



Univerza v Mariboru

Fakulteta za elektrotehniko,
računalništvo in informatiko

Zbornik strokovne konference

ROSUS 2017

Računalniška obdelava slik in
njena uporaba v Sloveniji 2017

Urednik:
Božidar Potočnik



Maribor, 23. marec 2017



Univerzitetna
založba
Univerze v
Mariboru

© **Univerzitetna založba Univerze v Mariboru**

Vse pravice pridržane. Brez pisnega dovoljenja založnika je prepovedano reproduciranje, distribuiranje, predelava ali druga uporaba tega dela ali njegovih delov v kakršnemkoli obsegu ali postopku, vključno s fotokopiranjem, tiskanjem ali shranjevanjem v elektronski obliki.

- Naslov:** ROSUS 2017: Računalniška obdelava slik in njena uporaba v Sloveniji 2017
(Zbornik 12. strokovne konference)
- Urednik:** izr. prof. dr. Božidar Potočnik (Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko)
- Recenzija:** izr. prof. dr. Simon Dobrišek (Univerza v Ljubljani, Fakulteta za elektrotehniko)
red. prof. dr. Aleš Holobar (Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko)
- Lektoriranje:** Zvezdana Sabol Golob, prof. slov. j.
- Tisk:** Tiskarna Saje d.o.o.
- Št. izvodov:** 50

CIP - Kataložni zapis o publikaciji
Univerzitetna knjižnica Maribor
<http://press.um.si>
004.932(497.4)(082)

STROKOVNA konferenca ROSUS (12 ; 2017 ; Maribor)
ROSUS 2017 : računalniška obdelava slik in njena uporaba v Sloveniji 2017 : (zbornik 12. strokovne konference) / urednik Božidar Potočnik. - Maribor : Univerzitetna založba Univerze, 2017

50 izv.

ISBN 978-961-286-018-9
1. Potočnik, Božidar
COBISS.SI-ID 91046913

Založnik:

Univerzitetna založba Univerze v Mariboru
Sloškovo trg 15, 2000 Maribor, Slovenija
tel. +386 2 250 42 42, fax +386 2 252 32 45
<http://press.um.si>, zalozba@um.si

Izdajateljica:

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko
Smetanova ulica 17, 2000 Maribor, Slovenija
tel. +386 2 220 70 00, faks +386 2 220 72 72
<http://www.feri.um.si>, feri@um.si

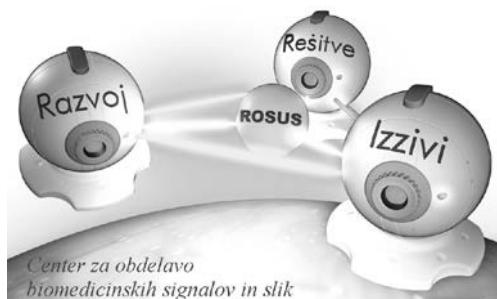
Cena: brezplačen izvod

Odgovorna oseba založnika: prof. dr. Igor Tičar, rektor Univerze v Mariboru



Univerza v Mariboru

Univerzitetna
založba
Univerze v
Mariboru



ROSUS 2017 - Računalniška obdelava slik in njena uporaba v Sloveniji 2017

(Zbornik 12. strokovne konference)

Urednik:
Božidar Potočnik

Marec 2017

ROSUS 2017 – Računalniška obdelava slik in njena uporaba v Sloveniji 2017

(Zbornik 12. strokovne konference)

BOŽIDAR POTOČNIK

Povzetek: ROSUS 2017 – Računalniška obdelava slik in njena uporaba v Sloveniji 2017 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, zabavništva in sorodnih področij. Zbornik konference ROSUS 2017 združuje 12 strokovnih prispevkov enaindvajsetih avtorjev, od tega 4 vabljena predavanja. Prispevki podajajo najnovejše dosežke slovenskih strokovnjakov s področij digitalne obdelave slik in strojnega vida, osvetlujejo 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

NASLOV UREDNIKA: dr. Božidar Potočnik, izredni profesor, Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Smetanova ulica 17, 2000 Maribor, Slovenija, e-pošta: bozidar.potocnik@um.si

