testu padejo. To podpira operativno tezo, da je jesenski prelet najprimernejši za referenčno kartiranje, spomladanski pa za zgodnje opozarjanje, ob nižji natančnosti.

Jesenske ocene presegajo spomladanske tako na učni kot na testni množici. To je pričakovano, saj so jeseni simptomi spektralno bolj kontrastni. Vinograd 3 kaže izrazit jesenski dvig na obeh sklopih (učni do 0,95; testni do 0,774), podobno 4 (učni

~0,80–0,816 jeseni; testni ~0,727 jeseni). Agregat “Skupaj pomlad” je pri večini vinogradov nižji na testu (npr. 1528147: 0,715 → 0,639), kar potrjuje omejitve zgodnjega zaznavanja brez posebnih prilagoditev. Lokaciji z majhnim številom podatkov (vinograda 2 in 4) na učni množici dosegata popolno natančnost ($F1 = 1,0$) v posameznih sezonah, kar se na testu normalizira na 0,667–0,857. Gre za tipičen učinek majhnih vzorcev in potencialno optimistično oceno zaradi prekrivnosti ali naključne razdelitve. Nasprotno pa vinogradi z večjim številom podatkov kažejo bolj stabilen sezonski vzorec in manj ekstremne vrednosti, vendar še vedno prisoten padec na testu (Tabela 1).

Tabela 1: F1-natančnosti klasifikacijskih modelov pri ločevanju zdravih in bolnih rastlin. Podani so rezultati na nizu za učenje in nizu za testiranje modelov. N: število podatkov; N train/N test: število podatkov v učenem in testnem nizu.

vinograd		N	Pomlad 2023	Jesen 2023	Pomlad 2024	Jesen 2024	Skupaj pomlad
Niz za učenje	1	382	0.755	0.733	0.727	0.737	0.715
	2	48	1	1	1	1	0.971
	3	236	0.82	0.837	0.847	0.95	0.739
	4	284	0.8	0.816	/	0.792	0.792
	5	148	0.88	0.84	0.857	0.906	0.99
	6	222	0.847	0.86	/	0.923	/
	7	48	/	/	1	1	/
Skupaj							0.721
N train / N test							
Niz za testiranje	1	267/115	0.69	0.727	0.76	0.727	0.639
	2	34/14	0.667	0.8	0.857	0.75	0.714
	3	165/71	0.645	0.65	0.743	0.774	0.646
	4	199/85	0.696	0.727	/	0.727	0.612
	5	104/44	0.759	0.727	0.727	0.75	0.723
	6	155/67	0.686	0.82	/	0.762	/
	7	34/14	/	/	0.8	0.857	/
Skupaj							0.6

Rezultati so v skladu z biološko in fenološko podlago zlate trsne rumenice. Jeseni so simptomi izrazitejši, zato je zaznavanje bolj zanesljivo. Zgodnje zaznavanje spomladi ostaja možno, vendar z nižjo natančnostjo, še posebej brez fenološke normalizacije in časovnih serij. Razkorak med učnimi in testnimi rezultati potrjuje, da razlike med lokacijami in sortami presegajo medletno variabilnost znotraj iste lokacije. Modeli, razviti na združenih podatkih, ki niso ozaveščeni o domenski strukturi (sorta/lokacija), se naučijo tudi neželene podpise okolja in sorte, kar znižuje prenosljivost. Spomladanski preleti so smiselni kot sistem zgodnjega zaznavanja, vendar je razmeroma zanesljivo določanje zlate trsne rumenice odvisno od sorte in lokacije, torej morajo biti modeli strojnega učenja temu prilagojeni. Jesenski preleti bi bili primerni za kartiranje okužb in usmerjanje fitosanitarnih ukrepov proti koncu sezone.

