

Zbornik strokovne konference

ROSUS 2022

Računalniška obdelava slik in
njena uporaba v Sloveniji 2022

Urednik:
Božidar Potočnik



Maribor, 17. marec 2022



ROSUS 2022 - Računalniška obdelava slik in njena uporaba v Sloveniji 2022

Zbornik 16. strokovne konference

Urednik
Božidar Potočnik

Marec 2022

Naslov <i>Title</i>	ROSUS 2022 - Računalniška obdelava slik in njena uporaba v Sloveniji 2022 <i>ROSUS 2022 - Computer image processing and its application in Slovenia 2022</i>
Podnaslov <i>Subtitle</i>	Zbornik 16. strokovne konference <i>Proceedings of the 16th Professional Conference</i>
Urednik <i>Editor</i>	Božidar Potočnik (Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko)
Recenzija <i>Review</i>	Aleš Holobar (Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko) Simon Dobrišek (Univerza v Ljubljani, Fakulteta za elektrotehniko)
Tehnični urednik <i>Technical editor</i>	Jan Perša (Univerza v Mariboru, Univerzitetna založba)
Oblikovanje ovitka <i>Cover designer</i>	Gašper Sedej (Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko)
Grafične priloge <i>Graphic material</i>	Avtorji prispevkov, 2022
Konferenca <i>Conference</i>	Strokovna konferenca ROSUS 2022: Računalniška obdelava slik in njena uporaba v Sloveniji 2022
Datum in kraj konference <i>Conference date & location</i>	17. marec 2022, Maribor, Slovenija
Programski odbor <i>Program committee</i>	Aleš Holobar, Damjan Zazula, Boris Cigale, Marjan Mernik, Peter Peer, Tomaž Tomažič, Matjaž Colnarič, Slobodan Ribarič, Vojko Flis, Saša Divjak, Stanislav Kovačič, Zdravko Kačič, Aleš Leonardis, Simon Dobrišek, Vitomir Štruc, Franc Solina, Andrej Šostarič, Franjo Pernuš, Jure Skvarč in Božidar Potočnik
Organizacijski odbor <i>Organization committee</i>	Gašper Sedej, Jurij Munda, Uroš Mlakar, Matjaž Divjak, Martin Šavc, Danilo Korže, Aleš Holobar in Božidar Potočnik
Založnik <i>Published by</i>	Univerza v Mariboru Univerzitetna založba Slomškov trg 15, 2000 Maribor, Slovenija https://press.um.si , zalozba@um.si
Izdajatelj <i>Issued by</i>	Univerza v Mariboru Fakulteta za elektrotehniko, računalništvo in informatiko Koroška cesta 46, 2000 Maribor, Slovenija https://feri.um.si , feri@um.si

Izdaja
Edition Prva izdaja

Vrsta publikacije
Publication type E-knjiga

Dostopno na
Available at <https://press.um.si/index.php/ump/catalog/book/597>

Izdano
Published Maribor, marec 2022



© Univerza v Mariboru, Univerzitetna založba
University of Maribor, University Press

Besedilo/ *Text*
© Avtorji in Potočnik, 2022

To delo je objavljeno pod licenco Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 4.0 Mednarodna. / *This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License.*

Uporabnikom se dovoli reproduciranje, distribuiranje, dajanje v najem, javno priobčitev in predelavo avtorskega dela, če navedejo avtorja in širijo avtorsko delo/predelavo naprej pod istimi pogoji. Za nova dela, ki bodo nastala s predelavo, je tudi dovoljena komercialna uporaba.

Vsa gradiva tretjih oseb v tej knjigi so objavljena pod licenco Creative Commons, razen če to ni navedeno drugače. Če želite ponovno uporabiti gradivo tretjih oseb, ki ni zajeto v licenci Creative Commons, boste morali pridobiti dovoljenje neposredno od imetnika avtorskih pravic.

<https://creativecommons.org/licenses/by-sa/4.0/>



SLOVENSKI PODJETNIŠKI SKLAD

Pokrovitelji

CIP - Kataložni zapis o publikaciji
Univerzitetna knjižnica Maribor

004.932 (497.4) (082) (0.034.2)

STROKOVNA konferenca ROSUS (16 ; 2022 ; Maribor)

ROSUS 2022 [Elektronski vir] : računalniška obdelava slik in njena uporaba v Sloveniji 2022 : zbornik 16. strokovne konference / urednik Božidar Potočnik. - 1. izd. - Maribor : Univerzitetna založba Univerze, 2022

Dostopno tudi na: <https://press.um.si/index.php/ump/catalog/book/597>
ISBN 978-961-286-575-7

doi: 10.18690/um.feri.7.2022

1. Potočnik, Božidar

COBISS.SI-ID 99918339

ISBN 978-961-286-575-7 (pdf)
978-961-286-576-4 (mehka vezava)

DOI <https://doi.org/10.18690/um.feri.7.2022>

Cena Brezplačni izvod
Price

Odgovorna oseba založnika prof. dr. Zdravko Kačič,
For publisher rektor Univerze v Mariboru

Citiranje Potočnik, B. (ur.). (2022). *ROSUS 2022 - Računalniška obdelava slik in njena uporaba v Sloveniji 2022: zbornik 16. strokovne konference*
Attribution Maribor: Univerzitetna založba. doi: 10.18690/um.feri.7.2022

Kazalo

Predgovor	
<i>Foreword</i>	1
Božidar Potočnik	
VABLJENA PREDAVANJA	3
Laboratorij za slikovno biometrijo	
<i>Imaging Biometrics Laboratory</i>	5
Peter Peer	
Priložnosti in prednosti digitalno podprtega kmetijstva	
<i>Possibilities and Advantages of Digitally Supported Agriculture</i>	7
Jurij Rakun	
POPOLDANSKA SEKCIJA	
Globoko učenje: Praktični nasveti strokovnjakov	19
Umetna inteligenca ali umetna inteligentnost: kje smo in kam gremo	
<i>Science Mapping of the Knowledge Base on Artificial Intelligence</i>	21
Andrej Kastrin	
DOPOLDANSKA SEKCIJA	
Industrijske aplikacije, Medicinske in biomedicinske aplikacije, Drugo, Študentske aplikacije	33
Segmentacija telesa z uporabo večciljnega učenja	
<i>Body Segmentation Using Multi-task Learning</i>	36
Julijan Jug, Ajda Lampe, Peter Peer in Vitomir Štruc	
Ločevanje široko in ozko listnih rastlin v podporo sistemu za zaznavo plevel	
<i>Separation of Wide and Narrow-leaf Plants in Support of Weed Detection System</i>	55
Urban Kenda in Jurij Rakun	
Detekcija vlitih navojev s pomočjo termovizije	
<i>Molded Thread Detection Using Thermal Imaging</i>	73
Anže Švigelj in Borut Batagelj	
Ocenjevanje starosti osebe na osnovi digitalnih posnetkov z uporabo modificirane mreže VGG-Face	
<i>Person Age Estimation Based on Digital Images Using Modified VGG-Face Network</i>	81
Tilen Krel in Božidar Potočnik	

Spoštovani!

Po petnajstih konferencah ROSUS 2006–2020 in z enoletnim premorom zaradi pandemije želimo s konferenco ROSUS 2022 nadaljevati s promoviranjem pomembnosti ekonomske koristi računalniške obdelave slik na področjih industrije, biomedicine in drugih poslovnih procesov. Vezi, ki smo jih na prejšnjih konferencah stkali med raziskovalci, razvijalci, ponudniki rešitev ter uporabniki računalniške obdelave slik v slovenskem prostoru, želimo še dodatno okrepiti, ob tem pa nuditi tudi možnosti sklepanja novih sodelovanj in svetovanja pri razreševanju konkretnih poslovnih oziroma raziskovalnih problemov.

Glavni poudarek namenjamo aplikacijam s področja računalniške obdelave slik, ki so že integrirane oziroma pripravljene za integracijo v poslovne procese. Na tej konferenci nadaljujemo globalni trend na področju računalniškega vida s popoldansko sekcijo z naslovom »Globoko učenje: Praktični nasveti strokovnjakov«, ki jo organiziramo v sodelovanju s podjetjem Kolektor. Demonstrirali bomo, da avtomatska obdelava v industriji lahko zaradi svoje natančnosti in hitrosti prinaša velike ekonomske koristi, hkrati pa nakazali, da aplikacije računalniške obdelave slik nudijo nove priložnosti za uveljavitev na trgu visokih tehnologij. Seveda ne smemo pozabiti na možnost diskusije ter predstavitev konkretnih problemov in potreb, ki

se porajajo pri uporabnikih, s katerimi bomo računalniško obdelavo slik in njeno koristnost še bolj približali avditoriju.

Naj sklenemo uvodne misli še s prisrčno zahvalo Javnemu skladu Republike Slovenije za podjetništvo, ki v okviru konference ROSUS 2022 predstavlja zanimive finančne instrumente za spodbujanje prenosa tehnoloških rešitev v podjetniško sfero. Izpostaviti želimo še medijskega pokrovitelja revijo IRT3000, ki je intenzivno promoviral konferenco ROSUS 2022 ter pomen strojnega vida v slovenskem prostoru.

Božidar Potočnik
predsednik konference
ROSUS 2020



VABLJENA PREDAVANJA

SLIKOVNA BIOMETRIJA NA POHODU

PETER PEER

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Ljubljana, Slovenija.

E-pošta: peter.peer@fri.uni-lj.si

Povzetek V zadnjem desetletju se je v Laboratoriju za računalniški vid na FRI UL oblikovala močna skupina, ki dela na področju biometrije. Skupina je močno povezana z Laboratorijem za strojno inteligenco na FE UL. Prvi ključni koraki so bili narejeni v okviru kompetenčnih centrov, kjer smo v oblaku naredili fuzijo modalnosti, obrazov in prstih odtisov. Vzporedno s tem se je odvijalo delo na razpoznavanju ljudi iz načina gibanja. Nato je delo na področju biometrije dobilo še dodaten zagon. Posvetili smo se povsem novi modalnosti uhljev, začeli delati na izzivu fotorealistične deidentifikacije obrazov, dodali beločnico, šarenico ter obočesno regijo kot naslednje tri sveže modalnosti. V zadnjem času se ukvarjamo tudi z detekcijo globokih ponaredkov, oceno pravičnega nošenja mask, virtualnim pomerjanjem oblačil, obdelavo prstnih sledi (latentov) in celovitim ohranjanjem zasebnosti obrazov. Na drugi stopnji študija smo uvedli tudi nov izbirni predmet Slikovna biometrija. Ta ima letos spet skoraj 90 slušateljev. Število članov skupine trenutno raste iz leta v leto, temu primerno tudi financiranje in publikacije na ključnih konferencah ter v revijah, nenazadnje pa se vpliv skupine pozna tudi pri organizaciji tekmovanj na teh konferencah ter tudi zmagah na sorodnih tekmovanjih. Predavanje bo osvetlilo prehojeno pot skozi ključne raziskovalne vsebine. Spletna stran: <https://fri.uni-lj.si/sl/laboratorij/lrv>.

Ključne besede:

računalniški
vid,
biometrija,
nove
modalnosti,
deidentifikacija,
globoki
ponaredki

IMAGE-BASED BIOMETRICS LABORATORY

PETER PEER

University of Ljubljana, Faculty of Computer and Information Science, Ljubljana,
Slovenia.

E-mail: peter.peer@fri.uni-lj.si

Abstract In the last decade, a strong team working in the biometrics field has been formed in the Computer Vision Laboratory (FRI UL). The group is strongly connected with the Machine Intelligence Laboratory (FE UL). The first steps were taken within competence centers, where we made a fusion of face and fingerprints modalities in the cloud. In parallel, we were working on gait recognition. Then, the work in the biometrics field got intensified. We focused on a new modality of ears, began to work on a photorealistic deidentification of faces, added the sclera, iris, and ocular region as next new modalities. Recently, we have been dealing with the deep fake detection, the assessment of correct face-masks wearing, virtual garment try-on, processing of fingermarks, and privacy-preserving face analytics. At the second Bologna cycle, we introduced a new elective course, Image-based Biometrics. This year it has almost 90 listeners again. The number of team members is growing from year to year, as well as funding and publications at key conferences and journals. Finally, the influence of the team is seen through organization of competitions at these conferences and victories in related competitions. The lecture will highlight the path traversed through key research topics. Website: <https://fri.uni-lj.si/en/laboratory/lrv>.

Keywords:

computer
vision,
biometrics,
new
modalities,
deidentification,
deep
fakes

PRILOŽNOSTI IN PREDNOSTI DIGITALNO PODPRTEGA KMETIJSTVA

JURIJ RAKUN

Univerza v Mariboru, Fakulteta za kmetijstvo in biosistemske vede, Hoče, Slovenija.
E-pošta: jurij.rakun@um.si

Povzetek Kmetijstvo je ena izmed panog, ki ob podpori digitalizacije doživlja ponoven preporod. Množica novih tehnoloških rešitev, kot so sistemi bližnjega ali oddaljenega zaznavanja, mobilne aplikacije, IoT in druge tehnologije ponujajo nove možnosti, ki jih izkoriščajo principi preciznega kmetijstva. Precizno kmetijstvo kmetovalcem omogoča selektivno, usmerjeno in natančno obdelavo, kar pa omogoča precejšnje prihranke pri porabi vhodnih surovin, manjšo ekološko obremenitev okolja in večji donos. V sklopu vabljenega predavanja je predstavljenih nekaj izbranih tem, ki zajemajo uporabo avtonomnih poljedelskih robotov, brezpilotnih letalnikov in naprednih senzorskih sistemov, ki temeljijo na postopkih digitalne obdelave signalov. V sklopu izbranih tem bodo predstavljene prednosti in priložnosti, ki jih postopki digitalizacije nudijo za sodobna kmetijska gospodarstva.

Ključne besede:

digitalizacija,
kmetijstvo,
precizno
kmetijstvo,
kmetijski
roboti,
senzorika,
digitalna
obdelava
signalov

POSSIBILITIES AND ADVANTAGES OF DIGITALLY SUPPORTED AGRICULTURE

JURIJ RAKUN

University of Maribor, Faculty of Agriculture and Life Sciences, Hoče, Slovenia.
E-mail: jurij.rakun@um.si

Abstract Agriculture is one of the disciplines that is experiencing a rebirth in support of the digitalisation. A multitude of new technological solutions, such as near or remote sensing systems, mobile applications, IoT and other technologies, offer new possibilities that are exploited by the principles of precision agriculture. Precise agriculture enables farmers to cultivate selectively, focused and precisely, which in turn enables significant savings in the consumption of input raw materials, lower ecological burden on the environment and higher yields. As part of the invited lecture, some selected topics are presented, including the use of autonomous agricultural robots, drones and advanced sensor systems based on digital signal processing procedures. The selected topics will present the advantages and opportunities that digitization processes offer for modern agricultural estates.

Keywords:

digitalization,
agriculture,
precision
agriculture,
agricultural
robotics,
sensors,
digital signal
processing

1 Uvod

Kmetijstvo se nahaja pred množico ključnih vprašanj; kako zagotoviti hrano za naraščajoče število prebivalstva, kako to napraviti ob omejenih možnostih, kot je velikost potencialno uporabnih kmetijskih površin in nenazadnje kako pridelek pridelati na trajnosten način, s čimer ne uničujemo narave in jo lahko prenesemo še v naslednje rodove.

Odgovor na to ponujajo tehnologije preciznega kmetijstva. Te zajemajo različne postopke, različne rešitve in spremenjeno miselnost na običajne pristope h kmetovanju. Tehnologije preciznega kmetijstva temeljijo na prostorski in časovni variabilnosti, ki omogočajo, da običajna kmetijska opravila opravimo selektivno, v pravih količinah in na pravih mestih. Na ta način lahko zagotovimo optimalne pogoje za rast rastlin in po drugi strani prihranimo na vhodnih surovinah. Večja kakovost pridelka in poraba manj surovin pa vpliva tudi na končno ceno in večji donos.

To delo povzema nekaj izbranih primerov s katerimi se ukvarjamo na Katedri za biosistemsko inženirstvo iz Fakultete za kmetijstvo in biosistemske vede, Univerze v Mariboru. Delo povzema naše izkušnje iz množice različnih projektov in jih povezuje z različnimi digitalizacijskimi postopki, ki vključujejo tudi postopke digitalne obdelave signalov.

2 Stanje na področju preciznega kmetijstva v Sloveniji

Precizno kmetijstvo se po svetu razvija različno hitro (Rihter, 2020). Ob tem srečamo kmetovalce, ki so pripravljeni nove tehnologije prevzeti že zelo zgodaj, do kmetovalcev, ki sledijo večini in tudi takih, ki se držijo predvsem ustaljenih praks in nove tehnologije prevzamejo, ko je to nujno. Seveda na vse to vplivajo tudi različni dejavniki, kot so starost prebivalstva, velikost in vrsta kmetijskih gospodarstev, subvencije in drugo.

V sklopu izvedene anketne raziskave preciznih tehnologij na slovenskih kmetijah, smo iskali razloge za njihovo uporabo oziroma neuporabo. Prišli smo do zanimivih dognanj. Ugotovili smo namreč, da dobre tri tretjine anketiranih, vprašani so bili lastniki kmetij različnih starosti z različno velikimi kmetijami in dejavnostmi, na svojih kmetijah še ni uporabljala preciznih tehnologij, vendar pa jih njihova uporaba v bližnji prihodnosti zanima, kar je vzpodbudno za prihodnost slovenskih kmetij.

Pri vprašanju o vplivih deležnikov na njihovo neuporabo je večina vprašanih pripisala trenutno zastavljenim političnim instrumentom, izobraževanju in svetovanju na področju preciznih tehnologij in inovacij v kmetijstvu. Slabih 70 % vprašanih meni, da bi v primeru nižjih cen oziroma nižje začetne investicije pogosteje kupovali naprednejše tehnologije za pomoč na kmetijah. Prav tako imajo pomisleke zaradi možnih nastalih vzdrževalnih stroškov in tehničnih podpor v primeru okvar.

Zanimalo nas je tudi, če anketirani poznajo prednosti uporabe preciznih tehnologij. Med poglavitne razloge za začetek uporabe so izbrali: prijaznost do uporabnika, sledljivost delovnih procesov in dobičkonosnost. Medtem ko so kot največji potencial navedli različne platforme za vodenje kmetijskih evidenc in drugih podatkov. Na podlagi analiz smo ugotovili, da je sprejemanje novih tehnologij iz strani kmetov odvisno tudi od drugih dejavnikov.

V Sloveniji imata na uporabo naprednejših tehnologij zagotovo vpliv velikostna in agrarna struktura kmetij, saj je povprečna velikost kmetijskega gospodarstva leta 2016 znašala le 6, 9 ha. Ugotovljeno je bilo, da države z večjimi kmetijami te tehnologije prevzemajo pogosteje in jih uporabljajo v večji meri, kar pa je sicer povezano s kmetijskim ekonomskim in proizvodnim obsegom v smislu višjih donosov. Na (ne) sprejemanje vplivajo tudi socialne strukture. Velik problem je visoka povprečna starost nosilcev kmetij v Sloveniji (leta 2016, 57 let). Bolj spodbudna je izobrazbena struktura, saj se povečuje delež gospodarjev z različnimi oblikami formalne kmetijske izobrazbe. Dokazano je, da so mlajši nosilci bolj spretni pri upravljanju in bolj razvojno naravnani, kar vpliva na vpeljavo novih, inovativnih tehnologij v njihovo okolje. Za nemoteno uporabo tehnologij preciznega kmetovanja so nujne osnovne internetne povezave, kar pa je zaradi slabe infrastrukture in dostopa do širokopasovnih povezav kmetom velikokrat onemogočeno, saj je v Sloveniji še vedno veliko t.i. belih lis predvsem na podeželju.

3 Izbrani projekti

V naslednjih podpoglavjih povzemamo nekaj izbranih projektov, ki se deloma nanašajo na kmetijstvo in na digitalno obdelavo signalov.

3.1 Pilotni projekt – razvoj nizkocenovnega multispektralnega sistema

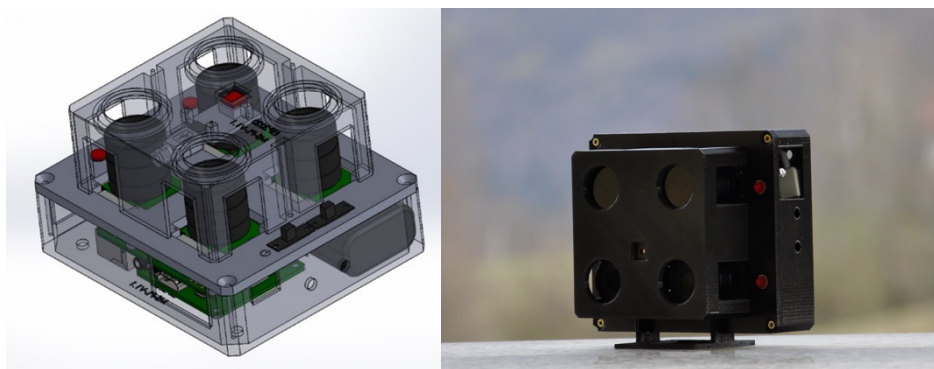
Del tehnologije preciznega kmetijstva temelji na periodičnih pregledih kmetijskih površin, ki se občajno izvajajo s pomočjo satelitskih posnetkov ali s pomočjo posebej opremljenih brezpilotnih letalnikov (Mongus, 2019). To pridelovalcu omogoča izboljšanje proizvodnega in tržnega potenciala. Nenazadnje pa te tehnologije nudijo tudi večjo skrb za okolje.

Izboljšanje proizvodnega potenciala se kaže v smeri obveznih pravo- oz. pred-časnih tretmajev rastlin s fitofarmacevtskimi sredstvi (FFS), ki se običajno nanašajo pavšalno, tehnologije preciznega kmetijstva pa omogočajo bolj premišljeno uporabo (Berk et al., 2019). Na podlagi meritev zajetih v vidnem in bližnjem IR spektru lahko namreč vidimo razlike med zdravimi rastlinami in rastlinami, ki so v stresu. Tako lahko rastlinam zagotovimo dovolj hrane ali pa jih zaščitimo ob zaznavi potencialnih žarišč, kar lahko opravimo na podlagi posameznih opazovanj ali na podlagi predvidevanj iz več preteklih sezon. S selektivno obdelavo lahko zmanjšamo cikle ali posamezno aplikacijo tretiranja s FFS. S tem pa zmanjšamo njihove negativne vplive, kot je npr. kopičenje v okolju, kjer se njihov negativni potencial lahko celo množi in posledično zmanjšuje sam proizvodni potencial tal.

Posredni rezultat takšne tehnologije ima seveda tudi tržni potencial (Lepej et al. 2016 in 2017), saj bolj kvaliteten pridelek dosega boljšo odkupno ceno in ima večji domet na trgu doma in v tujini (npr. omejitve glede kumulative uporabe fito-farmacevtskih sredstev). Po drugi strani odločna in pravočasna, premišljena dejanja pridelovalca storjena na podlagi analize predstavljenega sistema omogočajo prihranke pri vhodnih surovinah (pogonska goriva, fito-farmacevtska sredstva, gnojila,...), kar še dodatno povečuje konkurenčno prednost, obenem z manjšimi negativnimi posledicami na okolje.

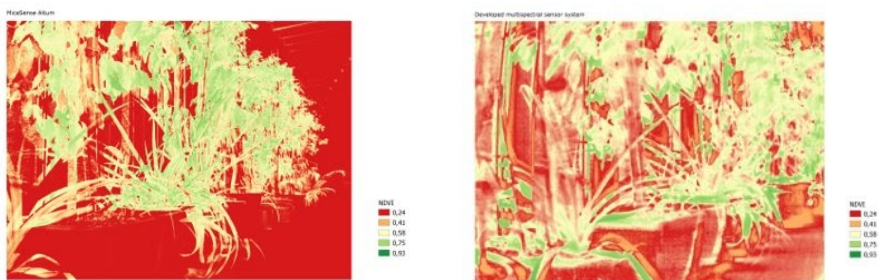
Imajo pa običajno uporabljeni pristopi tudi svoje omejitve. Satelitski posnetki so omejeni glede na pogostost zajetih podatkov, vremenske pogoje ob zajemu in nenazadnje glede na ločljivost, ki znaša nekje okoli 10 m x 10 m na en slikovni element. Te omejitve lahko rešimo z uporabo brezpilotnih letalnikov in multi- ali hiper-spektralnih kamer. Slabost te tehnologije pa se skriva v sami ceni in usposobljenosti operaterjev.