KAZALO

VABLJENA PREDAVANJA	1
Matej Kenda , TopIT, Nova Gorica <i>Arhiviranje medicinskih slik in možnost uporabe arhiva za raziskave</i>	3
Matej Logar , Kolektor Vision, Idrija <i>Strojni vid v industrijskem okolju – izzivi in priložnosti</i>	10
Jure Škrabar , Kolektor Orodjarna, Idrija <i>Robustnost industrijskih sistemov strojnega vida</i>	15
Samo Šela , Kolektor Orodjarna, Idrija <i>Trendi na področju strojnega vida</i>	21
INDUSTRIJSKE APLIKACIJE	29
Aleš Hladnik , NTF, Ljubljana <i>Računalniška obdelava slike v papirništvu in tiskarstvu</i>	31
Teja Jereb , Urban Pavlovčič , Matija Jezeršek , FS, Ljubljana <i>Optična detekcija napak na visoko-reflektivnih struženih površinah</i>	38
Domen Tabernik , Danijel Skočaj , FRI, Ljubljana <i>Pregled programskih orodij za globoko učenje z vidika uporabe v industrijskih aplikacijah</i>	38
Janez Langus , Tomaž Šuštar , C3M, Ljubljana <i>NanoEye – program za semi-avtomatično analizo slik</i>	52
Peter Uršič , Domen Tabernik , Rok Mandeljc , Danijel Skočaj , FRI, Ljubljana <i>Detekcija velikega števila kategorij prometnih znakov</i>	59
MEDICINSKE IN BIOMEDICINSKE APLIKACIJE	68
Franc Janžekovič , Ahmad Mohmoudi , Tina Klenovšek , FNM, FS, Maribor, Mashhad, Iran <i>Ocena napake pri analizi objektov z geometrijsko morfometrijo</i>	69

ŠTUDENTSKE APLIKACIJE	77
Nejc Škerjanc, Peter Peer, FRI, Ljubljana <i>Obogatena resničnost gibanja uporabnika v realnem času na mobilni napravi s pomočjo podatkov globinskega senzorja</i>	79
Gašper Sedej, Martin Šavc, FERI, Maribor <i>Zakrivanje oseb na posnetkih gradbišča.....</i>	89

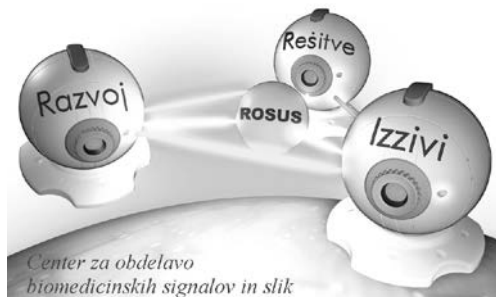
Spoštovani!

Po enajstih konferencah ROSUS 2006–2016 želimo tudi s konferenco ROSUS 2017 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.

Tudi letos namenjamo glavni poudarek aplikacijam s področja računalniške obdelave slik, ki so že integrirane oziroma pripravljene za integracijo v poslovne procese. Industrijsko sekcijo smo pod vodstvom dr. Jureta Skvarča temeljito prenovili. 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 pristrčno zahvalo Javnemu skladu Republike Slovenije za podjetništvo, ki je v okviru konference ROSUS 2017 predstavil zanimive finančne instrumente za spodbujanje prenosa tehnoloških rešitev v podjetniško sfero.

dr. Božidar Potočnik
predsednik konference
ROSUS 2017



Programski odbor ROSUS 2017:

dr. Aleš Holobar
dr. Damjan Zazula
dr. Boris Cigale
dr. Marjan Mernik
dr. Peter Peer
dr. Tomaž Tomažič
dr. Matjaž Colnarič
dr. Slobodan Ribarič
dr. Vojko Flis
dr. Saša Divjak

dr. Stanislav Kovačič
dr. Zdravko Kačič
dr. Aleš Leonardis
dr. Simon Dobrišek
dr. Mitja Lenič
dr. Franc Solina
dr. Andrej Šoštarič
dr. Franjo Pernuš
dr. Jure Skvarč
dr. Božidar Potočnik

Organizacijski odbor 2017:

Gašper Sedej
Jurij Munda
dr. Danilo Korže
dr. Vojko Glaser
Martin Šavc
Jernej Kranjec
dr. Aleš Holobar
dr. Božidar Potočnik

POKROVITELJI



ROSUS 2017
<http://rosus.feri.um.si/rosus2017/>

VABLJENA PREDAVANJA

ARHIVIRANJE MEDICINSKIH SLIK IN MOŽNOST UPORABE ARHIVA ZA RAZISKAVE

Matej Kenda

TopIT programska oprema, d.o.o.