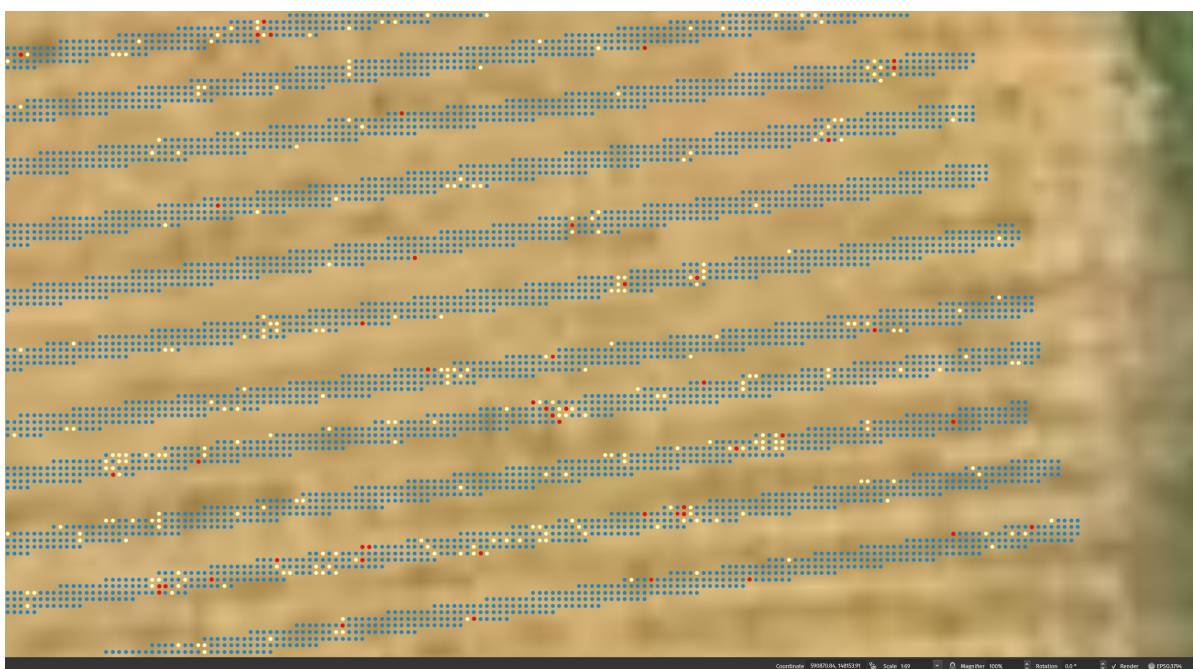
Vročinske karte

Vročinske karte prikazujejo prostorsko porazdelitev verjetnosti, da posamezna rastrska celica (angl. pixel) vsebuje okuženo rastlino. Barvna lestvica prehaja od nizkih verjetnosti (modro/zelene) prek mejnih (rumeno/oranžne, ~40–60 %) do visokih (rdeče, >61 %). Opazni so jasni skupki rdečih celic, ki označujejo žarišča povišanega tveganja, pogosto obdana z rastrskimi celicami z mejnimi verjetnostmi (40–60 %). Ta vzorec je skladen s prostorsko avtokorelacijo, okužbe se pojavljajo v lokalnih grozdih, medtem ko prehodni pasovi (rumena) nakazujejo robna območja žarišč ali mešanje zdravih in prizadetih rastlin.

Posamezne oziroma izolirane rastrske celice z višjo verjetnostjo se pogosto pojavijo ob robovih segmentiranih podatkov zaradi kombinacije robnih učinkov in mešanja signalov. Na mejah segmentov so piksli delno “mešani” med tarčno površino vinske trte in ozadjem (tla, listje sosednjih rastlin, žice opore), zato se njihov spektralni podpis razlikuje od homogenega notranjega dela segmenta. K temu prispevajo spremenjene razmere osvetlitve in sence na robovih krošenj, ki znižujejo razmerje signal/šum. Rezultat so “špičke” verjetnosti na posameznih pikslih ob mejah, kjer model zazna lokalne anomalije, ki niso nujno posledica dejanske okužbe, temveč posledica spektralnega mešanja in negotovosti na segmentnih robovih. Zato takšne izolirane celice interpretiramo previdno in jih vrednotimo v kontekstu sosednjih pikslov ter prostorske koherence celotnega segmenta.



Slika 6: Primer vročinske karte verjetnosti okužb. Višje verjetnosti okužb so pogostejše v severnem delu vinograda. Modra: verjetnost okužbe < 40 %. Rumena: 41–60 %. Rdeča: > 61 %.



Slika 7: Izrez in povečava vročinske karte. Jasno so vidne grupacije rastrskih celic, kjer je verjetnost okužbe višja. Modra: verjetnost okužbe < 40 %. Rumena: 41–60 %. Rdeča: > 61 %.

Zaključki

Naši rezultati utemeljujejo metodološki nabor za zgodnje odkrivanje na osnovi hiperspektralnih podatkov, pri čemer so ključni skrbna predobdelava, PLS-podprta redukcija dimenzionalnosti in sezonsko ter domensko ozaveščeno učenje, da se doseže čim bolj zanesljivo zgodnje prepoznavanje bolezni.

Odkrivanje rastlin okuženih z zlato trsno rumenico je zanesljivejše jeseni, ko so spektralni kontrasti simptomov izrazitejši, medtem ko je spomladi zgodnje opozarjanje sicer možno, vendar z nižjo natančnostjo in je občutljivejše na domenske učinke. Razkorak med učnimi in testnimi rezultati potrjuje, da razlike med lokacijami in sortami presegajo medletno variabilnost znotraj iste lokacije. Modeli, ki niso ozaveščeni o sorti in lokaciji, prevzamejo tudi neželene okoljske podpise, kar znižuje prenosljivost. V praksi bi jesenski preleti lahko služili za referenčno kartiranje in usmerjanje ukrepov, spomladanski pa za zgodnje opozarjanje z zavedanjem večje negotovosti.

Spektralne razlike med zdravimi in bolnimi rastlinami so majhne in skoncentrirane predvsem v NIR/SWIR, v vidnem delu so neizrazite. Derivati 2. reda poudarijo subtilne spremembe v območju 700–1000 nm, kar potrjuje informacijsko težo red-edge in NIR do 1000 nm. Pretvorba s FFT in izbor prvih 50 frekvenčnih komponent učinkovito povzameta glavne trende ter zmanjšata šum in redundanco, kar olajša nadaljnje modeliranje. t-SNE ne pokaže naravne ločitve med skupinama, zato je za izluščenje šibkih, a sistematičnih razlik potrebna nadzorovana ekstrakcija značilnk, kot je PLS, pri kateri moramo število latentnih spremenljivk uravnotežiti med zajeto informacijo in tveganjem prenaučenosti, da ostane model stabilen in prenosen.

Vročinske karte potrjujejo prostorsko avtokorelacijo z žarišči višjih verjetnosti, ki so pogosto obdana s pasovi mejnih vrednosti. Posamezni piksli z višjo verjetnostjo se pojavljajo predvsem ob robovih segmentov zaradi spektralnega mešanja in robnih učinkov, zato jih je treba interpretirati previdno in v kontekstu sosednjih pikslov ter koherence segmenta.