V sklopu ukrepa M16.2 tako izvajamo pilotni projekt, v katerem preučujemo možnost razvoja in uporabe nizkocenovne rešitve, ki bi tehnologijo lahko napravila bolj dostopno za uporabnike. V ta namen smo izdelali svoj nizkocenovni multispektralni senzor, ki temelji na uporabi dostopnih računalniških komponent. Te zajemajo računalnik RaspberryPI 3, eno 5 MP RGB kamero z nameščenim IR filtrom za opazovanje vidnega dela svetlobnega spektra, tri 5MP NOIR kamere z nameščenimi filtri v vidnem in NIR spektru, 1 FLIR PureThermal 2 Lepton kamero, multiplekstersko vezje za povezavo vseh kamer in akumulatorsko enoto za napajanje celotnega sistema. Slika 1 prikazuje načrt in izdelano multispektralno kamero, slika 2 pa primerjavo dobljene slike v primerjavi s sliko iz komercialne multispektralne kamere Altum Micasense.



Slika 1: Načrt (levo) in izdelana multispektralna kamera (desno).

Vir: lasten.



Slika 2: Izračun NDVI indeksa s pomočjo meritev kamere Altum Micasense (levo) in NDVI indeks izračunan iz nizkocenovno dostopne kamere (desno).

Vir: lasten.

3.2 Rovitis 4.0

Del prihodnosti kmetijstva bodo nedvomno avtonomni roboti (Lepej, 2017), ki bodo opravljali različna dela na poljih, v vinogradih, v sadovnjakih in drugje. Delo bo opravljeno hitreje, natančneje in bolj varno, saj fizična prisotnost ljudi ne bo več potrebna. Roboti bodo delo opravili avtonomno, najprej ob nadzoru operaterja, kasneje pa tudi brez neposrednega nadzora operaterja. S tem se bo spremenila tudi paradigma nakupa kmetijske mehanizacije, kjer so kmetovalci v zadnjih letih posegali po vedno večjih, težjih in dražjih strojih, ki so s svojo močjo načeloma res lahko opravili več, a jih je bilo v primeru odpovedi stroja težko nadomestiti. Uporaba avtonomnih strojev pa ponuja drugo možnost; več cenovno dostopnih avtonomnih naprav, ki delajo v skupini in v primeru odpovedi enega izmed strojev, delo opravijo ostali. V sklopu projekta Rovitis 4.0 (Pantano, 2020) smo izračunali, da se strošek nakupa takšnega stroja s ceno okoli 80 000 eur v Italiji povrne že v približno 4 letih, saj za njegovo uporabo ne potrebujemo dodatnega traktorista, en operater pa lahko istočasno opravlja nadzor več takšnih strojev.



Slika 3: Avtonomni vinogradniški robot Rovitis 4.0.

Vir: lasten.

Robot Rovitis 4.0, prikazan na sliki 3, je bil financiran kot italijanski EIP AGRI projekt, kjer je Univerza v Mariboru nastopala kot edini tuj partner znotraj vseh dosedanjih njihovih EIP AGRI razpisov. Naša naloga je bila razvoj programske opreme na avtonomnem robotu, kjer smo uporabili meta robotski sistem ROS, znotraj tega vzpostavili podporo gonilnikov za vso uporabljeno senzoriko in visokonivojsko logiko, ki jo za delovanje robot potrebuje. Ta del je zajemal vozlišča, ki skrbijo za fuzijo senzorskih podatkov, lokalizacijo sistema v prostoru, načrtovanje in sledenje poti, podporo za delovanje v strukturiranih okoljih podprto s proženjem aktuatorskih sistemov, »teach & repeat« funkcionalnost in drugo. Ob tem je potrebno poudariti, da sistem uporablja nizkocenovne komponente, kar sistem res naredi potencialno bolj dostopen za široko množico uporabnikov, a po drugi strani zahteva podporo ustrezne visokonivojske programske logike, ki nekoliko manj robustno delovanje enega sistema kompenzira z drugim.

3.3 Transfarm 4.0

Projekt Transfarm 4.0 (Interreg CE) naslavlja tematiko uporabe tehnologij preciznega kmetijstva. V ta namen partnerji iz petih različnih srednjeevropskih držav vzorčimo področje, ugotavljamo potencialne priložnosti in prenašamo dobre prakse na mednarodnem nivoju. Del aktivnosti pa zajema izvedbo treh pilotnih podprojektov, kjer se prvi nanaša na tematiko ISOBUS aplikacij, drugi na bližnje zaznavanje in tretji obdelavo gruče podatkov (»big data«).

Drugi pilotni projekt tako naslavlja izzive bližnjega zaznavanja lastnosti kmetijskih površin oz., še natančneje, lastnosti rastlin, ki jih želimo ustrezno tretirati. V projektu naslavljamo tudi enega izmed ciljev Evropske komisije, ki je postavila cilj, da do leta 2030 zmanjšamo količino porabljenih FFS za 50 %. V ta namen smo ob pomoči LiDARsko podportega sistema izdelali krmilno enoto, ki vklaplja elektro-magnetne ventile na mestih, kjer je zaznana prisotnost rastlin, in izklaplja na mestih, kjer rastlin ni.

Da je to mogoče, sistem uporablja 2 LiDARski tipali, kjer je prvo nameščena horizontalno, drugo pa vertikalno. Namen horizontalnega tipala je zagotoviti podatke o premiku sistema. Podatek o premiku je pomemben, saj se senzorski sistem in šobe ne nahajajo eden ob drugemu, saj bi delovanje šob motilo meritve ali pa zaradi morebitne korozivnosti FFS celo negativno vplivalo na senzor. Zato sistem oceni lokacijo zajema podatka in ugotovi, kdaj se je pršilnik premaknil toliko, da so

na tej lokaciji prisotne šobe. Ob tem se sistem zanaša na meritve SLAM algoritma, ki smo ga razvili (Lepej, 2016). Alternativa SLAM algoritmu bi bila uporaba RTK-GPS sistema, ki pa bi sistem naredilo še bolj zapleten, dražji in posledično slabše dostopen za končne uporabnike. Ob podpori lokalizacije pa lahko uporabimo še meritve vertikalno nameščenega LiDARskega sistema, ki v prostoru »tipa« na katerih višinah so prisotne rastline in na katerih ne. V primeru, da sistem krošnjo zazna, proži ustrezne elektro-magnetne ventile, ko se pršilnik premakne in ko so šobe na pravi poziciji.

Sistem ob podpori omogoča precej bolj natančen nanos FFS. Seveda so prihranki specifični glede na lastnosti posameznega vinograda oz. sadovnjaka. Preliminarni izsledki kažejo, da v običajnih sadovnjakih prihranimo nekje okoli 30 % FFS, v bolj ekstremnih razmerah, kjer nasad pestijo različne bolezni, kot je npr. esca, in del rastlin v vrsti manjka, pa lahko prihranimo 50 % ali več FFS.



Slika N+3: Pršilnik z nameščenim senzorskim sistemom za zaznavo krošenj in lokalizacijo sistema.

Vir: lasten.

3.4 Farmbeast

Ena izmed pomembnih komponent, ki bo v prihodnosti krojila področje preciznega kmetijstva, je priprava ustreznega kadra. Ti bodo nove tehnologije uporabljali, vzdrževali in seveda tudi razvijali. V ta namen na Katedri za biosistemsko inženirstvo že od leta 2008 izvajamo obštudijski projekt Farmbeast (Kajbič et al., 2021). Gre za projekt, kjer k reševanju izzivov sodobnega kmetijstva pristopamo interdisciplinarno, z združevanjem znanj s področja računalništva, elektrotehnike, strojništva in sorodnih drugih ved z namenom rešiti izzive sodobnega kmetijstva.

V sklopu projekta Farmbeast študenti razvijajo manjšega študentskega robota, s katerim se vsako leto pomerimo s podobnimi roboti na mednarodnem tekmovanju Field Robot Event (FRE, 2022). Roboti na tekmovanju FRE so zaradi zahtev ožji od 0,75 m, navadno manjši od 0,5 m in z maso običajno nekje do 20 kg. Vseeno pa na njih tečejo prava robotska okolja (ROS (Stanford Artificial Intelligence Laboratory et al., 2018) in algoritmi, ki so lahko del večjih, tudi komercialnih robotskih sistemov.

Farmbeast tako uporablja distribuiran računalniški sistem, z manj zmogljivim ARM računalnikom (RaspberryPI 3B), namenjen nizkonivojskim procesom, in računalnik z i7 procesorjem za visokonivojske procese (Intel NUC7I7BNH), ki ga prvi vklaplja in izkaplja po potrebi. Robot zajema množico različnih senzorskih sistemov, ki vključujejo MEMS inercialsko enoto, večkanalni LiDARski sistem (Velodyne VLP-16), dve 5 Mpix PoE kameri, dodati SICK TIM310 LiDAR za podporo orodijem, odometrijo, itd. Robota poganjajo štirje BLDC motorji s skupno močjo 800 W in dodatni štirje DC motorji za zasuk posameznega pogonskega sklopa, ki omogoča tudi različne načine vožnje, kot so skid-steer, ackermann in drugi.

Cilj projektne skupine je razviti oz. izboljšati manjšega avtonomnega robota, ki bo oz. je sposoben opravljati na polju različna opravila. To pa je tudi cilj tekmovanja FRE, kjer se študenti z izpopolnjenim robotom pomerijo z drugimi roboti oz. skupinami v 4 + 1 disciplinah. Prve štiri discipline zajemanje osnovo navigacijo, kjer se robot samodejno premika po polju v medvrstnem prostoru rastlin. V sklopu druge naloge se robot samodejno premika po polju, kjer lahko del rastlin tudi manjka, premika pa se po v naprej določenem zaporedju, kar pomeni, da se v danem trenutku mora »zavedati« kje se nahaja. V sklopu tretje naloge roboti izvajajo različna opravila povezana z zaznavo plevelov, z zaznavo obolelih rastlin oz. podobno. V sklopu

četrte naloge pa te rastline ustrezno tretira tako, da jih poškropi, odstrani ali naredi kaj tretjega, odvisno od teme tekmovanja, ki so jo organizatorji določili za trenutno leto.



Slika N+4: Robot Farmbeast, ki je nastal v sklopu interdisciplinarnega študentskega dela.

Vir: lasten.

Običajnim štirim nalogam pa sledi še dodatna disciplina Freestyle, kjer lahko skupine še posebej izrazijo svojo kreativnost in ideje, ki jih imajo na področju robotike v kmetijstvu. Običajno skupine v ta namen izdelajo nove priključke, ki robotom omogoča opraviti dodatna opravila, na hitrejši način, z bolj natančno obdelava, vse to pa običajno popolnoma samodejno.

5 Zaključek

Kmetijstvo je strateška panoga, ki jo moramo ohraniti tudi v prihodnje. To pa bo verjetno mogoče le, če bomo sledili tehnološkemu razvoju, kar bo omogočalo domačim pridelovalcem hrane, da bodo ostali primerljivi v svetovnem merilu. Ob podpori tehnologije bodo lahko pridelali visokokvalitetno hrano, hrano v dovolj

velikih količinah, hrano, ki bo cenovno dostopna, ob vsem tem, pa bodo kmetovali na trajnosten način, ki bo omogočal nadaljevanje kmetijske panoge tudi prihodnjim rodovom.

V sklopu prispevka so predstavljeni nekateri izzivi preciznega kmetijstva, kamor nedvomno sodi tudi računalništvo in, še natančneje, tudi postopki digitalne obdelave signalov. Glede na te so bili predstavljeni štirje projekti, ki pa še potekajo ali pa so se nadavno končali. Vsem tem pa je skupno to, da pri vsakem izmed njih vidimo dodatne možnosti, ki bi rešitve lahko še izboljšale, vse te dodatne rešitve, pa bodo nekoč v pomoč domačim pridelovalcem hrane, pa tudi še širše.

Literatura

- Berk, P., Stajanko, D., Hočevvar, M., Malneršič, A., Jejčič, V., Belšak, A. (2019). Plant protection product dose rate estimation in apple orchards using a fuzzy logic system. *PloS one*. April 2019, vol. 14, no. 4, e0214315, ISSN 1932-6203.
- Field Robot Event, (2022),. <https://www.fieldrobot.com/event/>, (15.2.2022).
- Kajbič, M., Kenda, U., Popič, G., Toš, D., Bernik, R., Lakota, M., Rakun, J. (2021) Farmbeast. V: GRIEPENTROG, H. W. (ur.), *Proceedings of the 18th Field Robot Event 2021, June 8th – 10th, 2021 : Conducted in conjunction with the DLG-Feldtage/ DLG Field Days*. University of Hohenheim, Technology in Crop Production, January 2022. Str. 78-87.
- Lepej, P., Rakun, J. (2016). Simultaneous localisation and mapping in a complex field environment. *Biosystems eng.* October 2016, vol. 150, str. 160-169, ISSN 1537-5110.
- Lepej, P., Rakun, J. (2017). Uporaba avtonomnih mobilnih robotov v kmetijstvu. V: MUNIH, Marko (ur.). *SI robotika*. Ljubljana: Slovenska matica, 2017. Str. 77-88, ilustr. ISBN 978-961-213-266-8. [COBISS.SI-ID 11702868]
- Pantano, M. (2020). Evento finale progetto, Rovtisi 4.0 – Webinar 4.12.2020, https://www.rovitisveneto.it/wp-content/uploads/2020/12/20200923_Rovitis40_conclusion_FINAL.pdf.
- Rihter, E. (2020). Pregled stanja in smernice razvoja preciznega kmetijstva v Sloveniji : magistrsko delo. Maribor: [E. Rihter], 2020. XIII, 67, [5] f., [COBISS.SI-ID 17701891]
- Mongus, Ž. (2019) Uporaba brezpilotnih letalnikov za določanje NDVI indeksa na rastlinah oljne ogrščice : diplomsko delo. Maribor: [Ž. Mongus], 2019. VI, 34. [COBISS.SI-ID 4616236]
- Stanford Artificial Intelligence Laboratory et al. (2018). Robotic Operating System. www.ros.org.

POPOLDANSKA SEKCIJA

Globoko učenje: Praktični nasveti strokovnjakov

UMETNA INTELIGENCA ALI UMETNA INTELIGENTNOST: KJE SMO IN KAM GREMO

ANDREJ KASTRIN

Univerza v Ljubljani, Medicinska fakulteta, Inštitut za biostatistiko in medicinsko informatiko, Ljubljana, Slovenija.
E-pošta: andrej.kastrin@mf.uni-lj.si

Povzetek Zdi se, da kljub nedorečeni definiciji področja, o umetni inteligentnosti (UI) danes vsi vse vemo. V ožjem, raziskovalnem smislu, se UI povezuje tako z razvojem novih računskih metod kot njihovo uporabo v praksi. V prispevku poročamo o rezultatih analize bibliografskih zapisov s področja UI od leta 2000 do danes. Predstavimo konstrukcijo omrežja sopojavnosti ključnih besed v času in analiziramo tematike, s katerimi so se ukvarjali raziskovalci v preteklih dveh desetletjih.

Ključne besede:

umetna
inteligentnost,
analiza
sopojavnosti,
tematski
diagram,
iMetrika,
zbirka
WoS

SCIENCE MAPPING OF THE KNOWLEDGE BASE ON ARTIFICIAL INTELLIGENCE

ANDREJ KASTRIN

University of Ljubljana, Faculty of Medicine, Institute of Biostatistics and Medical
Informatics, Ljubljana, Slovenia.
E-mail: andrej.kastrin@mf.uni-lj.si

Keywords:

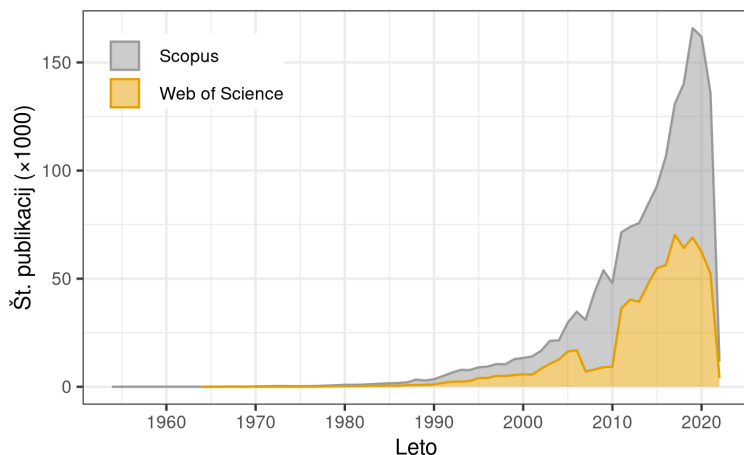
artificial
intelligence,
coword
analysis,
thematic
diagram,
iMetrics,
WoS
database

Abstract It seems that despite the vague definition of artificial intelligence (AI), we are all now top experts in the field. In a narrow sense, AI refers to the development of new computational methods and their application in practice. In this paper, we report the results of an analysis of bibliographic records in the field of AI from 2000 to the present. We present the construction of a keyword co-occurrence network over time and analyze the topics that researchers have worked on over the last two decades.

1 Uvod

Z umetno inteligentnostjo se danes srečujemo že praktično na vsakem koraku. Domuje ne le v znanosti in industriji, temveč na tak ali drugačen način kroji življenje slehernega posameznika. V znanosti praktično ne najdemo področja, ki k sodobnemu pojmovanju umetne inteligentnosti ni pristavilo vsaj kančka svojstvenosti.

Površen vpogled v količino in hitrost pridobivanja znanstvenega védenja nam običajno nudi že hiter sprehod po eni od splošnonamenskih bibliografskih zbirk. Za ilustracijo obsega znanja na področju umetne inteligentnosti je na sliki 1 prikazana rast št. bibliografskih zapisov povezanih s predmetno oznako *Artificial Intelligence*. Prikaz je narejen ločeno za zbirki Scopus in Web of Science (WoS). Denimo, da zanemarimo sicer pomembno razliko v št. publikacij med zbirkama in opazujemo le hitrost naraščanje publikacij v času. Brez zadržkov bomo hitro prišli do zaključka, da je rast literature v zadnjih treh desetletjih enormna.



Slika 1: Rast št. bibliografskih zapisov povezanih s tematiko umetne inteligentnosti v zbirkah Scopus in Web of Science

Vir: lasten

Redno sledenje novim objavam je na tako obsežnem in kompleksnem področju kot je umetna inteligenca, praktično nemogoče. Empirična evidenca pa opozarja, da je vsaj površinsko poznavanje ustroja znanstvenega védenja *conditio sine qua non* za “dobre” raziskovalne dosežke in nove ustvarjalne preboje na področju umetne

inteligentnosti (Chen, 2011). Priročen instrumentarij, ki nam pri iskanju in razumevanju skritih zakonitosti v bibliografskih podatkih lahko pomaga, najdemo v ti. znanosti o znanosti (angl. *Science of Science*).

Za znanost o znanosti—v nadaljevanju bomo uporabljali kratico SciSci—lahko rečemo, da ima dolgo preteklost, a kratko zgodovino. V zadnjih letih je področje, predvsem zaradi uspešnega interdisciplinarnega sodelovanja glavnih akterjev (med katerimi najdemo tudi zveneča imena kot je npr. Albert-László Barabási), dobilo nov raziskovalni zagon (Wang & Barabási, 2021). V prvi vrsti se to odraža v objavah v eminentnih znanstvenih časopisih; revija *Science* pa je predstavitvi discipline nedavno namenila celo številko (glej npr. Fortunato in sod., 2018). Kvantitativni metodološki aparat SciSci s pridom črpa s področij bibliometrike in scientometrike.¹

Na področju SciSci ločimo med pristopi za ocenjevanje uspešnosti raziskovalnih entitet (npr. raziskovalcev, raziskovalnih skupin) in pristopi, ki omogočajo vpogled v strukturo in dinamiko razvoja znanja na izbranem znanstvenem področju. Med zadnjimi najpogostejše srečamo ti. kartiranje znanosti (Small, 2006), analizo socitiranosti dokumentov (Small, 1973) in analizo sopojavnosti (ključnih) besed oz. terminov (Callon in sod., 1991). Nekateri od teh so bili s pridom uporabljeni tudi za pregled znanstvenega védenja na področju umetne inteligentnosti (Alonso in sod., 2018; Darko in sod., 2020; Yu in sod., 2019).

Namen prispevka je dvojen:² (i) z analizo sopojavnosti ključnih besed v znanstveni literaturi želimo predstaviti dinamiko razvoja področja umetne inteligentnosti v svetu in (ii) bralcu želimo nakazati pomen in pomembnost metodologije SciSci za spremljanje in načrtovanje uspešne raziskovalne poti.

¹ Poleg obeh naštetih je moč v literaturi zaslediti še poimenovanja kot so npr. informetrika, webometrika, altmetrika, zato Milojević in Leydesdorff (2013) predlagata uporabo krovne skovanke iMetrika, s katero zajamemo pisano paleto postopkov in metod, ki so bile razvite pod okriljem različnih pristopov k proučevanju znanosti.

² V okviru predavanja bomo problematiko razvoja umetne inteligentnosti naslovili širše in osvetlili tudi dosežke domače raziskovalne skupnosti.

2 Metode

2.1 Pridobivanje podatkov

Raziskovalec ima danes na voljo pisano paleto bibliografskih zbirk, ki se med seboj razlikujejo ne samo po namembnosti in obsegu indeksiranih publikacij, pač pa tudi po kvaliteti vnešenih bibliografskih zapisov in predvsem ceni dostopa. Zbirkam WoS, Scopus in PubMed so se v preteklem desetletju pridružile prostodostopne alternative, med katerimi velja izpostaviti zlasti zbirke Dimensions, OpenCitations, Semantic Scholar in OpenAlex. Sami smo podatke pripravili na osnovi zbirke WoS, ki po naših izkušnjah zagotavlja optimalno razmerje med kvaliteto bibliografskih zapisov in obsegov vključenih publikacij. Podatke smo zajeli 1. februarja 2022, pri priklicu pa smo se omejili na zapise objavljene v angleškem jeziku in uvrščene v kategorijo *Computer Science, Artificial Intelligence*.³

V zbirki WoS smo prešteli 747.018 zapisov z oznako *Computer Science, Artificial Intelligence*, med katerimi je bilo 742.376 (99 %) publikacij v angleškem jeziku. V nadaljevanju smo izdvojili samo zapise za publikacije objavljene v obdobju 2000–2020. Za potrebe identifikacije tematik v času smo 20-letno obdobje še dodatno razbili v štiri manjša obdobja, v vsakem razbitju pa upoštevali samo prvih 2000 zapisov z najvišjim št. citatov.

2.2 Omrežje sopojavnosti ključnih besed

Iz vsakega zapisa smo izluščili ključne besede, ki so jih pripravili avtorji (polje DE v zapisu WoS) in jih po potrebi normalizirali. Upoštevali smo samo ključne besede, ki se pojavijo v najmanj petih bibliografskih zapisih. Zaradi preglednosti prikazov smo upoštevali le prvih 250 najpogostejše zastopanih ključnih besed. Ta delovni korak smo zaključili s pripravo štirih matrik sopojavnosti ključnih besed (ena matrika za vsako od štirih časovnih rezin) razsežnosti 250×250 .

Matriko sopojavnosti lahko predstavimo kot neusmerjeno uteženo omrežje, v katerem posamezno vozlišče označuje eno od 250 izdvojenih ključnih besed, povezave med vozlišči pa vzpostavi sopojavnost parov ključnih besed v bibliografskem zapisu. Utež na povezavi predstavlja št. sopojavitev izbranega para

³ Zbirka WoS vsakemu bibliografskemu zapisu pripiše eno ali več predmetnih oznak (angl. *Subject Category*). WoS trenutno obsega 252 predmetnih oznak s področij naravoslovja, družboslovja, humanistike, tehnike in umetnosti.

ključnih besed preko vseh zapisov. Surove frekvence sopojavnosti smo normalizirali po obrazcu (Van Eck & Waltman, 2007):

$$a_{ij} = \frac{mc_{ij}}{c_{ii}c_{jj}} \quad \forall i \neq j,$$

kjer je c_{ij} št. zapisov, v katerih se sopojavita ključni besedi i in j , c_{ii} št. zapisov v katerih se pojavi ključna beseda i , m pa celotno št. zapisov v izbrani časovni rezini.