E-pošta: matej.kenda@topit.si

URL: <http://www.topit.si>

POVZETEK: *Zakonodaja v EU in drugje po svetu predpisuje načine in roke arhiviranja medicinske dokumentacije. V skladi s predpisi o varovanju osebnih podatkov je potrebno dokumentacijo arhivirati na način, ki zagotavlja dostop do informacij samo pooblaščenim osebam. Večletno arhiviranje velikih količin medicinskih slik v digitalni obliki lahko predstavlja poseben tehnološki izziv in stroške za medicinske ustanove. Predstavljeno je arhiviranje medicinskih datotek kot samostojna storitev. Velik centralni arhiv predstavlja tudi možnost dostopa pacientov do svojih osebnih podatkov, omogočanje dostopa do vsebin tretjim osebam in uporabe anonimiziranih podatkov za raziskovalne namene.*

1. UVOD

Podjetje TopIT sodeluje z nekaj podjetji iz Nemčije pri razvoju programske opreme in rešitev za dolgoročno arhiviranje velikih količin nestrukturiranih podatkov. Srečujemo se z izzivi zagotavljanja zanesljivosti, zasebnosti in integritete arhiviranih podatkov, primerne izbora formatov arhiviranja, hitrih sprememb tehnologij in staranja nosilcev podatkov (npr. diskov, trakov in oblčnih arhivov).

Dolgoročno arhiviranje medicinskih slik je eno od specializiranih področij, na katerem pripravljamo rešitve s pomočjo programske opreme, ki ustreza strogim zahtevam nemške zakonodaje.

Arhiv medicinskih slik lahko zmanjša število preiskav, ki jih mora opraviti pacient, zagotavlja boljše diagnoze z uporabo podatkov za drugo mnenje in omogoča uporabo teh podatkov v znanstvene namene.

2. IZZIVI ARHIVIRANJA MEDICINSKIH SLIK IN UPORABA ARHIVA V RAZISKOVALNE NAMENE

2.1 Dolgoročno arhiviranje podatkov v digitalni obliki

Veliko informacij je v zadnjih letih na razpolago v digitalni obliki. Nekateri od teh podatkov so digitalizirani, drugi pa obstajajo samo v digitalni obliki.

Pri načrtovanju sistemov za arhiviranje digitalnih podatkov moramo biti pozorni na dve pomembni zahtevi:

- Podatke bo možno prebrati z medijev, na katerih so shranjeni. Tehnološki razvoj se bo nadaljeval, kar pomeni, da bodo nosilci shranjevanja podatkov izboljšani in nadomeščeni z drugačnimi.
- Prebrane podatke bo možno pravilno interpretirati. Pomembna je izbira formata oziroma načina zapisa podatkov in šifriranih algoritmov.

Nekateri vidiki teh zahtev se lahko rešijo s pravilnim načrtovanjem programske opreme in celotnega sistema za dolgoročno arhiviranje:

- Pripravo postopkov za prenos arhiviranih podatkov na novejšo tehnologijo, npr.:
 - prepis podatkov s trakov tipa LTO generacije 4 na 6 ali novejše,
 - prenos podatkov iz diskovnih polj v shrambo v oblaku.
- Nešifrirane datoteke je možno pretvoriti v novejše formate kot del procesa vzdrževanja arhiva.

Pri načrtovanju sistema za digitalno arhiviranje se srečamo tudi z zahtevami za:

- zagotavljanje učinkovitega, varnega in zanesljivega prenosa velikih količin podatkov preko omrežja,
- zagotavljanje pravic dostopa do podatkov (avtentikacija, avtorizacija),
- zagotavljanje varovanja osebnih podatkov s šifriranjem podatkov med prenosom in pri hrambi.

Zaželeno je, da se pri shranjevanju v digitalni arhiv izberejo formati datotek, ki so standardizirani in namenjeni temu, da jih bo mogoče prebrati tudi v prihodnosti. Del odgovornosti za pravilno izbiro nosijo tudi stranke, ki se odločijo za uporabo digitalnega arhiva.

Primeri takih formatov datotek so:

- besedilne datoteke (ASCII ali Unicode),
- strukturirane besedilne datoteke (npr. XML),
- PDF/A,
- TIFF, JPEG, MPEG-4,
- ODF.