2.3 Omrežje sopojavnosti ključnih besed

Za identifikacijo raziskovalnih tematik v omrežju smo uporabili Louvainov algoritem gručenja v skupine oz. skupnosti (Blondel in sod., 2008). Algoritem temelji na optimizaciji kriterijske funkcije (ti. modularnosti), s katero merimo razmerje med gostoto povezav znotraj izbrane skupnosti in gostoto povezav zunaj skupnosti. V vsakem od štirih omrežij smo nato izluščili skupnosti vozlišč, za nadaljnjo analizo pa obdržali le tiste, pri katerih je modularnost znašala vsaj 0,7 (tj. skupnosti, ki so močno homogene, hkrati pa zelo heterogene med seboj).

2.4 Tematski diagram

Tematski diagram prikazuje porazdelitev izluščenih skupnosti ključnih besed v ravnini, ki jo opišemo s središčnostjo (angl. *centrality*) in gostoto (angl. *density*) (Callon in sod., 1991). S središčnostjo merimo stopnjo izbrane skupnosti (tj. št. sosednjih skupnosti). Višja kot je središčnost skupnosti, močnejša je aktivnost raziskovalne tematike (in njena interdisciplinarna vpetost), ki jo skupnost zrcali. Središčnost izračunamo po obrazcu:

$$c_L = \sum_{i \in L} \sum_{j \in M} w_{ij} \times e_{ij},$$

kjer je i izbrano vozlišče v skupnosti L , M množica preostalih skupnosti, j pa kazalec na izbrano vozlišče v M . Utež na povezavi med vozliščema i in j označimo z w_{ij} . Indikatorska spremenljivka e_{ij} ob povezavi vozlišč i in j zavzame vrednost 1 in 0 sicer. Z gostoto nasprotno merimo homogenost posamezne skupnosti. Izrazimo jo kot

razmerje med dejanskim ter največjim možnim št. povezav v skupnosti in jo izračunamo po obrazcu:

$$D_L = \frac{2E}{N(N-1)},$$

kjer je E št. povezav in N št. vozlišč v skupnosti L . Gostota skupnosti odraža razvojni potencial raziskovalne tematike.

Glede na porazdelitev središčnosti in gostote lahko vsako od identificiranih tematik uvrstimo v enega od štirih kvadrantov pravokotnega koordinatnega sistema:

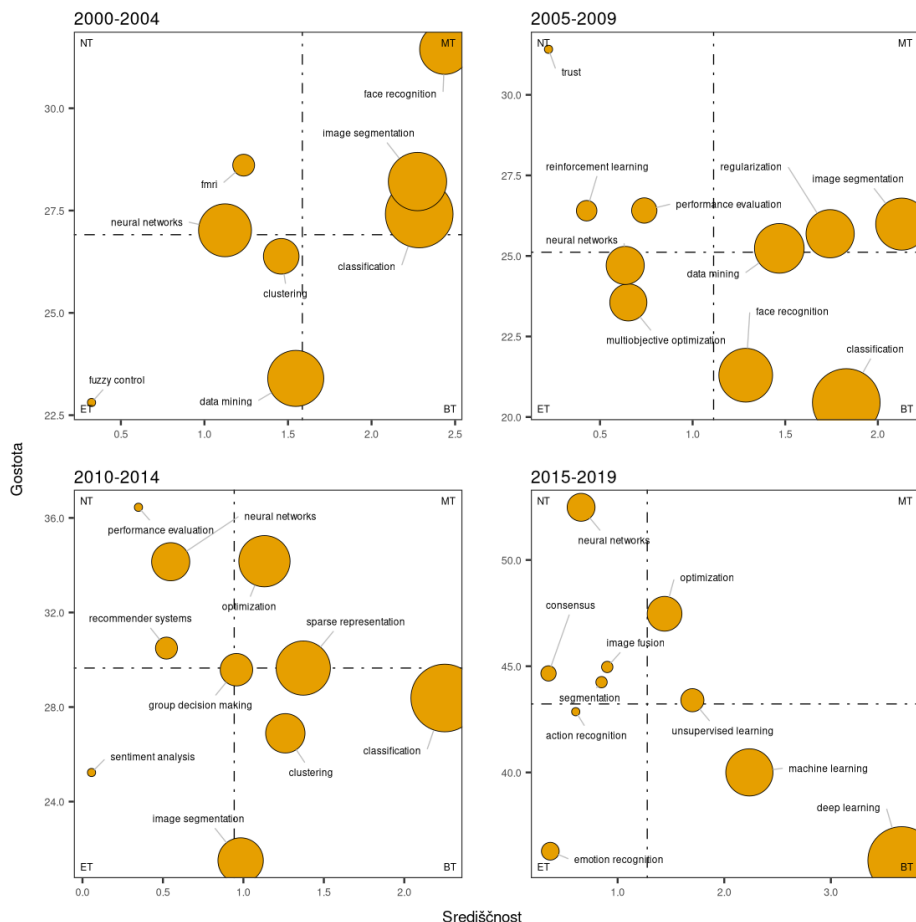
- (I) Skupnosti z močno izraženo središčnostjo in gostoto, ki odražajo pomembne in dobro razvite raziskovalne tematike;
- (II) Skupnosti s šibko izraženo središčnostjo in močno izraženo gostoto, ki zrcalijo sicer dobro razvite raziskovalne tematike, a brez povezanosti z ostalimi področji;
- (III) Skupnosti s šibko izraženo središčnostjo in gostoto odražajo obrobne in slabo razvite raziskovalne tematike, ki so bodisi že v zatonu bodisi se v znanstveni skupnosti šele porajajo;
- (IV) Skupnosti z močno izraženo središčnostjo, a šibko gostoto, ki so za izbrano raziskovalno področje sicer pomembne, nimajo pa razvojnega zagona.

2.5 Ponovljivost rezultatov

Programska koda, ki omogoča reprodukcijo predstavljenih rezultatov, je prostodostopna v avtorjevem podatkovnem prostoru na spletišču GitHub: <https://github.com/akastrin/rosus-2022>.

3 Rezultati in razprava

Na sliki 2 je prikazana semantična ravnina, ki jo razpenjata središčnost in gostota, definirana zgoraj v razdelku 2.4. Za vsako od štirih časovnih rezin v obdobju 2000–2020 je na prikazu predstavljena porazdelitev identificiranih skupnosti oz. raziskovalnih tematik, ki jih te zrcalijo. V nadaljevanju sledi kratek vsebinski pregled raziskovalnih tematik v posameznih časovnih rezinah.



Slika 2: Tematski diagrami z identificiranimi raziskovalnimi tematikami s področja umetne inteligentnosti v obdobju 2000–2020

Vir: lasten

Začnimo z obdobjem 2000–2004, v katerem smo identificirali 43.003 bibliografskih zapisov povezanih s predmetnim področjem *Artificial Intelligence*. Porazdelitev zapisov narašča od 5.790 v letu 2000, do 12.617 v letu 2004. Kar dobrih 46 % publikacij je bilo objavljenih v zbirkah *Lecture Notes in Computer Science* in *Lecture Notes in Artificial Intelligence* pri založbi Springer. Po št. citatov (46.933) se na vrh lestvice uvršča članek z naslovom *Random Forests*, ki je bil objavljen v reviji *Machine Learning* (Breiman, 2001). Med 10 najpogostejše citiranimi najdemo še naslova *Latent Dirichlet Allocation* (Blei in sod., 2003) in *SMOTE: Synthetic Minority Over-sampling Technique* (Chawla in sod., 2002). V množici izluščenih ključnih besed smo identificirali osem

tematik; najmanjša (*fuzzy control*) združuje 46, največja (*classification*) pa 253 ključnih besed. Slednja se, skupaj s tematikama *image segmentation* in *face recognition*, ponaša tako z visoko središčnostjo, kot gostoto. Za tematiko *neural networks* je značilna podpovprečna središčnost, ki v kombinaciji s povprečno gostoto govori v prid razvitemu, a šibko aktivnemu področju raziskovanja nevronske mreže. Obrnjeno je v tem obdobju tematika *data mining* sicer aktivna, vendar se še razvija.

V časovni rezini 2005–2009 smo identificirali 57.268 bibliografskih zapisov s proučevanega domenskega področja. Št. zapisov upada od leta 2005, v katerem je bilo objavljenih 16.296 prispevkov, do leta 2009, v katerem smo zabeležili le 9.041 vnosov v WoS. Med publikacijami najvišji mesti zasedata že omenjeni Springerjevi publikaciji iz serije *Lecture Notes*. Primat po št. citatov (12.582) pripada članku *Visualizing Data using t-SNE* (van der Maaten & Hinton, 2008), na neskromnem sedmem mestu pa najdemo tudi prispevek domačega avtorja (Demšar, 2006), ki je po naših poizvedbah najpogostejše navajanje članek s širšega področja umetne inteligentnosti, znanosti o podatkih in uporabne statistike. Louvainov algoritem je iz omrežja sopojavnosti izdvojil 10 raziskovalnih tematik, med katerimi je najmanjša (*trust*) vključevala le 10, največja (*classification*) pa 310 ključnih besed. Pozoren bralec bo opazil, da tematika *image segmentation* v primerjavi s časovno rezino 2000–2004 ostaja v istem kvadrantu. Podobno je pozicionirana tudi tematika *neural networks*, ki pa je po obsegu nekoliko manjša, saj se je del pomensko sorodnih ključnih besed porojil v tematiko *reinforcement learning*. Tematika *data mining* je v obdobju 2005–2009 postala bolj aktivna in raziskovalno zanimiva, hkrati pa beležimo tudi njen hitrejši razvoj.

Pomaknimo se po časovnem traku naprej. Obdobje 2010–2014 je v zbirki WoS zastopano z 170.973 predmetnimi zapisi, kar je v primerjavi s predhodnim obdobjem skoraj trikrat več. Od 9.246 zapisov v letu 2010, se števec ob koncu leta 2014 ustavi pri 47.449 bibliografskih vnosih. Pomembna sprememba nastopi v porazdelitvi tipov objav, saj so raziskovalci kar 64 % prispevkov objavili v obliki konferenčnega članka. Št. člankov kitajskih avtorjev tudi prvič preseže št. člankov, ki jih objavijo raziskovalci iz ZDA. Na lestvici najpogostejše citiranih izstopata aplikativna prispevka *Scikit-Learn: Machine Learning in Python* (Pedregosa in sod., 2011) in *LIBSVM: A Library for Support Vector Machines* (Pedregosa in sod., 2011) ter konferenčni članek *Generative Adversarial Nets* (Goodfellow in sod., 2014). Prva dva naslova sta dovolj povedna, da bo vsebino zlahka prepoznal že slušatelj dodiplomskega študija, v zadnjem pa avtorji predstavijo idejo generativnih

kontradiktornih mrež (angl. *Generative Adversarial Nets*). Med izluščenimi ključnimi besedami smo podobno kot zgoraj identificirali 10 raziskovalnih tematik. Najmanjšo skupnost (*sentiment analysis*) sestavlja 10 elementov, največjo (*classification*) pa 250 ključnih besed. V kvadrantu I se pojavi nova tematika (*optimization*), ki predstavlja pomembno gonilno silo razvoja. Podobno se razvojni potencial okrepi pri tematiki *classification*, ki hkrati pridobi tudi večjo pozornost raziskovalcev. Nasprotno ugotavljamo počasen zaton tematike *image segmentation*, pri kateri je stopnja razvoja in povezovanja s sorodnimi področji zelo nizka. Vsebinsko široka tematika *sparse representation* je povezana s prizadevanjem raziskovalcev po obvladovanju masivnih podatkov. V kvadrantu III se poraja nova tematika *sentiment analysis*.

Zadnje obdobje je po obsegu vključenih bibliografskih enot najbogatejše. V rezini 2015–2019 smo v zbirki WoS identificirali 313 602 zapisov. Porazdelitev zapisov po letih je v primerjavi s prejšnjimi časovnimi rezinami bolj enakomerna; leta 2015 beležimo 54.879 zapisov, največje št. (69.639) pa leta 2017. Delež prispevkov, objavljenih v obliki konferenčnega članka se še poveča (70 %). Najpogosteje citiran je konferenčni članek z naslovom *Deep Residual Learning for Image Recognition* (He in sod., 2016). Vseh 10 najpogosteje citiranih člankov se ukvarja z globokim učenjem v povezavi s širšim področjem računalniškega vida. V postopku priprave tematskega diagrama smo identificirali 10 skupnosti ključnih besed. Najmanjša skupnost (*action recognition*) vključuje 11 ključnih besed, najobsežnejša (*deep learning*) pa 357 ključnih besed. Tematika *deep learning* v povezavi s tematiko *machine learning* tvori bazično področje delovanja raziskovalcev. Pomembnih gonilnih tematik v tem obdobju ne identificiramo. Povedna je transformacija tematike *sentiment analysis*, ki smo jo identificirali v obdobju 2010–2014, v tematiko *emotion recognition*, ki zrcali prizadevanja po obvladovanju večrazrednega napovedovanja/razvrščanja emocij v različnih modalnostih.

4 Zaključek

V prispevku smo na kratko osvetlili pot razvoja umetne inteligentnosti v zadnjih dvajsetih letih. Predstavljena analiza nikakor ni izčrpna, je pa dovolj povedna, da bo bralec dobil prvi vpogled v dinamiko razvoja področja. Zaradi specifičnosti podatkovnega modela bibliografske zbirke WoS, v analizo nismo vključili raziskovalnih področij, ki umetno inteligentnost pravzaprav definirajo (npr. filozofija, vedenjske in kognitivne znanosti).

Literatura

- Alonso, J. M., Castiello, C., & Mencar, C. (2018). A bibliometric analysis of the explainable artificial intelligence research field. V J. Medina, M. Ojeda-Aciego, J. L. Verdegay, D. A. Pelta, I. P. Cabrera, B. Bouchon-Meunier & R. R. Yager (Ur.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations* (str. 3–15). Springer. doi:10.1007/978-3-319-91473-2_1
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993–1022.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. doi:10.1088/1742-5468/2008/10/P10008
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. doi:10.1023/A:1010933404324
- Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics*, 22(1), 155–205. doi:10.1007/BF02019280
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1–27. doi:10.1145/1961189.1961199
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. doi:10.1613/jair.953
- Chen, C. (2011). *Turning points: The nature of creativity*. Springer.
- Darko, A., Chan, A. P. C., Adabre, M. A., Edwards, D. J., Hosseini, M. R., & Ameyaw, E. E. (2020). Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities. *Automation in Construction*, 112, 103081. doi:10.1016/j.autcon.2020.103081
- Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research*, 7, 1–30.
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., & Barabási, A.-L. (2018). Science of Science. *Science*, 359(6379), eaao0185. doi:10.1126/science.aao0185
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, 2672–2680.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. doi: 10.1109/CVPR.2016.90
- Milojević, S., & Leydesdorff, L. (2013). Information metrics (iMetrics): A research specialty with a socio-cognitive identity? *Scientometrics*, 95(1), 141–157. doi:10.1007/s11192-012-0861-z
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-Learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830.
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265–269. doi:10.1002/asi.4630240406
- Small, H. (2006). Tracking and predicting growth areas in science. *Scientometrics*, 68(3), 595–610. doi:10.1007/s11192-006-0132-y
- van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(86), 2579–2605.
- Van Eck, N. J., & Waltman, L. (2007). Bibliometric mapping of the computational intelligence field. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 15(05), 625–645. doi:10.1142/S0218488507004911
- Wang, D., & Barabási, A.-L. (2021). *The science of science*. Cambridge University Press. doi:10.1017/9781108610834

Yu, D., Xu, Z., & Fujita, H. (2019). Bibliometric analysis on the evolution of applied intelligence. *Applied Intelligence*, 49(2), 449–462. doi:10.1007/s10489-018-1278-z

DOPOLDANSKA SEKCIJA

Industrijske aplikacije

Medicinske in biomedicinske aplikacije

Drugo

Študentske aplikacije

SEGMENTACIJA TELESA Z UPORABO VEČCILJNEGA UČENJA

JULIJAN JUG,¹ AJDA LAMPE,¹ PETER PEER¹ IN
VITOMIR ŠTRUC²

¹ Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Ljubljana, Slovenija.

E-pošta: julijan.jug@gmail.com, ajda.lampe@fri.uni-lj.si, peter.peer@fri.uni-lj.si

² Univerza v Ljubljani, Fakulteta za elektrotehniko, Ljubljana, Slovenija.

E-pošta: vitomir.struc@fe.uni-lj.si

Povzetek Segmentacija je pomemben del številnih problemov računalniškega vida, ki vključujejo človeške podobe, in je ena ključnih komponent, ki vpliva na uspešnost vseh nadaljnjih nalog. Več predhodnih del je ta problem obravnavalo z uporabo večciljnega modela, ki izkorišča korelacije med različnimi nalogami za izboljšanje uspešnosti segmentacije. Na podlagi uspešnosti takšnih rešitev v tem prispevku predstavljamo nov večciljni model za segmentacijo/razčlenjevanje ljudi, ki vključuje tri naloge, tj. (i) napoved skeletnih točk, (ii) napoved globinske predstavitve poze in (iii) segmentacijo človeškega telesa. Glavna ideja predlaganega modela Segmentacija-Skelet-Globinska predstavitev (ali na kratko iz angleščine SPD) je naučiti se boljšega modela segmentacije z izmenjavo znanja med različnimi, a med seboj povezanimi nalogami. SPD temelji na skupni hrbtenici globoke nevronske mreže, ki se razcepi na tri glave modela, specifične za naloge, in se uči z uporabo cilja optimizacije za več nalog. Učinkovitost modela je analizirana s strogimi eksperimenti na nizih podatkov LIP in ATR ter v primerjavi z nedavnim (najsodobnejšim) večciljnim modelom segmentacije telesa. Predstavljene so tudi študije ablacije. Naši eksperimentalni rezultati kažejo, da je predlagani večciljni (segmentacijski) model zelo konkurenčen in da uvedba dodatnih nalog prispeva k večji skupni uspešnosti segmentacije.

Ključne besede:

računalniški
vid,
segmentacija,
razločanje
delov
telesa,
večciljno
učenje,
mreža
ResNet-101

BODY SEGMENTATION USING MULTI-TASK LEARNING

JULIJAN JUG,¹ AJDA LAMPE,¹ PETER PEER¹ &
VITOMIR ŠTRUC²

¹ University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia.

E-mail: julijan.jug@gmail.com, ajda.lampe@fri.uni-lj.si, peter.peer@fri.uni-lj.si

² University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenia.
E-mail: vitomir.struc@fe.uni-lj.si

Abstract Segmentation is an important step in many computer vision problems involving human images and one of the key components that affects the performance of all downstream tasks. Several prior works have approached this problem using a multi-task model that exploits correlations between different tasks to improve segmentation performance. Based on the success of such solutions, we present in this paper a novel multi-task model for human segmentation/parsing that involves three tasks, i.e., (i) keypoint-based skeleton estimation, (ii) dense pose prediction, and (iii) human-body segmentation. The main idea behind the proposed Segmentation-Pose-DensePose model (or SPD for short) is to learn a better segmentation model by sharing knowledge across different, yet related tasks. SPD is based on a shared deep neural network backbone that branches off into three task-specific model heads and is learned using a multi-task optimization objective. The performance of the model is analysed through rigorous experiments on the LIP and ATR datasets and in comparison to a recent (state-of-the-art) multi-task body-segmentation model. Ablation studies are also presented. Our experimental results show that the proposed multi-task (segmentation) model is highly competitive and that the introduction of additional tasks contributes towards a higher overall segmentation performance.

Keywords:

computer
vision,
segmentation,
human
body
parsing,
multi-task
learning,
ResNet-101
net

1 Uvod

V zadnjih letih je bil na področju računalniškega vida dosežen velik napredek. Sodobni generativni modeli, kot so GAN (angl. generative adversarial network), so omogočili ustvarjanje fotorealističnih slik s prepričljivo vizualno kakovostjo. Veliko raziskav je osredotočeno tudi na uporabo takšnih modelov. Eden takšnih izzivov je generiranje fotorealističnih podob ljudi v želenih oblačilih ali problem virtualnega pomerjanja (Han, 2018; Fele, 2022). Takšne aplikacije imajo velik potencial za uporabo v spletnih trgovinah z oblačili in izboljšanje uporabniške izkušnje spletnega nakupovanja.

Z razvojem globokih nevronske mreže je bil velik preskok tudi na področju semantične segmentacije (Chen, 2018; Wang, 2020). Vendar pa je na nekaterih področjih, kot je segmentacija človeškega telesa, še veliko prostora za izboljšave. Trenutno najboljši modeli segmentacije še vedno ne delujejo tako dobro, kot bi morali za uporabo v aplikacijah, kot je navidezno pomerjanje oblačil. Večino težav trenutnim modelom povzročajo slike, posnete v slabših svetlobnih pogojih, in delno prikriti pogledi na subjekt.



Slika 1: Na tem primeru vidimo, da dodatni nalogi skeleta in globinske predstavitve doprineseeta koristne kontekstualne in strukturne informacije o človeškem telesu. Druga slika prikazuje segmentacijsko masko, ki jo izdelata naš večciljni model, ki vsebuje naloge za segmentacijo in napoved skeleta. Tretja slika prikazuje segmentacijsko masko, ki jo je ustvaril naš večciljni model z dodatno nalogo napovedi globinske predstavitve poze. Vidimo lahko, da dodatna naloga bistveno izboljša uspešnost segmentacije.

Vir: lasten.

Pred kratkim je bilo opravljenih veliko raziskav na temo izboljšanja takšnih modelov z uporabo dodatnih informacij za spodbujanje in podporo modelov segmentacije. Z zagotavljanjem dodatnih kontekstualnih informacij se domneva, da lahko model bolje razume vsebino slike in človeško anatomijo.

Obstoječe delo je torej usmerjeno v združevanje modelov segmentacije z drugimi sorodnimi nalogami v tako imenovani večciljni arhitekturi.

Najpogostejše obstoječi modeli vključujejo napoved skeletnih točk kot podporno nalogo, na primer (Gong, 2017).

Prejšnje raziskave so tudi pokazale, da uporaba večciljnega učenja prispeva h kakovosti segmentacije ljudi. Na podlagi tega vpogleda v tem prispevku raziskujemo možnosti za razširitev tovrstnih modelov z dodatnimi nalogami, ki bi lahko dodatno pripomogle k procesu segmentacije.

Medtem ko večina obstoječega dela vključuje napoved skeletnih točk kot podporno nalogo, se naše delo osredotoča na izboljšanje kakovosti rezultatov segmentacije z uporabo dodatne naloge. V ta namen predlagamo novo arhitekturo večciljnega modela, ki poleg naloge segmentacije telesa vključuje še nalogo napovedi skeletnega položaja oziroma drže in globinske predstavitve poze.

Predlagan model smo poimenovali SPD. Črke v imenu predstavljajo naloge: S segmentacija (angl. segmentation), P skelet (angl. pose) in D globinska poza (angl. dense pose). Predlagamo večciljno arhitekturo, ki temelji na skupni nevronske mreži z uporabo treh specializiranih vej na vrhu, po eno za vsako od izbranih nalog. Namen takšnega pristopa je izboljšati nalogo segmentacije.

Predlagani model ocenjujemo na naboru podatkov LIP in ATR in poročamo o zelo spodbudnih rezultatih.

Izvajamo tudi obsežne študije ablacije, da podpremo našo hipotezo, da dodajanje nalog izboljša splošno učinkovitost modela.

Glavni prispevki našega dela so:

- Predstavljamo SPD, nov večciljni model za segmentacijo človeškega telesa, ki vključuje naloge za napoved skeleta in napovedovanje globinske predstavitve poze.

- Pokazali smo, da dodajanje dodatnih nalog izboljša zmogljivost za primarno nalogo.

2 Sorodna dela

Ena izmed bolj specializiranih aplikacijskih področij semantične segmentacije je segmentacija človeškega telesa in oblačil. Potreba po takih algoritmih za segmentacijo izhaja iz zahtev različnih sistemov računalniškega vida, povezanih z analizo človeške podobe, kot je navidežno pomerjanje oblačil (Han, 2018; Fele, 2022) ali ponovna identifikacija (Zhao, 2013).

V zadnjem času je bilo opravljenih veliko raziskav na temo segmentacije ljudi (Liang, 2015; Liang, 2016) z uporabo globokih konvolucijskih nevronske mreže. Pomanjkljivost teh modelov je v tem, da ne upoštevajo strukture oziroma anatomije človeškega telesa, zato segmentacije pogosto vsebujejo napake, ki so s človeškega vidika nerazumne.

Veliko raziskav se je zato osredotočilo na reševanje tega problema z vključitvijo dodatnih informacij v postopek segmentacije, povezane s telesno držo in anatomijo.

Eden od načinov za uvedbo podpornih informacij v model je pristop učenja z več nalogami, kjer se model hkrati uči reševanje več nalog. Zaradi dobrih rezultatov v zadnjih letih se je večciljno učenje pogosto uporabljalo v različnih aplikacijah za obdelavo naravnega jezika in računalniškega vida (Kokkinos, 2017; Eigen, 2015; Bischke, 2019). Gong *et al.* (Gong, 2017) je na primer predlagal model, ki generira semantične segmentacijske maske in položaje skeletnih točk na podlagi generirane segmentacije. Model je optimiziran na podlagi kakovosti segmentacij in lokacij sklepov in s tem zagotovi, da se model uči semantično dosledne predstavitve človeškega telesa. Liang *et al.* (Liang, 2019) gradi na tem pristopu z uporabo skupne osnovne mreže, ki mu sledita dva manjša modula, specializirana za napoved skeletnih točk in semantične segmentacije. Moduli so zgrajeni na dvostopenjski način tako, da najprej generirajo približne in nato natančne rezultate in si pri tem delijo vmesne približne rezultate.

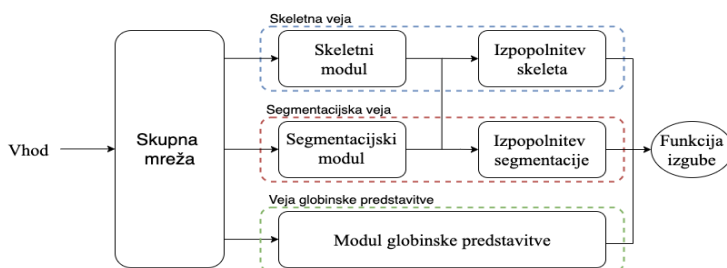
Njihov predlagan model, imenovan JPPNet, dosega impresivne rezultate in prepričljivo presega prejšnje delo. Vendar pa je še vedno prostor za izboljšave v modelovi predstavitvi telesa, saj nekatere strukturne pomanjkljivosti še vedno ostajajo.

3 Metodologija

Predlagamo večciljni model, imenovan SPD, za segmentacijo človeškega telesa, ki se uči na podlagi treh nalog: generiranje segmentacijske maske, napoved položaja skeletnih točk in napoved globinske predstavitve telesa (Guler, 2018). Model je navdihnjen z uspehi obstoječih večciljnih modelov, kot je JPPNet (Liang, 2019), za katere se je izkazalo, da zagotavljajo konkurenčne rezultate, obenem pa imajo tudi zaželene arhitekturne značilnosti.

3.1 Pregled arhitekture modela

Slika 2 prikazuje osnovno arhitekturo našega modela, ki je sestavljena iz skupne mreže za ekstrakcijo značilnosti in treh ločenih vej: (i) za segmentacijo človeškega telesa, (ii) za napoved skeletnih točk in (iii) za napovedovanje globinske predstavitve. Glavni cilj modela je zagotoviti učinkovito segmentacijo telesa, zato je segmentna veja obravnavana kot glavna komponenta modela, preostali dve veji pa opravljata pomožne naloge. Glavni model hrbtenice, ki je skupen vsem nalogam, je ResNet-101 (He, 2016) globoka rezidualna mreža, ki je sestavljeno iz 101 konvolucijskih plasti, razporejenih v 5 rezidualnih blokov. V modelu SPD si del te hrbtenice delijo tri veje, kar deluje kot povezava med tremi obravnavanimi nalogami.



Slika 2: Visokonivojski arhitekturni diagram predlaganega modela SPD. Skupno mrežo ResNet si delijo tri specializirane veje modela, zasnovane za segmentacijo človeškega telesa, napovedovanje skeletnih točk in napoved globinske predstavitve poze.

Vir: lasten.

Tri veje omogočajo definicijo treh ločenih učnih ciljev, ki se nato skupaj uporabijo za učenje modela. Natančneje, skupna funkcija izgube se izračuna kot utežena vsota treh izgub, specifičnih za naloge, t.j.:

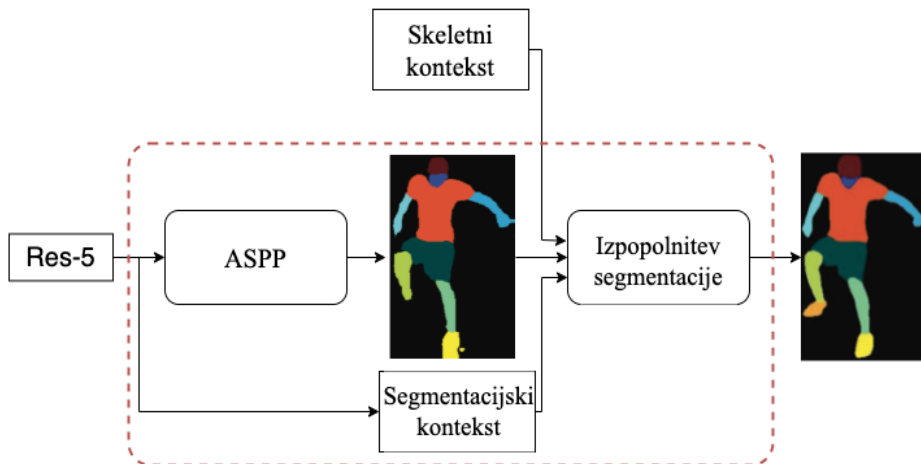
$$\mathcal{L} = \lambda_s \mathcal{L}_s + \lambda_p \mathcal{L}_p + \lambda_d \mathcal{L}_d,$$

kjer so λ_s , λ_p in λ_d uteži, ki ustrezajo funkcijam izgube segmentacije $\mathcal{L}_s$, funkciji izgube skeletnih točk $\mathcal{L}_p$ in izgubi globinske predstavitve $\mathcal{L}_d$. Empirično smo izbrali višjo utež za funkcijo izgube segmentacije in nižje vrednosti za drugi dve nalogi, da bi zagotovili prednost segmentacijske naloge. Na podlagi predhodnih poskusov smo izbrali vrednosti $\lambda_s = 1$, $\lambda_p = 0,8$ in $\lambda_d = 0,6$, da zagotovimo dober kompromis med tremi nalogami.

3.2 Segmentacija

Običajno se za učenje segmentacije ljudi uporabljajo samo informacije iz anotacijskih mask. V našem pristopu v model vključimo tudi kontekstualne informacije skeletnih točk neposredno v segmentacijsko mrežo. Slika 3 prikazuje visokonivojski pregled komponent v segmentacijski veji modela.

Kot je razvidno, se izhod petega rezidualnega bloka uporablja kot prvotna predstavitev za segmentacijski modul. Za generiranje prvotne segmentacijske maske se na vrhu ekstrahiranih značilk ResNet uporablja dodatna plast združevanja poroznih prostorskih piramid (ASPP). ASPP nad vhodnimi podatki izvede zajem več konvolucij pri različnih stopnjah vzorčenja in velikostih mask, pri čemer zajema predmete in kontekstualne informacije na različnih skalah. Vzporedno s komponento ASPP ustvarimo segmentacijski kontekst, tako da obdelamo izhod pete rezidualne plasti skozi dve dodatni konvolucijski plasti. Ta kontekst se kasneje uporablja v drugi fazi segmentacijske veje skupaj z drugimi viri informacij za nadaljnjo izpopolnjevanje rezultatov.



Slika 3: Pregled segmentacijske veje modela SPD. Veja je sestavljena iz dveh delov. Prvi generira začetno segmentacijo na podlagi značilnosti, ki jih ustvari skupni model, medtem ko drugi začetno segmentacijo izpopolni z uporabo različnih vrst vhodnih informacij – tudi iz drugih vej.

Vir: lasten.

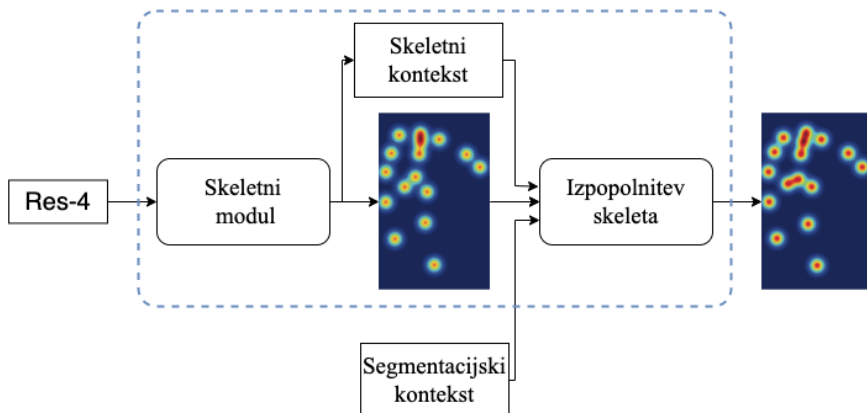
Izpopolnitvena mreža v drugem delu segmentacijske veje vzame kot vhod segmentacijski kontekst in prvotne (grobe) segmentacijske maske in te vhode združi s tako imenovanim skeletnim kontekstom, ki jo ustvari skeletna veja modela. Temu sledijo štirje konvolucijski nivoji, ki služijo zajemu lokalnega konteksta in učenja ključnih povezav med skeletom in segmentacijo. Rezultat teh konvolucijskih plasti je preoblikovan in usmerjen še skozi drugo komponento ASPP. Zadnja ASPP komponenta ustvari končne segmentacijske maske. Izguba segmentacije, se izračuna na koncu te veje in je definirana na nivoju slikovnih elementov kot navzkrižna soft-max entropija, tj:

$$\mathcal{L}_s = \frac{1}{M} \sum_{k=1}^K \sum_{m=1}^M y_m^k \times \log(h_\theta(x_m, k)),$$

kjer je M število vzorcev, K število razredov segmentacije, y_m^k je ciljna klasifikacija za vzorec m in razred k . Vhodni vzorec je označen z x , model napovedi pa z h .

3.3 Veja skeletnih točk

Slika 4 prikazuje visokonivojski pregled nad komponentami, ki sodelujejo pri ustvarjanju predstavitev skeleta. Za razliko od segmentacijske veje je vhod v skeletno vejo izhod četrtega rezidualnega bloka skupne hrbtenice.



Slika 4: Pregled skeletne veje modela SPD. Veja je sestavljena iz dveh delov, kjer prvi generira začetno napoved skeletnih točk na podlagi značilnosti, ki jih ustvari skupen model, medtem ko drugi del prvotno napoved izpopolni z uporabo različnih vrst vhodnih informacij – tudi iz drugih vej.

Vir: lasten.

Začetni skeletni modul v tej veji je sestavljen iz 8 konvolucijskih plasti, prvih šest pridobi skeletne značilnosti, zadnji dve pa generirata prvo različico skeletnih točk v obliki tenzorja s šestnajstimi koordinatami sklepov. Podobno kot v veji segmentacije se v drugi fazi te veje uporablja izpopolnitvena mreža, ki vzame prvotne napovedi skeleta, skeletni kontekst in segmentacijski kontekst, in nato aplicira 4 konvolucije za zajem lokalnega konteksta. Na koncu se za generiranje bolj natančnih koordinat skeletnih točk uporabita dve dodatni konvolucijski plasti.

Funkcija izgube L2 je definirana na koncu skeletne veje, t.j.:

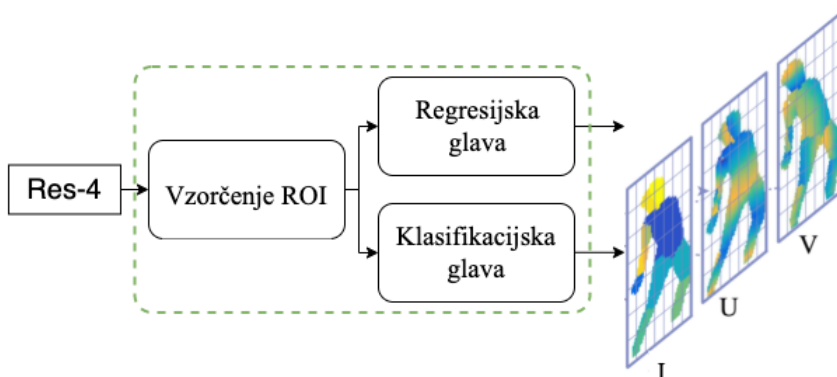
$$\mathcal{L}_p = \frac{1}{2N} \sum_{i=1}^N \|p_i - p'_i\|^2,$$

kjer N predstavlja število definiranih sklepov, p_i' napovedane koordinate sklepov in p_i anotirane koordinate sklepov.

3.4 Veja globinske predstavitve telesa

Slika 5 predstavlja arhitekturo veje globinske predstavitve telesa. Podobno kot v skeletni veji za vhod uporabimo izhod četrtega rezidualnega bloka skupne mreže. Za ResNet mrežo sledi modul za vzorčenje regij interesna (ROI), ki se uporablja za (kaskadni) zajem lokalnih kontekstov na različnih skalah. Nad modulom za vzorčenje ROI je globinska glava, sestavljena iz dveh namenskih CNN glav, klasifikacijske in regresijske.

Prva glava se uporablja za dodelitev slikovnih elementov ustreznemu segmentu telesa, to je klasifikacija komponente I . Druga glava določa položaj slikovnih elementov znotraj ustreznih segmentov, torej se uporablja za določanje komponent U in V .



Slika 5: Slika prikazuje visokonivojski diagram arhitekture veje globinske predstavitve telesa.

Vir: lasten, del slike povzet po (Güler,2018)

Funkcija izgube za vejo globinske predstavitve je sestavljena iz dveh delov. Prvi del se nanaša na komponento I in je izračunan na enak način kot pri glavni nalogi segmentacije, to je z uporabo navzkrižne entropije. Drugi del, ki se nanaša na koordinate U in V , pa se izračuna s pomočjo Huberjeve funkcije izgube. Celotna funkcija izgube te veje je torej izražena kot

$$\mathcal{L}_d = \sum_{m=1}^M CSE(x_I) \cdot L_1(x_U, x_V),$$

kjer so x_I, x_U, x_V komponente predstavitve globine, CSE je transversna entropijska funkcija za segmentacijski del in L_1 je Huberjeva funkcija izgube za koordinatni del.

4 Eksperimenti in rezultati

V tem delu predstavimo izbrane zbirke podatkov. Opišemo protokol, uporabljen za vrednotenje predlaganega modela SPD, in predstavimo mere uspešnosti, ki smo jih uporabili za vrednotenje nalog modela. Nato komentiramo in analiziramo rezultate modela. Izvedemo tudi ablacijsko analizo in pokažemo prispevek posameznih nalog h končni natančnosti segmentacijskega modela. Na koncu predstavimo tudi kvalitativne primere segmentacijskih mask in jih analiziramo.

4.1 Podatkovne zbirke

Izbira nabora podatkov igra pomembno vlogo pri učenju predlaganega modela SPD. Za naše namene smo uporabili več podatkovnih zbirk, ki vsebujejo slike ljudi v različnih oblačilih, situacijah, kontekstih in položajih telesa. Poseben izziv našega pristopa k modeliranju z več nalogami je potreba po zbirki podatkov, ki vsebuje več različnih vrst anotacij.

Za učenje večciljnega modela, ki vključuje generiranje segmentacij, položajev skeleta in globinskih predstavitev poze, potrebujemo nabor podatkov, ki vsebuje vse tri vrste anotacij. V ta namen smo izbrali zbirko LIP (Gong, 2017), ki vsebuje anotacije segmentacije in skeleta za več kot 50.000 slik.

Za anotacije globinskih predstavitev poze smo uporabili bazo podatkov COCO (Lin, 2014), katere podmnožica je zbirka LIP. Združili smo anotacije obeh zbirk podatkov, da smo pridobili referenčne podatke, potrebne za učenje modela SPD. V končni zbirki imamo anotacije z globinsko predstavitvijo za vse vhodne slike, 19-razredno anotacijo segmentacijskih mask in 16-točkovno oznako za skeletne točke.

4.2 Metrike uspešnosti

Po standardni metodologiji ocenjevanja uporabljamo štiri merila uspešnosti za poročanje o uspešnosti za segmentacijske naloge, to so Jaccardov indeks IoU , natančnost, priklic in mera $F1$ (Rot, 2020; Emeršič, 2021).

Prvo merilo je Jaccardov indeks ali uteženo povprečje razmerij prekrivanja in unije površin. Mera IoU je definirana kot:

$$IoU = \sum_{i=1}^K \frac{S_i' \cap S_i}{S_i' \cup S_i},$$

kjer S' predstavlja označeno območje in S anotirano območje razreda i -ega primerka in K število označenih referenčnih razredov.

Ko gledamo na semantično segmentacijo kot na problem klasifikacije na ravni slikovnih pik, je točnost opredeljena kot razmerje pravilno razvrščenih slikovnih pik med vsemi slikovnimi pikami, razvrščenimi v razred, medtem ko je priklic delež pravilno razvrščenih slikovnih pik med vsemi slikovnimi pikami, ki pripadajo razredu, tj:

$$Pr = \frac{TP}{TP + FP}, \quad Rec = \frac{TP}{TP + FN},$$

kjer TP , FP , TN in FN označujejo resnične pozitivne, lažno pozitivne, resnično negativne in lažno negativne.

Rezultat $F1$ mere je harmonično povprečje med točnostjo in priklicem:

Za nalogo napovedi skeleta poročamo o povprečni evklidski razdalji (mED) med predvideno p_i' in referenčno pozicijo sklepov p_i . Mera je definirana kot:

$$mED = \sum_{i=1}^N d_{L_2}(p_i, p_i'),$$

kjer je $d_{L_2}(\cdot)$ funkcija evklidske razdalje, $N=16$ pa je skupno število označenih skeletnih točk.

Za nalogo napovedovanja globinske predstavitve poze uporabljamo merilo podobnosti geodetskih točk med ustvarjenimi in referenčnimi točkami globinske predstavitve, kot je definirano v (Guler, 2018). Mera je definirana kot:

$$GPS = \frac{1}{|P|} \sum_{p_i \in P} \exp \left(\frac{-d(\hat{p}_i, p_i)^2}{2k(p_i)^2} \right).$$

V zgornji definiciji P predstavlja niz označenih površinskih točk, $|\cdot|$ je kardinalnost množice, $\hat{p}_i$ označuje i -to napovedano točko na površini in p_i ustrezno anotirano točko na površini osebe. Funkcija d predstavlja geodetsko razdaljo med točkami in k normalizacijski faktor, specifičen za vsak del telesa.

4.3 Segmentacijski rezultati in ablacijska analiza

S predlaganim modelom SPD želimo izboljšati rezultate obstoječih modelov segmentacije telesa. Natančneje, gradimo na nedavnem pristopu JPPNet iz (Liang, 2019) in zato uporabimo ta model za primerjavo. Tabela 1 prikazuje rezultate segmentacije na zbirki LIP.

Kot je razvidno, na naboru podatkov LIP model SPD doseže IoU rezultat 0,547 v primerjavi z modelom JPPNet, ki ima rezultat 0,538. Glede na mero $F1$ je predlagani model boljši od JPPNet za približno 5%. Podobne izboljšave zmogljivosti so opažene tudi pri natančnosti in priklicu.

Tabela 1: Rezultati segmentacije in ablacije na naboru podatkov LIP. Puščice poleg metrik nakazujejo kakšna vrednost predstavlja boljši rezultat.

Eksperiment	Model	Metrike uspešnosti			
		$IoU \uparrow$	$Pr \uparrow$	$Rec \uparrow$	$F1 \uparrow$
Primerjava	JPPNet	0,538	0,68	0,66	0,66
	SPD (naš)	0,547	0,76	0,68	0,71
Ablacijska analiza	SP	0,535	0,74	0,52	0,63
	SD	0,478	0,67	0,50	0,57
	S	0,483	0,62	0,49	0,54

Za nadaljnje preverjanje delovanja SPD na neodvisnem naboru podatkov z značilnostmi, ki se razlikujejo od učnih podatkov, smo naš model evalvirali tudi na naboru podatkov ATR. Rezultati segmentacije v tabeli 2 ponovno kažejo, da je SPD boljši od JPPNet glede na vse obravnavane mere uspešnosti. Opažene izboljšave učinkovitosti pripisujemo interakciji treh različnih nalog, kar našemu modelu omogoča, da se bolj učinkovito nauči segmentirati slike v različnih situacijah in svetlobnih pogojih.

Tabela 2: Rezultati segmentacije in ablacije na naboru podatkov ATR. Puščice poleg metrik nakazujejo kakšna vrednost predstavlja boljši rezultat.

Eksperiment	Model	Metrike uspešnosti			
		<i>IoU</i> ↑	<i>Pr</i> ↑	<i>Rec</i> ↑	<i>F1</i> ↑
Primerjava	JPPNet	0,464	0,66	0,67	0,66
	SPD (naš)	0,472	0,67	0,70	0,68
Ablacijska analiza	SP	0,423	0,69	0,53	0,60
	SD	0,340	0,59	0,44	0,50
	S	0,291	0,50	0,56	0,52

Da bi prikazali pomen vseh nalog v arhitekturi večciljnega modela SPD, smo izvedli ablacijsko študijo, kjer so iz modela odstranjene različne naloge. Za ta eksperiment so implementirani in naučeni trije dodatni modeli, to so: *(i)* model SPD brez naloge napovedovanja globinske predstavitve (SP v nadaljevanju), *(ii)* model SPD brez naloge za napovedovanje položaja skeletnih točk (SD v nadaljevanju) in *(iii)* model SPD brez obeh nalog, povezanih s pozo (S v nadaljevanju). Rezultati tega poskusa so predstavljeni v tabelah 1 in 2 za nabora podatkov LIP oziroma ATR. Vidi se, da vsaka dodana naloga modelu nudi nove uporabne informacije za izboljšanje rezultatov segmentacije. Odstranitev naloge za napoved globinske predstavitve povzroči padec uspešnosti segmentacije pri vseh merah uspešnosti. Odstranitev naloge za napoved skeletnih točk ima še večji škodljiv učinek na uspešnost. Če sta obe nalogi odstranjeni, opazimo najbolj pomembno poslabšanje zmogljivosti, kar kaže, da obe nalogi, povezani s pozo, zagotavljata pomembne informacije za nadaljnje izboljšanje rezultatov segmentacije. Zanimivo je, da opazimo večje padce zmogljivosti na naboru podatkov ATR kot na LIP. To je verjetno posledica dejstva, da je bil model usposobljen na delu podatkov v LIP, zato so pomožne naloge bolj kritične, ko se spremenijo značilnosti podatkov.

4.4 Rezultati pomožnih nalog

Ker je SPD učen na večciljni način, generira tudi napovedi skeletnih točk in predstavitev vhodnih slik z globinsko predstavitevjo poze. Za boljše razumevanje obnašanja modela tukaj poročamo o rezultatih za naloge napovedovanja skeletnih točk in globinske predstavitve na testnem delu nabora podatkov LIP.

Napoved skeleta. Za prvi poskus ovrednotimo tri modele, predlagani SPD, referenčni JPPNet in SPD model brez naloge napovedovanja globinske predstavitve, to je SP. Na testnih podatkih LIP ima model JPPNet najnižjo vrednost mED 51,2 slikovnih pik, sledi model SPD z vrednostjo 55,01 slikovnih pik. Najšibkejši model v tem poskusu je model SP z vrednostjo mED 56,82 slikovnih pik. Ti rezultati kažejo, da dodatek naloge z globinsko predstavitevjo jasno izboljša tudi učinkovitost naloge napovedovanja skeletnih točk. Vendar so končni rezultati slabši od JPPNet zaradi dejstva, da so naloge segmentacije dobile višjo prioriteto pri uravnovešanju funkcije izgube.

Napoved globinske predstavitve. Tretja naloga, ki se izvaja znotraj modela SPD, je napoved globinske predstavitve telesa. Ker JPPNet ne ustvarja napovedi globinske predstavitve, poročamo samo o rezultatih za celoten model SPD in model brez naloge za napovedovanje položaja na podlagi skeletnih točk, to je SD.

Na testnih podatkih LIP model SPD dosega oceno GPS 48,2%, model SD pa 50,1%. Rezultati kažejo, da dodajanje naloge za napovedovanje položaja skeletnih točk ne pomaga izboljšati učinkovitosti napovedi globinske predstavitve.