Kadar ni mogoče izbrati formata za dolgoročno arhiviranje, je potrebno zagotoviti, da bo ob izteku obdobja arhiviranja še vedno na voljo strojna in programska oprema, s katero bo mogoče prebrati take datoteke.

Potrebo po dolgoročnem shranjevanju datotek imajo med drugimi knjižnice, muzeji, raziskovalne, arhivske in medicinske ustanove.

2.2 Posebnosti arhiviranja medicinskih slik

Arhiviranje medicinskih slik in z njimi povezanih podatkov, kot so izvidi, predstavlja posebno področje, ki ima svoje specifične zahteve.

Močno je poudarjena zahteva po zagotavljanju zasebnosti in nadzoru pravic dostopa do podatkov, ker so to osebni podatki, ki se ščitijo z zakonodajo o varstvu osebnih podatkov.

Za shranjevanje medicinskih slik se ponavadi uporablja format zapisa DICOM (Digital Imaging and Communications in Medicine). Ta vsebuje sliko in pripadajoče attribute, ki opisujejo pacienta, za katerega je bila preiskava narejena, ter podrobnosti o preiskavi.

Standard DICOM opisuje tudi komunikacijski protokol, ki ga uporabljajo modalitete za pošiljanje slik. Primeri modalitet so naprave za rentgen, magnetno resonanco (MR) in računalniško tomografijo (CT).

V državah Evropske unije in drugod zakonodaja predpisuje medicinskim ustanovam, koliko časa morajo hraniti slike in pripadajoče dokumente. Razlogov je več:

- spremljanje zdravljenja,
- zahteve zdravstvenih zavarovalnic,
- zahteve inšpekcijskih služb,
- potrebe organov pregona in sodišč.

Zakonodaja v različnih državah predpisuje različne časovne roke. Slovenski Zakon o zbirkah podatkov s področja zdravstva (ZZPPZ) predpisuje različne dobe hranjenja za različne vrste podatkov. Večinoma je predpisan rok 10 ali 15 let, za nekatere vrste 10 let po smrti oziroma če ta ni znan 100 let od nastanka podatka, lahko pa celo trajno hranjenje.

2.3 Analiza količin podatkov

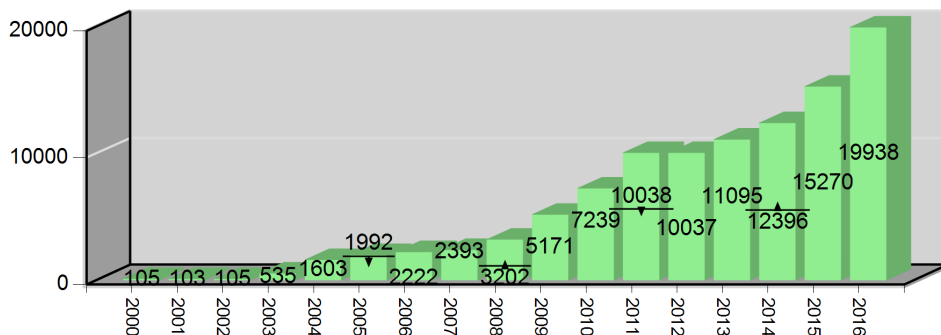
V podjetju TopIT smo opravili analizo velikosti datotek iz nekaj nemških klinik in na podlagi teh podatkov ugotovili povprečno velikost slike, ki znaša od 0,2 MB do 3,1 MB in povprečno količino podatkov za eno preiskavo, ki znaša od 27 MB do 118 MB.

Primer UKC Maribor: velikost CR je približno 7MB, CT pribl 200 kB, MR 150 kB do 250 kB. Število slik na preiskavo: 4 do 10 za skelet, CT 500 do 1000, MR 1000 do 1500, največ do 5000.

Modalitete so precej drage naprave, zato je poskrbljeno, da so čimbolj optimalno v uporabi. Glede na nam dostopne informacije trajajo preiskave od nekaj minut do pol ure.

Če predpostavimo, da je modaliteta uporabljena 6 ur na dan in 20 dni v mesecu, lahko izračunamo, da vsaka »proizvede« nekaj 10 GB podatkov na mesec.

Količina podatkov, ki so jih shranili v letu 2016 v dveh srednje velikih kliničnih centrih na Bavarskem, je bila 2,5 TB oz. 4,6 TB. Ta v zadnjih letih narašča.