4.5 Kvalitativni rezultati

V tem delu predstavimo in analiziramo kvalitativne rezultate, ki jih generira segmentacijska veja modela SPD.

Slika 6 prikazuje primerjavo rezultatov segmentacije, ki jih generirata SPD in JPPNet, skupaj z izvornimi vhodnimi slikami in anotacijskimi maskami segmentacije za tri izbrane slike iz nabora podatkov LIP.

Vhodna slika



Anotacija



JPPNet



SPD



Slika 6: Primerjava rezultatov segmentacije, ki jih generira predlagani model SPD in konkurenčni JPPNet na izbranih slikah iz nabora podatkov LIP. V prvi vrstici so prikazane izbrane vhodne slike, v drugi vrstici segmentacijske anotacije, in v tretji ter četrti vrstici rezultati modelov JPPNet oziroma SPD.

Vir: lasten.

Prva slika prikazuje teniškega igralca in osebo v ozadju, ki je neizostrena in delno zakrita. Vidimo, da je model SPD edini, ki je pravilno zaznal le igralca v ospredju. Referenčni model ima težave z osebo v ozadju, saj je zelo blizu igralca v ospredju.

Razlika v kakovosti segmentacije je vidna tudi pri definiciji prstov na desni roki, kjer je model SPD veliko bolje prepoznal posamezne prste kot model JPPNet. Drugi primer prikazuje žensko, ki je delno skrita za stolom. V tem primeru model JPPNet izpusti celoten predel nog, čeprav je še vedno delno viden skozi stol. Kljub prekrivanju model SPD prepozna položaj noge in jo pravilno označi. Druga edinstvena značilnost te slike je razvrstitev zgornjega dela oblačila. Zgornji del ženskega telesa je označen kot zgornji oblačilni razred, model JPPNet ga napačno klasificira kot plašč, medtem ko model SPD pravilno razvršča območje kot razred zgornjih oblačil, kar je posledica kontekstualnih informacij, ki jih zagotavljata drugi dve nalogi. Na tretji sliki vidimo moškega, ki deska na vodi. V tem primeru model JPPNet daje najboljšo segmentacijo glede na anotacije, saj ustrezno označi zgornji del oblačila in ga loči od hlač. Naš model celotno območje uvršča med enodelne kombinezone, kar je glede na videz slike iz človeške perspektive smiselna klasifikacija.

5 Zaključek

V tem delu smo predstavili večciljni segmentacijski model, imenovan SPD. Poleg primarne naloge segmentacije telesa model vključuje tudi nalogo napovedi skeletnih točk in napovedovanja globinske predstavitve telesa. Segmentacijski del modela je bil ovrednoten na podatkovnih zbirkah LIP in ATR, pri obeh zbirkah podatkov pa je SPD dosegel boljše rezultate kot referenčni model JPPNet. Poleg tega je bilo s strogimi študijami ablacije dokazano, da so modeli, ki so upoštevali manjše število nalog, povzročili slabšo učinkovitost.

V analizi ablacije smo predstavili prispevek vsake od nalog in ugotovili, da skupna uporaba skeleta in globinske naloge dodaja večjo vrednost kot uporaba katere koli od njiju samostojno. Za nadaljnje izboljšanje rezultatov načrtujemo raziskovanje dodatnih nalog v učnem postopku, ki bi lahko dale dodatne napotke za postopek segmentacije.

Opomba

To raziskavo so delno podprli projekt ARRS J2-2501 "Globoki generativni modeli za lepoto in modo (DeepBeauty)", raziskovalni program ARRS P2-0250(B) "Meroslovje in biometrični sistemi" in ARRS Research Program P2-0214 "Računalniški vid".

Literatura

- X. Han, Z. Wu, Z. Wu, R. Yu, and L. S. Davis, "Viton: An image-based virtual try-on network," in IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 7543–7552.
- B. Fele, A. Lampe, P. Peer, and V. Štruc, "C-vton: Context-driven image-based virtual try-on network," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), January 2022.
- L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," in Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4. IEEE, Apr. 2018, pp. 834–848. [Online]. Available: <https://doi.org/10.1109/tpami.2017.2699184>
- J. Wang, K. Sun, T. Cheng, B. Jiang, C. Deng, Y. Zhao, D. Liu, Y. Mu, M. Tan, X. Wang, W. Liu, and B. Xiao, "Deep high-resolution representation learning for visual recognition," vol. 43, no. 10. IEEE, Oct. 2021, pp. 3349–3364. [Online]. Available: <https://doi.org/10.1109/tpami.2020.2983686>
- K. Gong, X. Liang, D. Zhang, X. Shen, and L. Lin, "Look into person: Self-supervised structure-sensitive learning and a new benchmark for human parsing," in Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Jul. 2017, pp. 556–567. [Online]. Available: <https://doi.org/10.1109/cvpr.2017.715>
- R. Zhao, W. Ouyang, and X. Wang, "Unsupervised salience learning for person re-identification," in Conference on Computer Vision and Pattern Recognition. IEEE, Jun. 2013, pp. 500–518. [Online]. Available: <https://doi.org/10.1109/cvpr.2013.460>
- X. Liang, S. Liu, X. Shen, J. Yang, L. Liu, J. Dong, L. Lin, and S. Yan, "Deep human parsing with active template regression," in Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 12. IEEE, Dec. 2015, pp. 2402–2414. [Online]. Available: <https://doi.org/10.1109/tpami.2015.2408360>
- X. Liang, X. Shen, D. Xiang, J. Feng, L. Lin, and S. Yan, "Semantic object parsing with local-global long short-term memory," in Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Jun. 2016, pp. 710–724. [Online]. Available: <https://doi.org/10.1109/cvpr.2016.347>
- I. Kokkinos, "UberNet: Training a universal convolutional neural network for low-, mid-, and high-level vision using diverse datasets and limited memory," in Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Jul. 2017, pp. 1380–1410. [Online]. Available: <https://doi.org/10.1109/cvpr.2017.579>
- D. Eigen and R. Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture," in International Conference on Computer Vision (ICCV). IEEE, Dec. 2015, pp. 800–820. [Online]. Available: <https://doi.org/10.1109/iccv.2015.304>
- B. Bischke, P. Helber, J. Folz, D. Borth, and A. Dengel, "Multi-task learning for segmentation of building footprints with deep neural networks," in International Conference on Image Processing (ICIP). IEEE, Sep. 2019, pp. 630–647. [Online]. Available: <https://doi.org/10.1109/icip.2019.8803050>
- X. Liang, K. Gong, X. Shen, and L. Lin, "Look into person: Joint body parsing & pose estimation network and a new benchmark," in Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 4. IEEE, Apr. 2019, pp. 871–885. [Online]. Available: <https://doi.org/10.1109/tpami.2018.2820063>
- R. A. Guler, N. Neverova, and I. Kokkinos, "DensePose: Dense human pose estimation in the wild," in CVF Conference on Computer Vision and Pattern Recognition. IEEE, Jun. 2018, pp. 1120–1135. [Online]. Available: <https://doi.org/10.1109/cvpr.2018.00762>

- K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
- T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in Computer Vision – ECCV. Springer International Publishing, 2014, pp. 740–755. [Online]. Available: https://doi.org/10.1007/978-3-319-10602-1_48
- P. Rot, M. Vitek, K. Grm, Z. Emeršič, P. Peer, and V. Štruc, "Deep sclera segmentation and recognition," in Handbook of vascular biometrics. Springer, Cham, 2020, pp. 395–432.
- Z. Emeršič, D. Sušanj, B. Meden, P. Peer, and V. Štruc, "ContextedNet: Context-aware ear detection in unconstrained settings," IEEE Access, vol. 9, pp. 145 175–145 190, 2021.

LOČEVANJE ŠIROKO IN OZKO LISTNIH RASTLIN V PODPORO SISTEMU ZA ZAZNAVO PLEVELA

URBAN KENDA¹ IN JURIJ RAKUN²

¹ Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, Slovenija.

E-pošta: urban.kenda@student.um.si

² Univerza v Mariboru, Fakulteta za kmetijstvo in biosistemske vede, Hoče, Slovenija.

E-pošta: juriy.rakun@um.si

Povzetek Sodobno kmetijstvo se srečuje z vedno višjo stopnjo avtomatizacije, katere cilj je pridelati več pridelka, ki dosega višjo kakovosti, vse to ob manjših negativnih učinkih na okolje in s potencialom ohranitve narave za naslednje generacije. V ta namen smo v sklopu študentskega projekta razvoja avtonomnega kmetijskega robota Farmbeast ustvarili sistem za selektivno škropljenje plevela, katerega del sta dva algoritma za ločevanje ozko- in širokolistnih rastlin, nujna za ločevanje plevela od pridelka. Ob uporabi testnega nabora slik je prvi algoritem uspešno zaznal širokolistni plevel s 53,3 % in ozkolistni s 93,3 % uspešnostjo medtem, ko drugi algoritem obe sorti plevela uspešno loči v 93,3 %. Delo opisuje tudi orodje, ki je bilo v sklopu projekta razvito in deluje na podlagi algoritma ter škropi s potrebnim herbicidom.

Ključne besede:

strojni
vid,
plevel,
škropljenje,
segmentacija,
robot,
ROS

SEPARATION OF BROAD AND NARROW-LEAF PLANTS IN SUPPORT OF WEED DETECTION SYSTEM

URBAN KENDA¹ & JURIJ RAKUN²

¹ University of Maribo, Faculty of Electrical Engineering, Computer Science and Informatics, Maribor, Slovenia.

E-mail: urban.kenda@student.um.si

² University of Maribor, Faculty of Agriculture and Life Sciences, Hoče, Slovenia.

E-mail: jurij.rakun@um.si

Abstract Modern agriculture is facing an ever-increasing level of automation, which aims to produce crops in higher quantity and higher quality. All this but causing fewer negative effects on the environment and with the potential to preserve the nature for future generations. For this purpose, we created a system for selective weeding, which was developed as part of the student project Farmbeast, where we focus our efforts in the development of an autonomous agricultural robot. The system includes two algorithms for separating narrow- and wide-leaf plants, necessary for separating weeds from crops. By using a test set of images, the first algorithm successfully detected wide-leaf weeds with 53.3% and narrow-leaved weeds with 93.3% effectiveness, while the second algorithm successfully separated the two weed sorts with 93.3% success rate. This work also describes the tools that were developed as part of the project and works on the basis of an algorithm, which sprays with the necessary kind of the herbicide.

Keywords:

machine
vision,
weeding,
spraying,
segmentation,
robot,
ROS

1 Uvod

V sodobnem času kmetijstvo vedno več pozornosti namenja ohranjanju okolja in narave. Eden izmed večjih dejavnikov onesnaževanja je uporaba fito-farmaceutskih sredstev (FFS), saj trenutna mehanizacija v večini primerov ne preverja doziranje FFS, to pa ima tudi negativne posledice, saj se s tem zastrupljajo tla, posledično pa tudi voda in pridelki, ki jih gojimo na teh področjih. V ta namen se razvija nova generacija kmetijstva – precizno kmetijstvo, ki opisuje vse metode in tehnologije, s katerimi kmetovalec strmi k natančnejšem oziroma preciznejšem obdelovanju polj. [1]

Vsekakor se bo nova generacija kmetijstva morala robotizirati, s čimer bodo roboti nadomestili delo človeka, sodobni kmet pa bo vse skupaj nadzoroval z enega mesta. Na trgu popolnih rešitev še ni, vendar pa že poznamo nekaj podjetji, ki nekako »orjejo ledino« po tem novem področju; eno izmed njih je podjetje Naio Technologies, ki so s svojimi tremi namenskim roboti nekako vodilni na tem področju [2]. Drugo, Ecorobotix s svojim avtonomnim robotom namenjenim za zatiranje plevela trdi, da robot uporablja do 90 % manj škropiva saj škropi le zaznan plevel, hkrati pa da manj kot 5 % rastlin pridelka zazna kot plevel [3]. Podobno trdi tudi podjetje Blue River Technology, ki je svojo rešitev prodalo vodilnemu proizvajalcu kmetijske mehanizacije John Deere - ju, kar za 305 milijonov ameriških dolarjev [4]. Slednji so sistem za prepoznavo in škropljenje, naredili kot priključek za obstoječe kmetijske stroje – traktorje.

Tudi naša skupina, ki deluje v sklopu projekta Farmbeast se zaveda problemov povezanih s kmetijstvom, zato je nastala želja, po razvoju sistema, ki omogoča selektivno škropljenje plevela, s čimer lahko zmanjšamo vpliv škropljenja na okolje. Del tega sistema je sistem ločevanja ozko- in širokolistnih rastlin, na podlagi katerega lahko uspešno ločimo del neželenih rastlin nasprotne vrste, torej ves ozkolistni plevel med širokolistnim pridelkom ali širokolisten plevel med ozkolistnim pridelkom. Drugi del pa predstavlja samo orodje, ki omogoča selektiven in točkovni nanos dveh različnih vrst FFS, za eno ali za drugo vrsto. Oboje je opisano v nadaljevanju.

2 Metode dela in uporabljen material

V nadaljevanju so opisane metode in material, uporabljene pri razvoju opisane rešitve in zajemajo tako strojni, kot tudi programski del.

2.1 Strojna oprema

Strojna oprema vključuje sprednji razširitveni priključek robota Farmbeast. Farmbeast [5] je avtonomen poljedelski robot, v osnovi namenjen za delo med vrstami pridelka. Glavni senzor za zaznavanje okolice je Velodyne Puck VLP-16 LiDAR [6], ki deluje kot oči robota. Robot se premika s štirimi pogonskimi kolesi, ki imajo še zmožnost obračanja po z osi, kar omogoči robotu zasuk na mestu, brez zdrsa koles.

2.1.1 Farmbeast orodje za odstranjevanje plevela – greben

Orodje za odstranjevanje plevela – greben je nameščen na spredjem delu robota, kot prikazuje slika 1, in tako omogoča odstranjevanje plevela pred robotom. V osnovi vsebuje dva modularna priključka: nitkasti rezalnik in laser, ki uničujeta oziroma odstranjujeta plevel pred robotom. V kolikor gre za mlado rastlino, jo lahko namreč termično odstranimo že z usmerjenim laserskim žarkom, starejšo pa lahko odrežemo. Po kemičnih pripravkih pa navadno posežemo, če gre za trdovratno vrsto plevela.

Greben in orodja so zasnovana tako, da prijemalo grebena lahko brez posredovanja človeka menja orodje. Dodatno omogoča pomik modulov po dveh oseh, s katerim se dodatno približa oziroma prestavi na pravilno razdaljo plevelu. Celoten greben krmili mikrokrmilnik Arduino Mega [7].



Specifikacije robota	
Teža:	45 Kg
Velikost	550x400x400
DxŠxV:	mm
Pogon:	4 BLDC motorji
Zaznavanje okolice:	Velodyne Puck VLP-16 LiDAR
Računalniki:	Intel NUC in Raspberry Pi 3B+
Napajanje:	LiFePO4, 26,4 V, 12 Ah

Specifikacije grebena	
Teža:	17 Kg
Velikost	250x450x350
DxŠxV:	mm
Pogon:	Koračni motorji
Zaznavanje okolice:	PI kamera
Računalniki:	Raspberry Pi 3B+ in Arduino Mega
Napajanje:	12V, iz robota

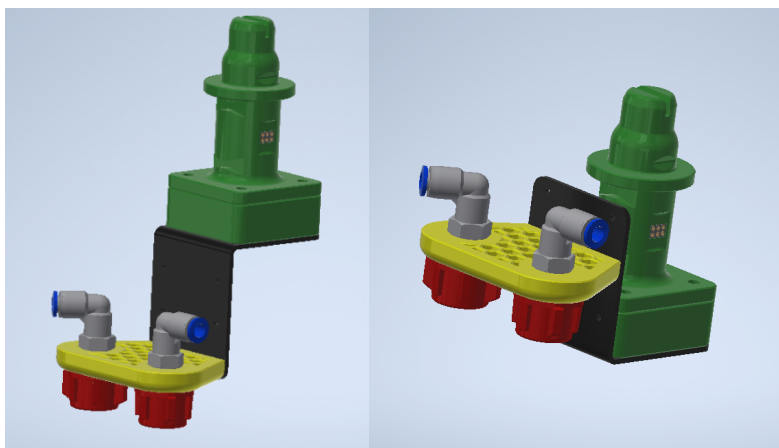
Slika 1: Robot Farmbeast s sprednjim priključkom

Vir: lasten

Na greben je nameščen računalnik Raspberry Pi 3B+ (RPI) [8], pi kamera, dve posodi za dva različna herbicida in dva ventila. Pi kamera je nameščena spredaj, na vrhu grebena in obrnjena tako, da opazuje področje pred grebenom. Zajem in obdelava slik se procesira na RPI računalniku, ki prav tako skrbi za vodenje grebena in proženje dodatnih ventilov.

2.1.2 Škropilno orodje

Orodje je zasnovano na principu predhodnih avtomatskih zamenljivih orodjih, dodatna jeklena konzola pa omogoča premik šob po višini, kar je še posebej koristno v fazi testiranja. Na konzolo je pritrjen 3D tiskani nosilec, na katerega sta nameščeni dve šobi, vsaka za eno vrsto herbicida. Dve varianti konfiguracije prikazuje slika 2.



Slika 2: Možne konfiguracije škropilnega orodja

Vir: lasten

Na škropljenje vpliva tudi pravilna izbira šob. V primeru napačno določene šobe lahko dosežemo preveliko ali premajhno doziranje škropiva, kar lahko povzroči preveliko osnaževanje ali pa se lahko s premajhnim doziranjem učinkovitost škropiva slabša. Izbrana šoba je tipa 468.526.5E podjetja Lechler [9], ki ima stožčast curek in kot škropljenja 90°. S takšno šobo lahko že pri višini 75 mm poškopimo krog premera 150 mm. Pri delavnem tlaku 3 bar šoba doseže pretok 2,35 l/min, s čimer s trenutnim volumnom posode, le to izpraznimo po 23 škropljenjih če le-ti trajajo 1 s. Razvit sistem torej služi v podporo demonstracije načina delovanja in potrditvi koncepta.

2.2 Programska oprema

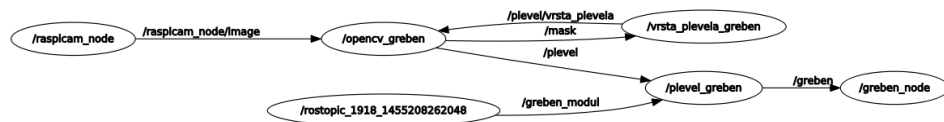
Sistem za zaznavo potencialnega plevela je v celoti programsko ločen od robota, zaradi česa lahko deluje neodvisno. V nadaljevanju je opisana uporaba funkcij iz knjižnice OpenCV [10], ki temeljijo na principu enostavne segmentacije in njene implementacije v ROS – u [11].

2.2.1 ROS »Robot Operating System«

Robotski operacijski sistem (ROS) je zbirka orodij, knjižnic in pretvornik z namenom olajšanja ustvarjanja zapletenih robotskih nalog, prav tako pa poenostavi komunikacijo med različnimi vozlišči.

ROS komunikacija omogoča vozliščem prejemanje in pošiljanje podatkov, tudi če so napisani v različnih programskih jezikih. Izmenjava podatkov temelji na objavljanju sporočil na teme, ki delujejo kot nekakšen kanal po katerih vozlišča objavljajo ali poslušajo sporočila. Eno vozlišče lahko hkrati objavlja in posluša več različnih tem, pomembno je le to, da se na temo objavlja oziroma v primeru poslušanja v naprej določi tip sporočila. Ker je objavljanje in naročevane na temo anonimno, vozlišča ne vedo, kdo je poslal oziroma prej sporočilo. V ta namen lahko uporabimo storitve, ki se uporabljajo za dejanja, ki imajo v naprej določen začetek in hkrati konec, vendar jih v naši raziskavi nismo potrebovali. Omenjeni ROS elementi so del ROS paketa, poleg teh se še v paketu lahko nahajajo: header datoteke, zagonske datoteke (ang. »launch files«), CMakeList datoteka in »Pacakage« datoteka. ROS paket strmi k širši funkcionalnosti, da postane uporaben za več projektov, hkrati pa ni preobširen in težak za porabo.

Za raziskavo smo uporabili štiri ROS pakete: raspicam, roserial [12], opencv in greben. Le paket greben je ustvarjen, medtem ko so ostali trije preneseni in po potrebi spremenjeni oziroma dopolnjeni. Oblika ROS grafa, ki prikazuje slika 3, prikazuje povezave med vozlišči in pot sporočil po temah.



Slika 3: ROS graf z vsemi vozlišči (elipse) in temami znotraj katerih se objavljajo podatki

Vir: lasten

V nadaljevanju bodo predstavljeni paketi in njihova vozlišča.

Paket raspicam omogoča zajem slike s pi kamere, vsebuje le eno vozlišče /raspicam_node, ki mu lahko nastavimo frekvenco, velikost in resolucijo zajemanja slik. Vozlišče zajeto sliko objavlja na temo /raspicam_node/Image v sporočilu tipa Image.msg.

V opencv paket smo dodali dve novi vozlišči /opencv_greben in pa /vrsta_plevela_greben. Prvo vozlišče po temi /raspicam_node/Image pridobi sliko in jo z upragovanjem in filtriranjem pretvori v masko, ki jo objavlja na temo /mask. Na slednjo temo je naročeno vozlišče /vrsta_plevela_greben, ki ločuje plevel po sorti, na podlagi prejete maske, hkrati pa podatek o sorti in velikosti plevela objavlja na temo /plevel/vrsta_plevela. Na katero je naročeno vozlišče /opencv_greben, ta na podlagi središča, velikosti in sorte pleveli objavlja na temo /plevel sporočilo s podatki o nahajanju plevela po x koordinati (širini), vrsti in velikosti plevela.

Paket greben vsebuje vozlišče /plevel_greben, zagonsko datoteko s katero zaganjamo vseh šest vozlišč iz štirih paketov. Vozlišče je naročeno na temo /plevel, hkrati pa na temo /greben_modul, preko katere lahko uporabnik določi željen modul. Nato vse prejete podatke pretvori v obliko primerno za sporočilo greben.msg ki se objavlja na temo /greben.

Namen Rosserial paketa je, da pretvori ROS sporočila v serijska sporočila namenjena serijski komunikaciji. V našem primeru vozlišče /greben_node, vzpostavi serijsko komunikacijo s mikrokrmilnikom Arduino, ki je nameščen na grebenu.

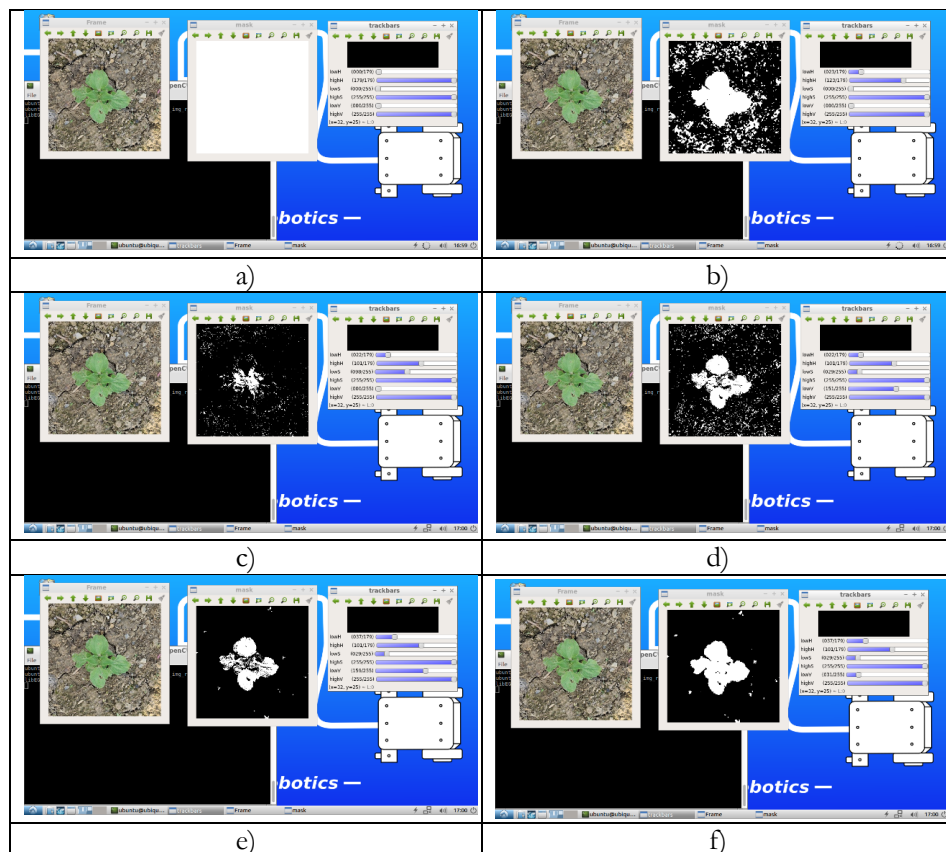
2.2.2 Delo z OpenCV knjižnicami

Vse obdelave slik in posnetkov so izvedene s pomočjo OpenCV knjižnice (»open source computer vision library«). V nadaljevanju pa so opisane osnovne operacije, ki so potrebne za prepoznavo in segmentacijo slik, ki loči koristne rastline od plevela.

2.2.2.1 HSV upravljanje

V algoritmu za zaznavo plevela najprej ustvarimo binarno matriko, na podlagi pragov v HSV prostoru. Pragovi so bili določeni empirično, kot ga opisuje naslednji odstavek.

Na začetku so pragovi nastavljeni med minimalnimi in maksimalnimi vrednostmi, zato zaznavamo celotno področje, kot to prikazuje sliki 4 (a). Slika 4 (b) prikazuje naslednji korak, kjer so pragovne vrednosti nastavljene tako, da zajemajo odtenke (»Hue«), ki določajo slikovne elemente z zeleno barvo. Slika 4 (c) in (d) prikazujeta vpliv barvne nasičenosti (»Saturation«), medtem ko sliki 4 (e) in (f) prikazujeta spreminjanje svetlosti (»Value«). Hkrati lahko med zaznavami opazimo, da se šum zmanjšuje, s čimer si dobro nastavimo pogoje za nadaljnjo filtriranje. Po končanem preizkušanju smo z območji: Hue 37 – 101, Saturation 29 – 255 in Value 31 – 255 dobili optimalno masko, z malo vsebovanega šuma.



Slika 4: Zaznava vsega (a), približno nastavljen prag barvnega odtenka (b), območje barvne nasičenosti premajhno (c), zmanjšanje območja svetlost (d), natančneje določen prag barvnega odtenka (e), povečanje območja svetlosti (f)

Vir: lasten

2.2.2.2 Morfološke operacije

Po HSV upravljanju sledijo morfološke operacije, ki se uporabljajo pri obdelavi slik in temeljijo na logičnih operatorjih [1]. Ti omogočajo, da na slikah izluščimo iskane informacije o lastnosti rastlin. Morfološke operacije potrebujejo dva vhodna elementa: sliko in pa strukturni element oziroma t.i. »kernel«. Pri našem preizkušanju sta bili uporabljeni dve osnovni morfološki operaciji: erozija in dilatacija, njuno zaporedje pa tvori dve drugi operaciji: odpiranje in zapiranje.

Erozija, kot že samo ime pove, odnaša določene slikovne elemente. Pri binarnih slikah se to odraža, da kernel, ki potuje skozi sliko spreminja vrednost iz 1 v 0, kadar niso vsi členi kernela enaki 1. Tako zaznanemu objektu odnaša robove.

Dilatacija je obratna operacija od erozije, ki namesto odnašanja robov, same robove razširi. Pri tej operaciji kernel prav tako potuje skozi sliko vendar spremeni vrednost iz 0 v 1 kadar je že samo en člen enak 1.

Operaciji odpiranja in zapiranja se običajno uporabita, kadar v binarnih posnetkih odpravljamo šum. Operaciji določimo tako, da si erozija in dilatacija sledita zaporedoma. Ob zaporedju erozija, dilatacija se tvori odpiranje, v obratnem vrstnem redu pa zapiranje.

V algoritmu so uporabljeni operaciji odpiranja in zapiranja, ki poskrbita da se šum v okolju odstrani, hkrati pa luknje v zaznanem plevelu zapolnijo.

2.2.3 Segmentacija plevela

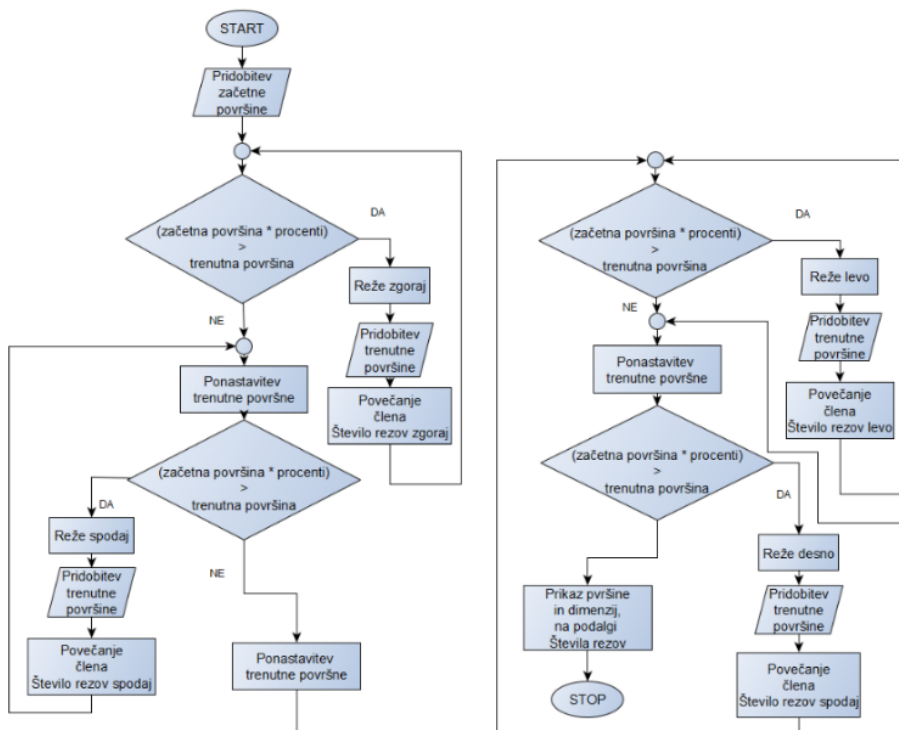
Sledi ločevanje plevela na široko- in ozko-listni plevel. Razvili smo dva algoritma za ločevanje: prvi primerja velikost plevela glede na zaznano površino, drugi pa zmanjšuje površino z izvajanjem erozije.

Oba algoritma sta preprosta, katerih namen je prikaz ideje oziroma enostavnega principa segmentacije, pri čemer se uporabljajo osnovne funkcije, ki niso tako zahtevane. V primeru obeh algoritmov sledimo ideji, da širokolistne rastline proizvedejo večjo količino grupiranih slikovnih elementov, ki zahteva več iteracij obdelave, kar pa pove za kakšen tip rastline gre. V primeru ozkolistne poljščine lahko tako odkrijemo širokolistne plevela ali ravno obratno.

2.2.3.1 Prvi način ločevanja

Pri prvem načinu se primerja velikost plevela z njegovo površino, razmerje med njima pa nam poda vrednost s katere lahko sklepamo, da bo ozkolistni plevel imel manjšo napram širokolistnemu, saj pri isti velikosti plevela bo zaznane površine manj.

Ker oblika plevela, ni vedno simetrična, smo se odločili, da velikost pridobimo z rezanjem slike dokler zaznana površina ni manjša za delež celote vseh potencialnih slikovnih elementov. Tako smo izločili del slike, na katerem rastline ni. To smo izvedli štiri krat za vsako stran posebej, kot prikazuje diagram poteka na sliki 5. Nato smo na podlagi števila rezov z vsake strani, dobili približno velikost plevela, ki pa smo jo na koncu skupaj z zaznano površino izpisali.



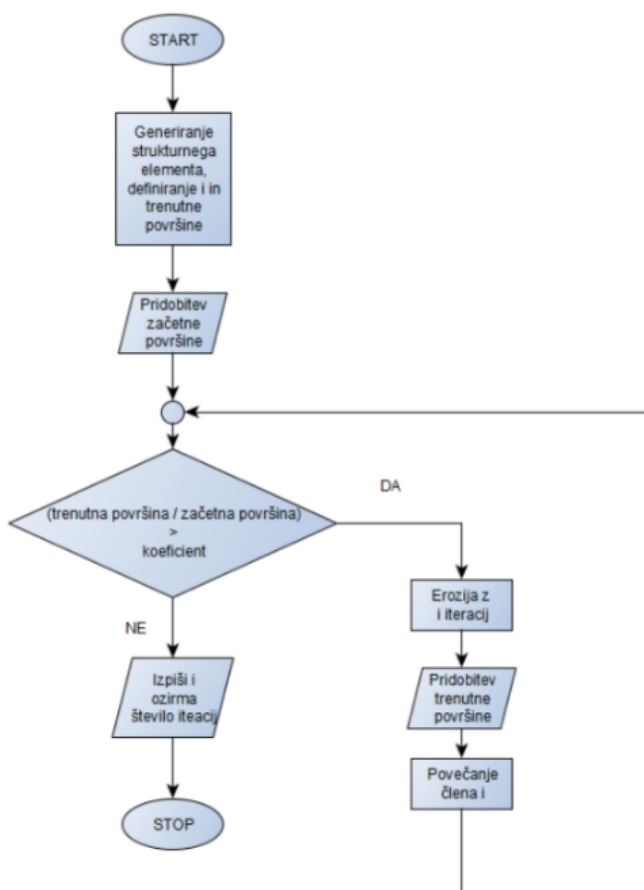
Slika 5: Diagram poteka prvega načina ločevanja

Vir: lasten

2.2.3.2 Drugi način ločevanja

Pri drugem načinu pa je ločevanje temeljilo na ideji, da pri isti površini obeh sort plevela, širokolistni potrebuje več iteracij erozije, da se zaznana površina zmanjša na določeno vrednost.

Algoritem izvaja erozijo, kjer po vsaki iteraciji preveri, če je trenutna površina ulomljena z začetno, manjša od določenega koeficienta oziroma procenta. Na koncu se število potrebnih iteracij izpiše, s čimer bomo kasneje določali vrsto plevela. Diagram poteka je prikazan na sliki 6.



Slika 6: Diagram poteka drugega načina ločevanja

Vir: lasten

Naknadno smo dodali korekcijski faktor, z namenom zmanjšati raztros vrednosti in s tem povečati uspešnost zaznave. Ta ob primeru plevela, ki ima znatno večjo ali manjšo površino, najprej primer plevela sorazmerno zmanjša oziroma poveča in šele nato začne izvajati zgoraj opisan algoritem.

3 Rezultati in ugotovitve

Postopka smo ovrednotili na podlagi 30 naključno izbranih primerov plevela, polovico ozko in drugo polovico široko listnega. Ker gre za prvo fazo raziskovanja tematike zaznave in prepoznave plevela, je potrebno poudariti, da so primeri bili izbrani takšni, da ne prihaja do prekrivanj plevela, hkrati pa, da so lahko prisotne rastline velikosti v rangi 1 cm, kot prikazuje slika 7. Parametri in rezultati primerov plevela so prikazani na tabelah 1 in 2.



Slika 7: Primer plevela, s prisotnostjo manjših rastlin

Vir: lasten

3.1 Rezultati prvega načina ločevanja

Prvi algoritem ločuje plevel na podlagi koeficienta dobljenega iz zaznane površine deljene s povprečjem obeh velikosti maske. Pri tem lahko na uspešnost segmentacije vplivamo s spreminjanjem vrednosti do koliko odstotkov prvotne površine algoritem reže masko.

Najboljši pogoji so doseženi pri rezanju do 98 % prvotne površine, kjer se v povprečju koeficienti ene sorte plevela razlikujejo za 36,5 % od druge sorte. Sam raztros vrednosti pa je velik in sicer, kar 7 primerov širokolistnega plevela ima vrednost koeficienta v območju ozkolistnega plevela. Kar na podlagi danih primerov znaša 53,3 % uspešnost segmentacije široko in 93,3 % ozko – listnega plevela.

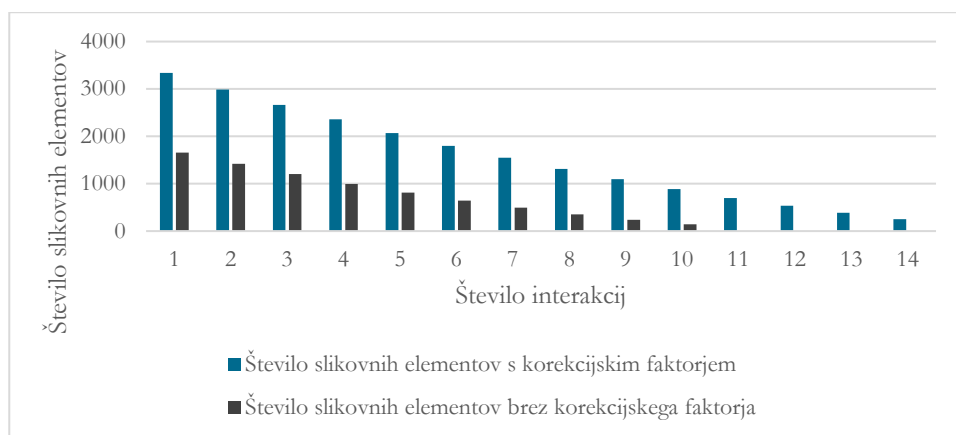
3.2 Rezultati drugega načina ločevanja

Drugi algoritem ločuje plevel na podlagi potrebnega števila iteracij erozije, da dosežemo določeno vrednost začetne zaznave. Pri čemer s spreminjanjem vrednosti začetne zaznave vplivamo na kvaliteto segmentacije.

Ugotovljeno je bilo, da z izvajanjem erozije do 10 % prvotne površine, ozkolistni plevel potrebuje 50 % manj iteracij v primeru širokolistnega. Na podlagi dobljenih rezultatov, prikazanih v tabelah 3.1 in 3.2, se je določila meja pri kateri se loči plevel, ta znaša 14 iteracij. Tako plevel z manj iteracij spada pod ozkolistne medtem, ko z večjim številom spada pod širokolistne. Opazimo lahko, da je raztros vrednosti še posebej velik pri širokolistnem plevelu, vendar pa so te vrednosti raztrošene na dobro stran. V osnovi z drugim načinom ločimo ozkolistni plevel 93,3 % in širokolistni 86,7 % uspešno.

Po vpeljanem korekcijskem faktorju, ki vpliva le na tiste primere, kjer je zaznana površina nad 11 tisoč 750 ali pod 3 tisoč zaznanih slikovnih elementov, se je raztros vrednosti zmanjšal, hkrati pa se je tudi izboljšala segmentacija in sicer sta bili obe vrsti plevela ločeni uspešno v 93,3 %.

Grafikon 1 prikazuje število iteracije 6. primera širokolistnega plevela pred in po korekciji.



Grafikon 1: Prikaz spreminjanja števila slikovnih elementov v odvisnosti od števila interakcij erozije

Vir: lasten

Tabela 1: Rezultati ločevanja ozkolistnega plevela, z rdečo barvo so prikazani napačno določeni primeri

Ozko listnata plevel		Ločevanje po prvem načinu z rezanjem do 98 % prvotne površine				Ločevanje na drugi način z izvajanjem erozije do 10 % prvotne površine	
		Velikost slike po rezanju		Povprežje velikosti	Koeficient	Brez korekcijskega faktorja	S korekcijskim faktorjem
Št.	Št. slikovnih elementov	x	y				
1	5561	224	181	202.5	2.75	9	9
2	2280	58	99	78.5	2.90	10	10
3	14552	193	190	191.5	7.60	18	16
4	4395	93	137	115	3.82	8	8
5	3801	98	110	104	3.65	12	12
6	8689	269	183	226	3.84	9	9
7	1143	293	133	213	0.54	7	13
8	8370	143	209	176	4.76	8	8
9	7876	165	206	185.5	4.25	9	9
10	7288	134	198	166	4.39	8	8
11	4235	187	92	139.5	3.04	13	13
12	10030	191	243	217	4.62	13	13
13	9051	165	239	202	4.48	9	9
14	7872	192	185	188.5	4.18	9	9
15	3381	130	106	118	2.87	7	7
				Povprežje vrednosti	3.8452897	9.933333333	10.2
				Standardni odklon	1.492506	2.939063667	2.569046516

Tabela 2: Rezultati ločevanja širokolistnega plevela, z rdečo barvo so prikazani napačno določeni primeri

Široko listnata plevel		Ločevanje po prvem načinu z rezanjem do 98 % prvotne površine				Ločevanje na drugi način do 10 % prvotne površine	
		Velikost slike po rezanju		Povprežje velikosti	Koeficient	Brez korekcijskega faktorja	S korekcijskim faktorjem
Št.	Št. slikovnih elementov	x	y				
1	2323	67	48	57.5	4.04	15	15
2	3803	86	76	81	4.70	15	15
3	6164	178	205	191.5	3.22	18	18
4	14629	165	167	166	8.81	16	14
5	8558	99	117	108	7.92	20	20
6	1658	67	133	100	1.66	10	14
7	6545	100	97	98.5	6.64	20	20
8	8681	209	253	231	3.76	18	18
9	11196	126	137	131.5	8.51	29	29
10	1273	69	117	93	1.37	9	13
11	7076	183	171	177	4.00	19	19
12	12462	132	169	150.5	8.28	19	19
13	21930	179	161	170	12.90	46	27
14	14368	199	228	213.5	6.73	24	22
15	11005	116	150	133	8.27	17	17
				Povprežje razmerji	6.0544311	19.66666667	18.66666667
				Standardni odklon	3.1559071	8.772250621	4.608480795

4 Zaključek

Del robotiziranih pripomočkov za odstranjevanje plevela zahteva ustrezno programsko podporo, ki napravi sisteme dejansko uporabne. V danem članku tako predstavljamo strojno rešitev in programsko rešitev za zaznavo in odstranjevanje plevela. Programska rešitev vključuje dva različna načina ločevanja koristnih rastlin od plevela. Ta sta bila preizkušena ob uporabi naključnega vzorca posnetkov, kjer pa je potrebno izpostaviti, da so bili vzorci zajeti ob optimalnih pogojih, brez prekrivanja drugih rastlin in takšni, ki prikazujejo pretežno cele rastline. Postopek zaradi tega še ni popolnoma primeren za uporabo v realnih pogojih in predstavlja izhodišče za nadaljnje raziskave.

Povzetek rezultatov kaže, da se drugi algoritem v primerjavi s prvim obnese veliko bolje, saj prvi algoritem, kljub dobri prepoznavi ozkolistnega plevela, skoraj pol širokolistnega plevela določi napačno. Drugi algoritem pa v več kot v 90 % določi uspešno obe sorti plevela, kar kaže na uspešnost delovanja algoritma in daje vzpodbudne rezultate za nadaljnje delo.

Za naslednje korake v smeri izboljšave sistema, ki bi rešitev napravil še bolj natančno in robustno, predlagamo uporabo bolj naprednih metod, kot so opazovanje vzorcev tekstur v slikovnem ali frekvenčnem prostoru, analizo oblik oz. uporabo različnih transformacij, kot je Houghove transformacija za krožnice, s katero bi poudarili delno okroglost širokolistnih plevelov ali pa mogoče uporabo klasifikacije s še naprednejšimi metodami, kot so nevronske mreže.

Literatura

- [1] Farmdok. Preceizno kmetijstvo. Dostopno na : <https://www.farmdok.com/sl/2017/05/27/precizno-kmetijstvo-farmdok-kmetijstvo/> [20.1.2022]
- [2] Naio Technologies. Dostopno na: <https://www.naio-technologies.com/en/home/> [21.1.2022]
- [3] Ecorobotix. AVO. Dostopno na: <https://ecorobotix.com/en/avo/> [25.1.2022]
- [4] Medium, DCVC. John Deere acquires Blue River Technology for \$305 million, bringing full stack AI to agriculture. Dostopno na: <https://medium.com/@dcvc/john-deere-acquires-blue-river-technology-for-305-million-bringing-full-stack-ai-to-agriculture-7ca8c25a5fe1> [25.1.2022]
- [5] Farmbest. Dostopno na: <https://farmbeast.um.si> [5.1.2022]
- [6] Velodyne lidar. Puck. Dostopno na: <https://www.velodynelidar.com/products/puck/> [19.1.2022]
- [7] Michael Margolis, Arduino Cookbook, druga izdaja. O'Reilly Media, Inc, USA, 2012.

- [8] Derek Molloy, *Exploring Raspberry Pi: Interfacing to the Real World with Embedded Linux*, prva izdaja. Wiley 2016.
- [9] Lechler. Axial-flow full cone nozzles for retaining nut, Series 468. Dostopno na: https://www.lechler.com/fileadmin/media/kataloge/pdfs/industrie/katalog/EN/03_vollkegel/lechler_full_cone_nozzles_series_468.pdf [14.1.2022]
- [10] Robert Laganier. *OpenCV 3 Computer Vision Application Programming Cookbook*, tretja izdaja. Packt Publishing Limited, 2017.
- [11] Morgan Quigley, Brian Gerkey, William D. Smart. *Programming Robots with ROS: A practical Introduction to the Robot Operating System*, prva izdaja. O'Reilly Media, Inc, USA, 2015.
- [12] Lentin Joseph. *ROS Robotics Projects*. Packt Publishing Limited, 2017.
- [13] Gonzales, Rafael C., Woods, Richard E. *Digital Image Processing*, druga izdaja. Prentice Hall, Inc. Upper Saddle River, New Jersey, 2002.

DETEKCIJA VLITIH NAVOJEV S POMOČJO TERMOVIZIJE

ANŽE ŠVIGELJ IN BORUT BATAGELJ

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Ljubljana, Slovenija.
E-pošta: as9467@student.uni-lj.si., borut.batagelj@fri.uni-lj.si

Povzetek Cilj članka je opis razvoja spletne aplikacije za detekcijo vlitih navojev segrevane deske za sneg s pomočjo termografske kamere v okolju SICK AppStudio, ki omogoča razvoj v programskem jeziku Lua na računalniku SICK SIM4000. V začetku sta predstavljeni področji računalniškega vida in termografskih kamer. V naslednjem koraku je predstavljena oprema in pripadajoče tehnologije, ki smo jih uporabili. V nadaljevanju pa je prikazana implementacija dveh algoritmov za detekcijo navojev: zaznavanje področij in Houghova transformacija kroga ter njuna primerjava.

Ključne besede:

termovizija,
Lua,
SICK
AppStudio,
termografska
kamera,
Houghova
transformacija

MOLDED THREAD DETECTION USING THERMAL IMAGING

ANŽE ŠVIGELJ & BORUT BATAGELJ

University of Ljubljana, Faculty of computer and information science, Ljubljana,
Slovenia.

E-mail: as9467@student.uni-lj.si, borut.batagelj@fri.uni-lj.si

Abstract The aim of this article is to concisely describe the development of a web application for detecting molded threads of a heated snowboard using a thermographic camera in the SICK AppStudio environment, which uses the Lua programming language on the SICK SIM4000 industrial computer. In the beginning, the areas of computer vision and thermographic cameras are described and presented. In the following chapter, we presented the equipment we worked on and the associated technologies we used. We continued with the description of the implementation of two algorithms we used for detecting threads: blob detection and Circle Hough Transform then compared the results.

Keywords:

thermal
imaging,
Lua,
SICK
AppStudio,
thermal
camera,
Hough
transform

1 Uvod

Termografija je dandanes ključnega pomena v industriji, policiji, medicini, vojski in drugje. Pri policijskem delu je lep praktičen primer tega iskanje oseb in predmetov ponoči in ob omejeni vidljivosti. Poleg termografije je danes med nami vse bolj prisoten tudi računalniški vid in se pojavlja na vse več področjih. V industriji nam na primer koristi pri ločevanju na slabe in dobre izdelke ter polizdelke. Računalniški vid je tudi vse bolj prisoten na merilcih hitrosti na različnih avtocestnih odsekih, saj je boljša in cenejša alternativa klasičnim fotoradarskim merilcem in omogoča branje tablic in njihovo shrambo v podatkovno bazo (Czajewski in Iwanowski, 2010).

V članku je predstavljeno iskanje vlitih navojev na deski za sneg (angl. snowboard) s termovizijo, ki je kombinacija zgoraj navedenih pojmov. Predstavljena bo obdelava slik in algoritma za detekcijo, z zajemom iz termografske kamere FLIR A615.