Slika 1: Skupna količina shranjenih podatkov po letih (GB) v vzorčni kliniki
Velike medicinske ustanove imajo večje število različnih modalitet: UKC Regensburg pribl 55, UKC Maribor 20 RTG, 2 CR, 2 MR.

Trenutno sta najbolj uporabljena dva tipa tehnologije za shranjevanje velikih količin podatkov:

- trdi diski, sestavljeni v diskovna polja,
- tračne kasete, vstavljene v tračne knjižnice.

Zanesljivi sistemi za diskovna polja (Storage Area Network, SAN) ali tračne knjižnice s pripadajočimi strežniki in osebjem, ki vzdržuje take sisteme, lahko predstavljajo za medicinsko ustanovo precejšen strošek. Večji (univerzitetni) klinični centri že imajo računske centre in včasih opravljajo to dejavnost v sklopu dejavnosti takega centra.

2.4 Dolgoročno arhiviranje medicinskih datotek kot storitev

Pri razvoju programske opreme sodelujemo s podjetjem Telepaxx, ki se je že pred 20 leti odločilo, da ponudi arhiviranje medicinskih datotek kot storitev, najprej v Nemčiji, nato pa tudi v drugih evropskih državah in Združenih državah Amerike.

Trenutno je arhiviranih 5.5 milijonov preiskav ali 11 PB (11.000 TB) podatkov. Za shranjevanje se uporabljajo tračne knjižnice Oracle StorageTek SL3000 z osmimi LTO6 tračnimi enotami in 1500 predali za trakove v šestih računskih centrih na različnih lokacijah.

Za zapis take količine podatkov bi bilo potrebno 13,800 tračnih kaset tipa LTO 4 z zmogljivostjo 800 GB. Zato datoteke že prepisujejo na tračne kasete LTO 6 s kapaciteto 2,5 TB.

Varovanje osebnih podatkov se zagotavlja s šifrirnimi algoritmi. Podatki se šifrirajo že na strani pošiljatelja, zato tudi pri morebitnem vdoru v računske centre dostop do podatkov ni mogoč.

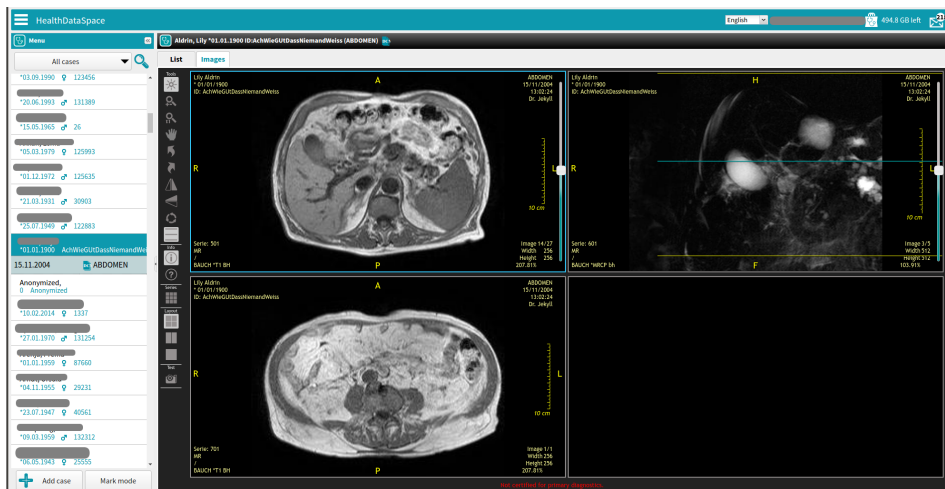
Pravkar poteka razvoj nove različice programske opreme, s katero bo lažje vzdrževati tako veliko količino podatkov.

2.6 Health Data Space

HealthDataSpace (HDS) je storitev, ki omogoča pacientom in pooblaščenim zdravnikom (osebni, lečeči) dostop in pregled medicinskih slik ter drugih zdravstvenih dokumentov v obliki medicinskega oblaka.

Pacient je pravno lastnik svojih medicinskih podatkov. HDS omogoča zbiranje datotek iz vseh bolnišnic in ambulant, ki uporabljajo ta sistem. Vsak pacient ima dostop do vseh svojih medicinskih preiskav in diagnoz, ki jih lahko deli z drugimi. Na ta način se je mogoče izogniti marsikateremu nepotrebnemu posegu, če pacient zdravniku lahko predloži vse predhodne preiskave in pridobi dodatno drugo mnenje.