Ciljna aplikacija teče na industrijskem računalniku SICK SIM4000 in je razvita v okolju SICK AppSpace v programskem jeziku Lua (Ierusalimschy, 2016). Deska je ob izdelavi v tovarni navadno premazana s smolo oziroma posebno zmesjo in se zato lukenj z navoji, ki služijo namestitvi vezi ne vidi. Ideja je, da celotno desko enakomerno segrejemo in s termografsko kamero preverimo, kje so ti navoji, saj so luknje narejene iz kovine in se ohlajajo počasneje kot preostali del deske.

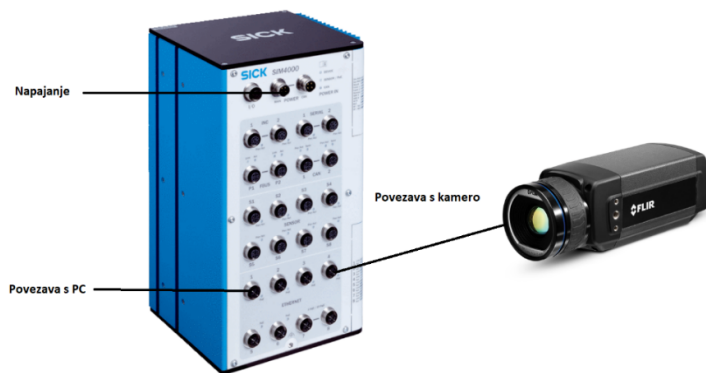
2 Uporabljene tehnologije

– Industrijski računalnik SICK SIM4000

Računalnik SICK SIM4000 je namenjen izdelavi in izvajanjem aplikacij povezanih z računalniškim vidom, ki ga je izdelalo nemško podjetje SICK, ki se v večini ukvarja z izdelavo kamer in senzorjev za področje računalniškega vida ter z razvojem namenskih aplikacij za uporabo njihovih izdelkov (SICK, 2021). Računalnik ima 24 priključkov (Slika 1), ki služijo povezavam s kamerami, senzorji, osebnim računalnikom za namen razvoja in režo za lokalno shranjevanje podatkov na SD kartico.

– Termografska kamera FLIR A615

Za zajem slik smo uporabili termografsko kamero FLIR A615 (Slika 1) z vidnim kotom 25°. Omogoča zajem slik v ločljivosti 640 x 480 slikovnih elementov, ki jih predstavi v 16-bitnem formatu. Naprava ima razpon zaznavanja temperature med 20 °C in 2000 °C z natančnostjo ± 2 °C (± 2 %) na posamezno branje temperature.



Slika 1: Prikaz povezav na industrijski računalnik SICK SIM4000 s povezano termografsko kamero FLIR A615.

Vir: lasten

– Razvojno okolje SICK AppStudio

Razvojno okolje je namenjeno razvijanju in poganjanju aplikacij na računalniku SICK SIM4000. Razvoj poteka v skriptnem jeziku Lua.

Med drugim smo za namene shranjevanja podatkov v bazo uporabljali tudi odprtokodno večplatformsko strežniško okolje Node.js in program za podatkovne baze MongoDB.

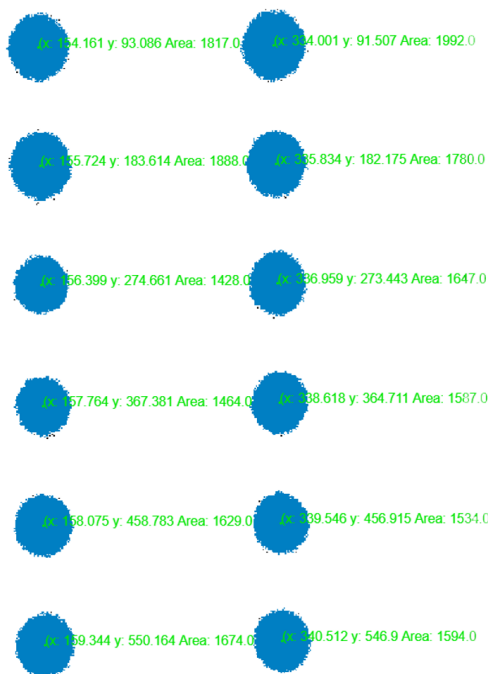
3 Razvoj aplikacije

Aplikacija ima naslednje funkcionalnosti: gumb za vklop in izklop kamere, nastavitve IP naslova kamere, nastavitve mejne vrednosti binarizacije, nastavitve parametra velikosti področja in raztezka posameznega področja, gumb za obdelavo slike, prikaz surove slike, prikaz obdelane slike z vidnimi koordinatami, shranjevanje slike,

obdelava slike z zaznavanjem področij in Houghove transformacije kroga, shranjevanje v bazo in vpogled v bazo.

3.2 Zaznavanje področij

Pri metodi zaznavanja področij (angl. blob detection) posneto sliko najprej pretvorimo v sivinsko in jo potem binariziramo, kar je postopek pretvorbe sivinske slike s 256 odtenki v sliko z dvema barvama, črno in belo, kar izloči nepomembne elemente nad nastavljeno mejno vrednostjo. Iz binarizirane slike nato izvlečemo vsa področja, ki imajo več kot 20 črnih slikovnih elementov skupaj in jih shranimo v tabelo področij. Nad tabelo smo nastavili še filtra površine (angl. area) in raztezka (angl. elongation), kar izloči vse neprimerne objekte iz tabele. Za določitev središča smo v okolju SICK AppSpace iz knjižnice PixelRegion uporabili funkcijo `Image.PixelRegion.getCenterOfGravity(blobs[i], image)`, ki na posameznih področjih določi težišče in vrne koordinate težišča oziroma v našem primeru središče (Slika 2).



Slika 2: Grafični izris področij in koordinat.

Vir: lasten.

3.3 Houghova transformacija kroga

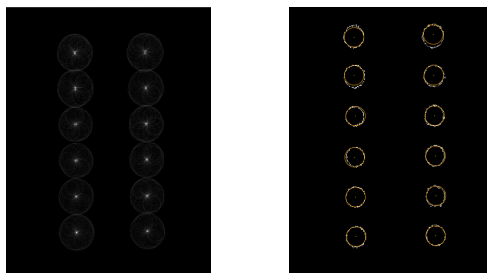
Pri Houghovi transformaciji kroga smo nad binarizirano sliko uporabili Cannyjev detektor robov, s pomočjo katerega dobimo sliko robov, ki z visoko verjetnostjo sovpadajo z robovi na binarizirani sliki. Nato se nad vsakim robnim slikovnim elementom nariše krožnica z želenim polmerom, kjer je robni slikovni element središče kroga, okoli celotnega roba, dokler ne pridemo na isto mesto. Posamezen krog lahko predstavimo v enačbi 1.

$$r^2 = (x - a)^2 + (y - b)^2 \quad (1)$$

V enačbi 1 a in b predstavljata središče kroga, r predstavlja polmer, x in y predstavljata točko na krožnici.

Rezultat je akumulatorsko polje vidno na prvi sliki 3. Ker krožnice v našem primeru niso popolne krivulje smo v akumulatorskem polju določili lokalne maksimume, ki so v našem primeru točke na akumulatorju, kjer je največje število stičišč akumuliranih krožnic oziroma kjer je jakost slikovnih elementov na akumulatorju najvišja. Lokalni maksimumi predstavljajo najboljše kandidate za središča krožnic (OpenCV, 2022). Algoritem lahko zapišemo s psevdokodo v šestih korakih:

1. Najdi robove na sliki s Cannyjevim detektorjem robov
2. Za vsako področje
 3. Za vsako točko na robu
 4. Nariši krožnico s središčem v tej točki na robu s polmerom r
 5. Poišči lokalni maksimum v akumulatorskem polju
6. Najdene parametre (r, a, b) , ki ustrezajo lokalnim maksimumom, preslikaj nazaj na izvirno sliko (Slika 3).



Slika 3: Akumulatorsko polje (levo) in izris najdenih krogov na vhodni sliki robov (desno).

Vir: lasten.

3.4 Primerjava obeh metod

Pri primerjavi obeh metod smo na sliki ročno določili središča za vsak posamezen navoj tako, da smo poljubno izbrali tri točke na posamezni krožnici in s pomočjo njih določili središče. Koordinate središč navojev po vseh treh metodah so vidne v tabeli 1. Na dnu tabele je prikazano še povprečno odstopanje posamezne metode od ročnega izračuna središč. Za izračun odstopanja smo uporabili formulo Pitagorovega izreka za računanje razdalje med dvema točkama v koordinatnem sistemu. Na podlagi podatkov in izračunov lahko vidimo, da je točnost določanja središč pri obeh metodah primerljiva, saj je razlika med njima v povprečju le desetinka slikovnega elementa in od ročno izmerjenih središč v povprečju odstopata le za en slikovni element.

Tabela 1: Primerjava razlik med ročnim izračunom središč in uporabljenima metodama.

št. navoja	ročno		zaznavanje področij	Hough
	x	y	odstopanje	odstopanje
1	115	167	1,6826	2,5838
2	206	165	0,9156	1,6466
3	296	166	2,0807	1,2236
4	388	165	1,5605	1,4040
5	481	164	0,8021	0,9960
6	571	163	0,7066	0,8884
7	113	345	1,3319	1,2606
8	205	345	0,2745	1,2964
9	298	347	2,0796	2,8351
10	387	346	2,1942	2,8317
11	480	346	0,7498	0,3417
12	572	343	1,5816	0,5450
povprečje:			1,3300	1,4877

Literatura

- Czajewski, W., in Iwanowski, M. (2010). Vision-based vehicle speed measurement method. V International Conference on Computer Vision and Graphics (str. 308-315). Springer, Berlin, Heidelberg.
- Ierusalimsky, R. (2016). Programming in Lua, Fourth Edition. Lua.org
- OpenCV, Open Source Computer Vision. (2022). Hough Circle Transform. https://docs.opencv.org/3.4.15/d4/d70/tutorial_hough_circle.html
- SICK, Sensor Intelligence. (3. 8. 2021) SIM4x00, Flexible. Intelligent. Communicative. https://cdn.sick.com/media/familyoverview/2/52/452/familyOverview_SIM4x00_g386452_en.pdf

OCENJEVANJE STAROSTI OSEBE NA OSNOVI DIGITALNIH POSNETKOV Z UPORABO MODIFICIRANE MREŽE VGG-FACE

TILEN KREL¹ IN BOŽIDAR POTOČNIK²

¹ Krelware s.p., Selnica ob Dravi, Slovenija.

E-pošta: tilen.krel@krelware.com.

² Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, Slovenija.

E-pošta: bozidar.potocnik@um.si.

Povzetek V članku predstavimo model konvolucijske nevronske mreže za ocenjevanje starosti osebe iz digitalnega posnetka. Kot osnova za naš model je bila uporabljena in modificirana obstoječa arhitektura konvolucijske nevronske mreže VGG-Face, namenjena razpoznavanju obrazov. Za učenje in testiranje sta bili uporabljeni bazi podatkov IMDB-WIKI in FG-NET. Na bazi podatkov IMDB-WIKI je bila dosežena povprečna napaka med dejansko in ocenjeno starostjo 6,7 leta, na bazi podatkov FG-NET pa je z validacijsko metodo »izpusti-eno-osebo« bila izračunana povprečna napaka med dejansko in ocenjeno starostjo 3,9 leta. Dobljeni rezultati so primerljivi oziroma le malo zaostajajo za najuspešnejšimi metodami za ocenjevanje starosti osebe iz digitalnega posnetka. Na tej osnovi naš model ocenjujemo kot primeren za uporabo v produkcijskih rešitvah.

**Prispevek
temelji na:**

Krel, T. (2021).
Ocenjevanje
starosti osebe na
osnovi digitalnih
posnetkov z
uporabo
konvolucijskih
nevronske mreže:
magistrsko delo,
Univerza v
Mariboru,
Fakulteta za
elektrotehniko,
računalništvo in
informatiko.
Maribor.

Ključne besede:

računalniški
vid,
konvolucijske
nevronske
mreže,
globoko
učenje,
ocenjevanje
starosti,
mreža
VGG-Face

PERSON AGE ESTIMATION BASED ON DIGITAL IMAGES USING MODIFIED VGG-FACE NETWORK

The paper is
based on:

Krel, T. (2021).

Ocenjevanje
starosti osebe na
osnovi digitalnih
posnetkov z
uporabo
konvolucijskih
nevronske mreže:
master's thesis,
University of
Maribor, Faculty
of Electrical
Engineering and
Computer Science.
Maribor.

Keywords:

computer
vision,
convolutional
neural
networks,
deep
learning,
age
estimation
VGG-Face
net

TILEN KREL¹ & BOŽIDAR POTOČNIK²

¹ Krelware s.p., Selnica ob Dravi, Slovenia.

E-mail: tilen.krel@krelware.com

² University of Maribor, Faculty of Electrical Engineering and Computer Science,
Maribor, Slovenia.

E-mail: bozidar.potocnik@um.si

Abstract In this article we present a convolutional neural network model for a person age estimation from an image. As a base for our model, we used and modified the existing convolutional neural network architecture VGG-Face, used for face recognition. For learning and testing, the IMDB-WIKI and FG-NET datasets were used. With the IMDB-WIKI dataset, we achieved the average error 6.7 years between the actual and the estimated age, while using the FG-NET dataset, we calculated the average error 3.9 years between the actual and the estimated age, whereas the »leave-one-person-out« validation method was employed. The obtained results are comparable to or only slightly behind the state-of-the-art methods for the age estimation from a digital image. On this basis, we evaluate our model as suitable for use in production solutions.

1 Uvod

Med uporabna področja računalniškega vida spada tudi področje analize oseb na osnovi posnetkov obrazov, kjer s postopki računalniškega vida iz posnetka obraza pridobimo želene informacije o osebi na sliki. Primeri uporabe informacij o obrazu so samodejno beleženje prisotnih študentov na predavanjih (Grover, 2019), samodejno preverjanje potnih listov na letališčih (MRZ, 2022) ali pa verifikacija identitete kot dodaten varnostni sloj pri dvigu gotovine na bankomatu (Murugesan, 2020).

Zelo koristna informacija, ki jo lahko pridobimo iz posnetka obraza, je starost osebe. Podatek o starosti osebe nam koristi na področjih, kot so nadzor dostopa do vsebin, biostatistika, profiliranje uporabnikov, pametno oglaševanje, simulacije staranja in varnostni sistemi.

Določanje starosti osebe je zahtevno opravilo že za človeka, razlog za to pa so različni dejavniki, kot so na primer življenjski slog osebe (kajenje, slabo prehranjevanje), ki lahko osebo navidezno postarajo, okoljski vplivi (življenje v onesnaženih okoljih, visoka izpostavljenost soncu), genetika, poraščenost obraza ali uporaba ličil, različni vzorci staranja pri različnih rasah itd. Na zajetih posnetkih k temu dodamo še dodatne dejavnike, kot so poza osebe, izraz na obrazu in osvetlitev fotografije. Vse to še dodatno oteži analiziranje slike.

Za nov preboj pri ocenjevanju starosti na osnovi digitalnih posnetkov so poskrbele metode strojnega učenja, natančnejše globoke konvolucijske nevronske mreže, ki so se izkazale kot dobra rešitev že pri veliko drugih težavah procesiranja vizualnih podatkov (Glavez, 2018), (Sultana, 2020).

V tej raziskavi predstavimo naš razvit model konvolucijske nevronske mreže za ocenjevanje starosti osebe iz digitalnega posnetka. Nad modelom izvedemo več eksperimentov. Najprej preizkusimo natančnost razvitega modela ob različnih nastavitvah hiperparametrov in izvedemo primerjavo s sorodnimi deli. Preverimo tudi, v kolikšni meri spol osebe na digitalnem posnetku vpliva na natančnost ocenjevanje starosti, oziroma, ali lahko različne vzorce staranja zajamemo tudi med učenjem konvolucijske nevronske mreže in jih izkoristimo za natančnejše

ocenjevanje starosti osebe na digitalnem posnetku. Rezultate in ugotovitve natančno kvantificiramo.

2 Algoritem za ocenjevanje starosti z uporabo modela VGG-Face

Za osnovo za svoj algoritem smo izbrali arhitekturo VGG-16, saj na njej temelji veliko uspešnih algoritmov s področja računalniškega vida (Parlhi, 2015), (Rothe, 2015). Izbrali smo različico modela VGG-16, imenovano VGG-Face, ki je bila razvita za razpoznavanje obrazov (Parlhi, 2015). Za to različico modela smo se odločili, ker je bil za namen razpoznavanja obrazov model učen na več kot milijonu digitalnih posnetkov obrazov. Takšen model se je naučil prepoznavati značilnice obrazov, ki so uporabne tudi za naš sistem za ocenjevanje starosti. V naši raziskavi uporabimo prenosno učenje (angl. transfer learning). Da bi ohranili že naučeno znanje, ki smo ga pridobili z uporabo uteži drugega, že naučenega modela, smo v našem modelu zamrznili vse plasti, razen zadnjih treh konvolucijskih. Uteži zamrznjenih plasti se med učenjem niso spreminjale (Krel, 2021).

Ocenjevanje starosti smo opredelili kot klasifikacijski problem s 101 razredom, ki predstavljajo leta od 0 do vključno 100. Zadnja konvolucijska plast v VGG-Face za potrebe razpoznavanja obrazov vsebuje 2.622 izhodov, ki smo jih nadomestili s svojimi klasifikacijskimi razredi.

Tako smo dobili model, ki na vhodu pričakuje digitalni posnetek velikosti 224 x 224 slikovnih elementov s tremi barvnimi kanali, izhod pa predstavlja 101 klasifikacijski razred. Model ima trinajst konvolucijskih plasti in skupno 134.674.341 parametrov, pri čemer je 119.959.653 parametrov možno učiti, 14.714.688 parametrov pa je zamrznjenih.

3 Eksperimentalno okolje

Izvorno kodo algoritma smo pripravili v splošno namenskem programskem jeziku Python. Za delo z nevronskimi mrežami smo uporabili odprtokodno programsko ogrodje Keras (Keras, 2022), ki služi kot vmesnik do odprtokodne knjižnice TensorFlow (Tensorflow, 2022), ki ponuja širok nabor operacij za učenje in delo z nevronskimi mrežami (Krel, 2021).

Ker je učenje konvolucijske nevronske mreže računsko zahtevno in težko izvedljivo na povprečnem osebem računalniku, smo za učenje uporabili storitev Google Colaboratory (Google Colaboratory, 2022). Google Colaboratory omogoča izvedbo kode Python v oblaku prek spletnega brskalnika. Storitev je priljubljena predvsem pri razvijalcih algoritmov strojnega učenja, saj omogoča dostop do zmogljivih grafičnih procesnih enot.

3.1 Uporabljene slikovne baze podatkov za ocenjevanje starosti

V tej raziskavi smo za učenje in eksperimente uporabili podatkovni bazi FG-NET in IMDB-WIKI.

Baza podatkov FG-NET (FG-NET, 2022) vsebuje 1.002 digitalnih posnetkov 82 različnih oseb. Razpon starosti oseb je med 0 in 69 let. Informacija o starosti osebe na digitalnem posnetku je shranjena v imenu slikovne datoteke v formatu: id_osebe_A_starost_osebe.jpg (primer: datoteka, poimenovana 065A09.jpg, vsebuje digitalni posnetek obraza osebe z identifikatorjem 065 in starostjo 9 let).

Baza podatkov IMDB-WIKI (IMDB-WIKI, 2022) vsebuje 524.230 digitalnih posnetkov obrazov različnih znanih oseb, samodejno pridobljenih s spletnih strani IMDB in Wikipedija. Posnetkom je dodana datoteka s končnico .mat, iz katere lahko z orodjem Matlab preberemo različne metapodatke o osebi na digitalnem posnetku (npr. datum rojstva, spol). Baza je razdeljena na več direktorijev z digitalnimi posnetki, metapodatki o posnetkih pa so zapisani v dveh ločenih datotekah za digitalne posnetke, pridobljene s spletne strani IMDB in za digitalne posnetke, pridobljene iz Wikipedije. Ker so posnetki pridobljeni samodejno s spleta, vsebujejo tudi napačne in neustrezne posnetke ter napačne metapodatke, zato moramo podatkovno bazo prečistiti.

4. Eksperimenti

Izvedli smo več različnih eksperimentov. Najprej smo večkrat ponovili učenje nevronske mreže za iskanje optimalnih nastavitvev hiperparametrov. Zatem smo izvedli eksperimente, s katerimi smo preverili vpliv spola na uspešnost ocenjevanja starosti ter opravili primerjavo našega algoritma s sorodnimi rešitvami (Krel, 2021).

4.1 Eksperiment 1: Učenje modela na podatkovni bazi IMDB-WIKI

Najprej smo preverili, kako dobro se je naš model sposoben naučiti ocenjevanja starosti. Za učenje in validacijo smo uporabili prečiščeno bazo IMDB-WIKI, kjer smo za učenje uporabili 95.121 digitalnih posnetkov iz podmnožice IMDB. Učenje smo izvedli v 500 epohah. V vsaki epohi smo naključno prebrali 10.240 digitalnih posnetkov, ki smo jih razdelili v razmerju 80:20; 8.192 posnetkov za učenje in 2.048 posnetkov za validacijo. Uspešnost smo ovrednotili s povprečno napako med dejansko in ocenjeno starostjo osebe nad 5.000 posnetkih iz podmnožice WIKI.

Eksperiment smo izvedli večkrat, s preizkusom različnih nastavitev hiperparametrov in optimizacijskih algoritmov. Preizkusili smo dva različna optimizacijska algoritma, in sicer stohastični gradientni sestop (angl. Stochastic Gradient Descent, SGD) ter optimizacijski algoritem Adam.

Pri optimizacijskem algoritmu SGD smo uporabili dva načina, in sicer s konstantno stopnjo učenja (angl. learning rate) in s spremenljivo stopnjo učenja, hiperparametre pa nastavili kot:

- moment (angl. momentum): 0,9,
- stopnjo učenja: 0,01,
- upad stopnje učenja (uporabljeno le pri modelu s spremenljivo stopnjo učenja): 0,001.

Pri optimizacijskem algoritmu Adam smo izbrali stopnjo učenja 0,01. Vrednosti vseh omenjenih hiperparametrov so nastavljene glede na priporočila oziroma privzete vrednosti v ogrodju Keras.

Učenje smo opravili v paketnem načinu, kjer smo kot del eksperimenta spreminjali število posnetkov v paketu, in sicer smo uporabili pakete s 64, 128 in 256 posnetki.

4.2 Eksperiment 2: Vpliv spola na uspešnost ocenjevanja starosti

V naslednjem eksperimentu smo preverjali, ali lahko s poznavanjem spola osebe izboljšamo natančnost ocenjevanja starosti. Ponovno smo uporabili podatkovno bazo IMDB-WIKI, kjer metapodatki vsebujejo tudi podatek o spolu osebe. Iz podmnožice IMDB smo tvorili novi podmnožici s posnetki samo moških (45.413 posnetkov) in s posnetki samo žensk (48.673 posnetkov). Učenje smo izvedli dvakrat od začetka, in sicer prvič le na podmnožici posnetkov s samo moškimi, pri čemer smo posnetke moških razdelili v razmerju 80 % (36.330 posnetkov) za učenje in 20 % (9.083 posnetkov) za validacijo, nato pa še na podmnožici posnetkov s samo ženskami, kjer smo posnetke razdelili v enakem razmerju 80:20 (38.938 posnetkov za učenje in 9.735 posnetkov za validacijo). Pri učenju obeh modelov smo uporabili optimizacijski algoritem SGD z momentom 0,9, stopnjo učenja 0,01 in pakete z velikostjo 128 posnetkov. Za ocenjevanje uspešnosti smo pripravili tri različne testne baze podatkov iz podmnožice WIKI, pri čemer smo v prvi izbrali 5.000 posnetkov obeh spolov, v drugi 5.000 posnetkov le moških in v tretji 5.000 posnetkov le žensk. Uspešnosti smo vrednotili s povprečno napako med dejansko in ocenjeno starostjo osebe nad vsemi tremi testnimi bazami ločeno za model, učen le na posnetkih moških in za model, učen le na posnetkih žensk (Krel, 2021).

4.3 Eksperiment 3: Primerjava s sorodnimi deli

V zadnjem eksperimentu smo primerjali naš algoritem in s sorodnimi deli. Za eksperiment smo uporabili model z optimizacijskim algoritmom SGD, momentom 0,9, stopnjo učenja 0,01 ter pakete velikosti 128 posnetkov. Model smo učili na posnetkih iz podatkovne baze FG-NET po metodi validiranja LOPO (angl. Leave-One-Person-Out).

4.4 Metrike za vrednotenje rezultatov

Za vrednotenje smo uporabili metriko povprečne absolutne napake MAE (angl. Mean Absolute Error) in metodo validiranja LOPO. Z metriko MAE merimo razliko med ocenjeno in dejansko starostjo osebe. Izračunamo jo kot:

$$MAE = \frac{1}{K} \sum_{i=1}^K |y'_i - y_i| \quad (1)$$

kjer y'_i pomeni ocenjeno starost osebe na i -tem posnetku, y_i dejansko starost osebe na i -tem posnetku in K število testnih posnetkov. Manjša izračunana napaka seveda pomeni boljši rezultat.

Primerjavo svojega algoritma z obstoječimi rešitvami smo opravili z metodo LOPO na bazi podatkov FG-Net, ki vsebuje 1.002 posnetkov 82 različnih oseb, posnetih pri različnih starostih. V eksperimentu smo pripravili isto število modelov, kot je različnih oseb v množici. Za vsakega izmed modelov smo pripravili testno in učno množico tako, da smo tvorili novo učno množico na način, da smo iz originalne učne množice izločili vse posnetke izbrane osebe. Te izločene posnetke, ki vsi pripadajo izbrani osebi, smo uporabili kot testno množico, na kateri smo izračunali MAE. Postopek smo ponovili za vse osebe v bazi podatkov. Na koncu smo izračunali še povprečje vseh izračunanih MAE (kar označimo z LOPO) kot:

$$LOPO = \frac{1}{L} \sum_{i=1}^L MAE_i \quad (2)$$

kjer L pomeni število vseh različnih oseb v bazi podatkov, MAE_i pa izračunano napako med ocenjenimi in dejanskimi starostmi na posnetkih i -te osebe.

5 Rezultati

V prvem eksperimentu smo našo mrežo učili na bazi podatkov IMDB-WIKI, pri čemer smo uporabili 95.121 posnetkov iz podmnožice IMDB za učenje, testirali pa na 5.000 še nevidenih posnetkih iz podmnožice WIKI. Tabela 1 prikazuje rezultate, ovrednotene z metriko MAE, za različne optimizacijske algoritme in različne velikosti paketov ob različnem številu epoh, pri čemer je za vsako kombinacijo nastavitve poudarjen najboljši rezultat (najmanjša metrika MAE).

Tabela 1: Rezultati eksperimenta 1. Podana metrika MAE (v letih), pri čemer je najboljši rezultat označen krepko.

Nastavitev učenja	Epoh 100	Epoh 200	Epoh 300	Epoh 400	Epoh 500
SGD (64 posnetkov v paketu)	6,99	6,80	6,72	6,85	7,02
SGD (128)	7,43	7,09	6,83	6,75	6,68
SGD (256)	8,94	7,68	7,20	6,99	6,83
SGD LR (64)	7,00	7,23	7,38	7,38	7,31
SGD LR (128)	7,10	7,12	7,34	7,45	7,58
SGD LR (256)	6,99	7,39	7,42	7,41	7,60
ADAM (64)	7,00	6,99	7,08	7,20	7,39
ADAM (128)	6,97	7,40	7,45	7,43	7,44
ADAM (256)	7,29	7,40	7,62	7,77	7,69

Opazimo, da z uporabo spreminjajoče se stopnje učenja (SGD s spremenljivo stopnjo učenja in Adam) dosežemo dobre rezultate že po 100 oziroma 200 epohah, vendar pa najboljši rezultat med vsemi dosežemo z uporabo učnega algoritma SGD s konstantno stopnjo učenja po 500 epohah.

Trajanje neprekinjenega učenja modela z učnim algoritmom SGD in 128 posnetki v paketu je trajalo 564 minut, učenje modela z učnim algoritmom Adam in 128 posnetki v paketu pa 447 minut. Med učenji prihaja do razlik v potrebnem času za učenje, saj se v okolju Google Colaboratory razpoložljiva sredstva dodeljujejo dinamično glede na zasedenost virov.

V drugem eksperimentu smo našo nevronske mrežo učili ločeno za posnetke moških in žensk na bazi podatkov IMDB-WIKI, pri čemer smo za učenje uporabili 45.413 posnetkov moških in 48.673 posnetkov žensk iz podmnožice IMDB. Tabela 2 prikazuje rezultate modela, učenega na posnetkih moških, in ovrednotenega z metriko MAE nad 5.000 še nevidenimi posnetki moških, žensk in obeh spolov skupaj iz podmnožice WIKI, tabela 3 pa prikazuje po enakem protokolu še rezultate modela, učenega na posnetkih žensk.

Tabela 2: Rezultati eksperimenta 2 (različica A): Model je bil naučen le na posnetkih moških. V tabeli je podana metrika MAE (v letih).

Testna množica WIKI	MAE
Moški	6,71
Ženske	10,45
Vsi	7,88

Tabela 3: Rezultati eksperimenta 2 (različica B): Model je bil naučen le na posnetkih žensk. V tabeli je podana metrika MAE (v letih).

Testna množica WIKI	MAE
Moški	9,20
Ženske	7,51
Vsi	8,71

V zadnjem eksperimentu smo naredili primerjavo med našo metodo in sorodnimi deli. Rezultate, ki smo jih nad bazo podatkov FG-NET ovrednotili z validacijsko metodo LOPO, prikazuje tabela 4. Najboljša metoda ima najnižjo vrednost MAE.

Tabela 4 Rezultati eksperimenta 3: Primerjava s sorodnimi deli na množici FG-NET s pomočjo validacije LOPO. V tabeli je podana metrika MAE (v letih).

Metoda	MAE
AGES (Geng, 2007)	6,77
CPNN (Geng, 2013)	4,76
<i>Človeški opazovalec (Han, 2015)</i>	4,70
MV (Pan, 2018)	4,10
Naša rešitev	3,91
DRFs (Shen, 2018)	3,85
DEX (Rothe, 2015)	3,09
BridgeNet (Li, 2019)	2,56

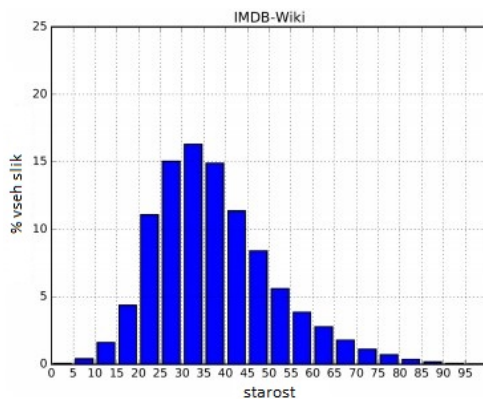
6 Diskusija in zaključek

V tem članku smo predstavili modificirano arhitekturo VGG-Face, namenjeno ocenjevanju starosti osebe na osnovi enega digitalnega posnetka. Pri učenju smo si pomagali s prenosnim učenjem. Rezultati testiranja so pokazali, da smo se približali sodobnim metodam, še bolj spodbudno pa je spoznanje, da smo na bazi podatkov FG-NET s pomočjo validacije LOPO dosegli boljšo oceno starosti kot človeški opazovalec.

Menimo, da je povprečna napaka ocene 6,7 leta na podatkovni bazi IMDB-WIKI in 3,9 leta na podatkovni bazi FG-NET dovolj dober rezultat, da ocenimo svoj model kot primeren za uporabo v produkcijskih rešitvah. Ob ocenjevanju starosti namreč le redkokdaj potrebujemo natančnejšo oceno starosti, pogosteje nas zanimajo določeni širši razponi let, na primer, v kakšne starostne skupine spadajo udeleženci dogodka, ki jih posnamemo pri vходу v dvorano, ali pa s kakšno gotovostjo lahko

ocenimo, da je oseba pri nakupu alkoholne pijače na avtomatu polnoletna (Krel, 2021).

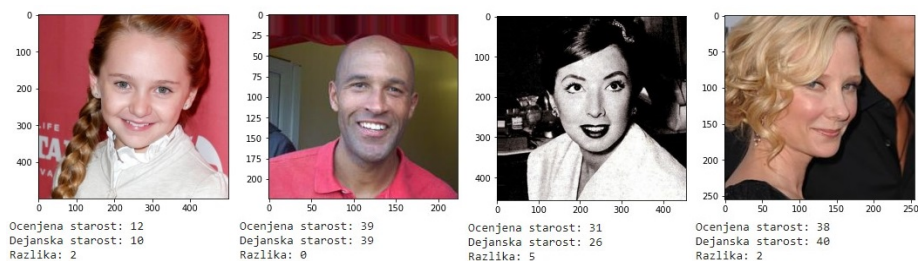
Kot glavni prednosti predlagane metode navajamo primerno uspešnost ocenjevanja starosti osebe in zmožnost hitrega prilaganja naše rešitve z uporabo prenosnega učenja. Med slabosti pa štejemo omejitve, da je naša rešitev trenutno zmožna obdelati le en obraz naenkrat, ter slabše ocenjevanje starosti mlajših otrok in starejših oseb. Slednje je nedvomno posledica pomanjkanja posnetkov v teh starostnih skupinah oziroma neenakomerne porazdelitve učnih vzorcev po starosti v bazi podatkov IMDB-WIKI (glej sliko 1). Pomanjkanje kakovostne baze podatkov z zadostnim in uravnoteženim številom označenih digitalnih posnetkov je sicer težava, ki jo omenjajo tudi avtorji sorodnih del (Rothe, 2015).



Slika 1: Porazdelitev posnetkov po starosti v bazi podatkov IMDB-WIKI

Vir: lasten

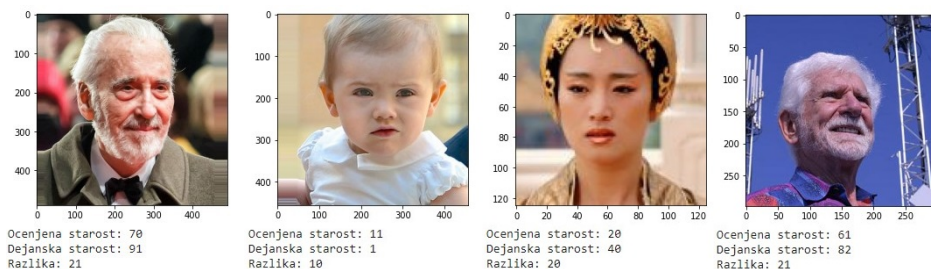
Na sliki 2 smo predstavili nekaj testnih primerov dobro ocenjenih starosti z modelom, učenim pri 500 epohah na podmnožici IMDB iz podatkovne baze IMDB-WIKI z učnim algoritmom SGD in 128 posnetki v paketu (eksperiment 1). Opazimo, da z našo rešitvijo dobro ocenimo starost osebe na posnetku za različen spol, položaj obraza ali izraz na obrazu.



Slika 2: Primeri dobrega ocenjevanja starosti osebe na štirih testnih posnetkih s pomočjo modificiranega modela VGG-Face

Vir: lasten

Na sliki 3 pa predstavljamo nekaj testnih primerov slabo ocenjenih starosti osebe, pri čemer smo uporabili enak naučen model, kot je bil opisan zgoraj. Opazimo, da gre najpogosteje za posnetke otrok, mlajših od deset let, ali pa za osebe, starejše od osemdeset let. Problem je v pomanjkanju podatkov v teh starostnih skupinah v bazi IMDB-WIKI (glej še tekst zgoraj). Pogosto je napačna ocenitev posledice razlike med navidezno in biološko starostjo osebe (na primer tretja slika z leve na sliki 3).



Slika 3: Primeri slabšega ocenjevanja starosti osebe na štirih testnih posnetkih s pomočjo modificiranega modela VGG-Face

Vir: lasten

V prihodnosti želimo ovrednotiti vpliv drugih hiperparametrov, ki jih v tem delu še nismo analizirali. Na tak način bi poskusili izboljšati natančnost ocenjevanja starosti in tudi hitrost učenja. Prav tako bi lahko odpravili omejitev zaznavanja in ocenjevanja le ene osebe na posnetek. Nadaljnje delo bomo usmerili še v preizkušanje drugih modernih arhitektur CNN za določanje starosti osebe.

Verjamemo, da bi natančnost naše rešitve še izboljšali, če bi pripravili ali pridobili večjo in kakovostnejšo učno množico.

Literatura

- Antipov G., Baccouche M., Berrani S., Dugelay J. Apparent age estimation from face images combining general and children-specialized deep learning models. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, (2016), str. 801-809.
- FG-NET dataset by Yanwei Fu. Dostopno na: https://yanweifu.github.io/FG_NET_data/index.html [1.2.2022]
- Galvez R., Bandala A., Dadios E., Vicerra R., Maningo J. Object Detection Using Convolutional Neural Networks, *TENCON IEEE Region 10 Conference*, (2018), str. 2023-2027.
- Geng X., Zhou Z., Miles K. Automatic age estimation based on facial aging patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29, (2007), str. 2234-2240.
- Geng X., Yin C., Zhou Z. Facial age estimation by learning from label distributions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35, (2013), str. 2401-2412.
- Google Colaboratory. Dostopno na: <https://colab.research.google.com/> [1.2.2022].
- Grover V., Chhabra N., Attendance Monitoring system Through Face Recognition, *2019 6th International Conference on Computing for Sustainable Global Development (INDIACom)*, (2019), str. 179-183.
- Han H., Otto C., Liu X., Jain A.K. Demographic estimation from face images: Human vs. Machine Performance, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37, (2015), str. 1148-1161.
- IMDB-WIKI – 500k+ face images with age and gender labels. Dostopno na: <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/> [1.2.2022]
- Keras. Dostopno na: <https://keras.io/> [1.2.2022].
- Kerl T. Ocenjevanje starosti osebe na osnovi digitalnih posnetkov z uporabo konvolucijskih nevronske mreže. *Magistrsko delo*, 2021.
- Li W., Lu J., Feng J., Xu C., Zhou J., Tian Q. BridgeNet: A continuity-aware probabilistic network for age estimation, *IEEE Computer Vision and Pattern Recognition*, (2019), str. 1146-1154.
- MRZ code extraction from visa and passport documents using convolutional neural networks. Dostopno na: <https://arxiv.org/abs/2009.05489> [1.2.2022].
- Murugesan M., Santhosh M., Sasi T., Sasiwarman M., Valanarasu I. Securing ATM Transactions using Face Recognition, *International Journal of Advanced Trends in Computer Science and Engineering*, 9, (2020), str. 1295-1299.
- Pan H., Han H., Shan S., Chen X. Mean-variance loss for deep age estimation from face, *IEEE Computer Vision and Pattern Recognition*, (2018), str. 5285-5294.
- Parlhi O., Vedaldi A., Zisserman A., Deep Face Recognition, *British Machine Vision Conference*, (2015), str. 41.1-41.12.
- Rothe R., Timofte R., Gool L. DEX: Deep EXpectation of Apparent Age from a Single Image. *IEEE International Conference on Computer Vision Workshop*, (2015), str. 252-257.
- Shen W., Guo Y., Wang Y., Zhao K., Wang B., Yuille A.L. Deep regression forests for age estimation, *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (Junij 2018), str. 2304-2313.
- Sultana F., Sufian A., Dutta P. Evolution of Image Segmentation using Deep Convolutional Neural Network: A survey, *Knowledge-Based Systems*, 201-202, (2020), str. 1-38.
- Tensorflow: An end-to-end open source machine learning platform. Dostopno na: <https://tensorflow.org> [1.2.2022]

ROSUS 2022 - RAČUNALNIŠKA OBDELAVA SLIK IN NJENA UPORABA V SLOVENIJI 2022: ZBORNİK 16. STROKOVNE KONFERENCE

BOŽIDAR POTOČNIK (UR.)

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko,
Maribor, Slovenija.

E-pošta: bozidar.potocnik@um.si.

Povzetek ROSUS 2022 – Računalniška obdelava slik in njena uporaba v Sloveniji 2022 je strokovna računalniška konferenca, ki jo od leta 2006 naprej vsako leto organizira Inštitut za računalništvo iz Fakultete za elektrotehniko, računalništvo in informatiko, Univerze v Mariboru. Konferenca povezuje strokovnjake in raziskovalce s področij digitalne obdelave slik in strojnega vida z uporabniki tega znanja, pri čemer uporabniki prihajajo iz raznovrstnih industrijskih okolij, biomedicine, športa, zabavnštva in sorodnih področij. Zbornik konference ROSUS 2022 združuje strokovne prispevke več deset avtorjev, od tega dva vabljena predavanja ter več demonstracijskih prispevkov. Prispevki podajajo najnovejše dosežke slovenskih strokovnjakov s področij digitalne obdelave slik in strojnega vida, osvetljujejo pa tudi trende in novosti na omenjenih strokovnih področjih. Velik poudarek prispevkov je na promoviranju ekonomske koristnosti aplikacij računalniške obdelave slik in vida v slovenskem prostoru. Takšne računalniške aplikacije zaradi visoke natančnosti, robustnosti in izjemnih hitrosti pri obdelovanju informacij nudijo namreč nove priložnosti za uveljavitev na trgu visokih tehnologij.

Ključne besede:

računalniška
obdelava
slik
strojni
vid,
biomedicina,
industrijske
aplikacije,
prenos
znanja

ROSUS 2022 - COMPUTER IMAGE PROCESSING AND ITS APPLICATION IN SLOVENIA 2022: PROCEEDINGS OF THE 16TH PROFESSIONAL CONFERENCE

BOŽIDAR POTOČNIK (ED.)

University of Maribor, Faculty of Electrical Engineering and Computer Science,
Maribor, Slovenia.

E-mail: bozidar.potocnik@um.si.

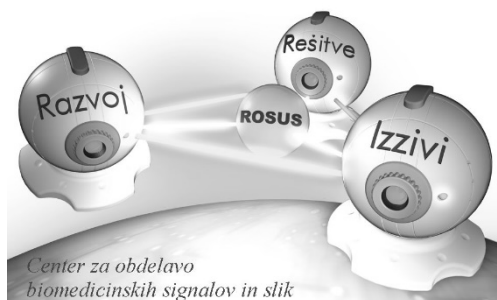
Abstract ROSUS 2022–Computer image processing and its application in Slovenia 2022 is a professional conference that, since 2006, has been organised every year by the Institute of Computer Science of the Faculty of Electrical Engineering and Computer Science, University of Maribor. The conference connects researchers in the fields of Image Processing and Machine Vision with users of this knowledge, whereby users are coming from diverse industrial environments, such as Biomedicine, Sport, Entertainment, and related fields. The proceedings of ROSUS 2022 combine scientific articles by dozens of authors, including two invited lectures and several demonstration articles. Contributions represent the latest achievements of Slovenian experts in the fields of Image Processing and Vision, and also highlight trends and novelties in these areas. Great emphasis is on promotion of the economic usefulness of Image Processing and Vision applications in the Slovenian region. Namely, such software, due to high precision, robustness, and exceptional speed in information processing, provides new opportunities for penetration on the high technology.

Keywords:

computer
image
processing,
machine
vision,
biomedicine,
industrial
applications,
knowledge
transfer.



SLOVENSKI PODJETNIŠKI SKLAD



ROSUS 2022
<http://rosus.feri.um.si>



SPS V ŠTEVILKAH

Σ

PLAN 2022

»V letu 2022
SPS načrtuje
pomagati
3.685
podjetjem z
152,96
mio EUR
finančnimi in
vsebinskimi
pomočmi.«

Σ

REZULTATI 2021

»V letu 2021 je
SPS pomagal
6.624
podjetjem z
235,11
mio EUR
finančnimi in
vsebinskimi
pomočmi.«

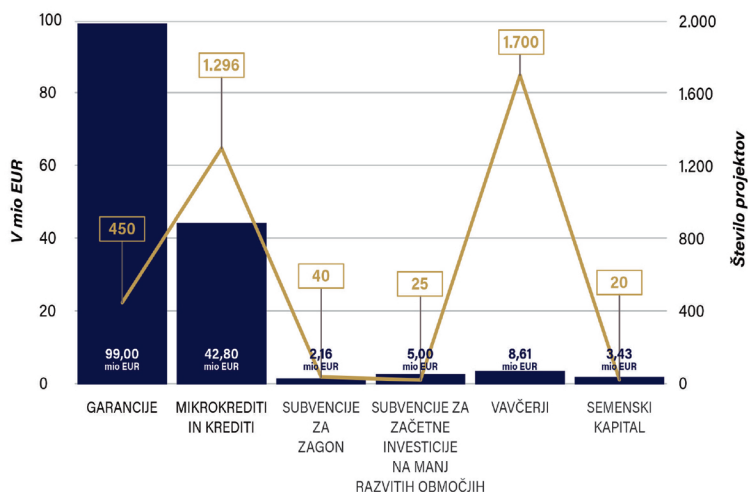
I. VLOGA SPS V LETU 2022

SPS tudi v letu 2022 ostaja:

1. vodilni ponudnik garancij za bančne kredite,
2. vodilni ponudnik (mikro)kreditov,
3. sooblikovalec start-up ekosistema v Sloveniji za doseganje globalnega uspeha start-up podjetij in novonastalih podjetij na problemskih območjih z zagonskimi spodbudami,
4. ključni soinvestitor semenskega in tveganega kapitala kot lastniškega financiranja v povezavi z vsebinsko podporo,
5. edini ponudnik spodbud malih vrednosti – vavčerjev,
6. sooblikovalec podjetniškega okolja za podjetniški sektor v Sloveniji,
7. povezovalac podjetniške mreže v mednarodnem okolju.

Plan finančne pomoči v letu 2022:

- o **152,96 mio EUR novih pomoči**
- o **cca 3.685 podjetij**, ki bodo v letu 2022 lahko koristila pomoč SPS-a
- o pri tem SPS ocenjuje, da bodo vložena sredstva spodbudila **cca 237,00 mio EUR novih investicij v gospodarstvu**, kar je primerljivo s preteklimi leti oz. z letom 2019.



PRODUKTI 2022	NAČRTOVANA VIŠINA SREDSTEV	OBJAVA/ PREDVIDENA OBJAVA
1. Garancije – P1 plus 2022 <i>(do 1.250.000 EUR/podjetje)</i>	99,00 mio EUR	April 2022
2. Mikrokrediti za obmejna problemska območja - P7R 2022 <i>(do 25.000 EUR/podjetje)</i>	7,00 mio EUR	April 2022
3. Zagonske spodbude za inovativna podjetja – P2 2022 <i>(do 54.000 EUR/podjetje)</i>	2,16 mio EUR	Februar 2022
4. Spodbude za začetne investicije na problemskih območjih – P4I 2022 <i>(do 300.000 EUR/podjetje)</i>	5,00 mio EUR	Februar 2022
5. Konvertibilna posojila za inovativna podjetja – SK75 2022 <i>(75.000 EUR/podjetje)</i>	1,13 mio EUR	Februar 2022
6. So –investiranje z zasebnimi investitorji – SI-SK 2022 <i>(do 400.000 EUR/podjetje)</i>	2,30 mio EUR	AKTIVEN 28. 1. 2022 Naslednji bližnji roki za oddajo vlog: 28. 2., 31. 3., 29. 4 in 31. 5. 2022
7. COVID krediti – P7 COVID 2022 <i>(do 50.000 EUR/podjetje)</i>	10,00 mio EUR	Maj 2022
8. Vavčerji ReactEU <i>(do 9.999,99 EUR/podjetje)</i>	8,61 mio EUR	Različno po sklopih 2021/2022
PRODUKTI iz leta 2021	RAZPISANA SREDSTVA	OBJAVA
9. Mikrokrediti za mikro in mala podjetja iz SS – P7 2021 <i>(do 25.000 EUR/podjetje)</i>	15,00 mio EUR	AKTIVEN 31. 12. 2021 Naslednji rok za oddajo vlog: 1. 3. 2022
10. COVID neposredni krediti za turizem in gostinstvo - P7C-2 2021 <i>(do 50.000 EUR/podjetje)</i>	10,8 mio EUR	ZAPRT 31. 12. 2021 450 vlog v ocenjevanju



Slovenski podjetniški sklad

Ulica Kneza Koclja 22
2000 Maribor
022341260
info@podjetniskisklad.si
www.podjetniskisklad.si



Univerza v Mariboru

Fakulteta za elektrotehniko,
računalništvo in informatiko

LSPO

Laboratorij za sistemsko
programsko opremo

<https://rosus.feri.um.si/rosus2022>
rosus.feri@um.si