HDS šifrira datoteke s kombinacijo simetričnega algoritma AES in asimetrične metode RSA. Ta kombinacija zagotavlja visok nivo varnosti in hkrati omogoča enostavno deljenje podatkov z drugimi člani storitve HDS.



Slika 2: Spletni uporabniški vmesnik HDS

2.7 Uporaba centralnega arhiva za drugo mnenje in raziskave

V načrtu je povezava arhivskih podatkovnih shramb s platformo Health Data Space. Tako bi vsak pacient dobil dostop tudi do podatkov, ki so jih bolnišnice arhivirale že pred njegovim pristopom v HDS.

HDS omogoča anonimizacijo medicinskih slik pri prenosu na uporabnikov računalnik ali prenosno napravo.

To odpira možnost dodatnih storitev, saj za dostop do anonimiziranih podatkov ne veljajo enaka pravila kot za varovanje osebnih podatkov. V nekaterih državah se mora pacient izrecno strinjati, da se njegovi anonimizirani podatki lahko uporabijo (t.i. opt-in), drugje pa morajo uporabo izrecno prepovedati (t.i. opt-out). Slovenski zakon o arhivskem gradivu v zdravstvu (ZAGOPP) omogoča dostop do arhivskih podatkov ob pisnem soglasju pacienta ali če iz podatkov ni mogoče razbrati identitete posameznega pacienta.

Primeri dodatnih storitev sta uporaba anonimiziranih podatkov za izdelavo dodatnih pogledov s pomočjo programske opreme (npr. 3D rekonstrukcije) in priprava drugega mnenja.

Ogromna zbirka medicinskih podatkov, ki pokriva široko geografsko področje in več deset letni časovni interval, je primerna tudi kot vir podatkov za raziskave ali strojno učenje z algoritmi, katerih rezultati bodo lahko služili kot pripomoček pri zgodnjem diagnosticiranju določenih bolezni.

HDS je kot varna platforma zanimiv za mnogo zagonskih podjetij s področja varovanja zdravja, ki ponujajo rešitve za mobilne naprave, ki zahtevajo varno hranjenje in izmenjavo podatkov, zbranih s pomočjo posebnih tipal.

3. ZAKLJUČEK

Zakonodaja in predpisi za hranjenje medicinskih slik in pripadajočih dokumentov v digitalni obliki postavljajo nove zahteve za ambulante in klinike ter priložnosti za storitve in platforme, ki lahko ustvarijo novo dodano vrednost na osnovi velike količine digitalnih podatkov.

LITERATURA

1. Gašper Šmid, Žarko Štrumbi (2015), Dostopnost in nedostopnost zdravstvene dokumentacije, *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja, Radenci 2015*.
2. Zakon o zbirkah podatkov s področja zdravstvenega varstva (ZZPPZ), *Uradni list RS, št. 65/00, Uradni list RS, št. 47/15*
3. Zakon o arhivskem gradivu, ki vsebuje osebne podatke o zdravljenju pacienta (ZAGOPP), *Uradni list RS, št. 85/16*
4. https://en.wikipedia.org/wiki/Linear_Tape_Open
Informacije o tračni tehnologiji Linear Tape Open

5. <https://www.telepaxx.com/>, <https://healthdataspace.org/>
Spletni strani ponudnika arhiviranja medicinskih podatkov

STROJNI VID V INDUSTRIJSKEM OKOLJU – IZZIVI IN PRILOŽNOSTI

Matej Logar

Kolektor Orodjarna d.o.o. – PC Vision

E-pošta: matej.logar@kolektor.com

URL: <http://www.kolektorvision.com>

POVZETEK: *Najpogostejši namen aplikacij strojnega vida v industriji je zmanjševanje stroškov končne kontrole. Poleg slednjega investicije v sisteme strojnega vida prinašajo dodatne pozitivne učinke, kot so zmanjševanje stroška reklamacij, izboljšanje kakovosti izdelkov, realno-časovno spremljanje trendov proizvodnje, posredno izboljšanje tehnoloških procesov in zmanjšanje izmeta, ter višanje produktivnosti.*

1. UVOD

Proizvodna podjetja so zaradi pritiskov konkurence primorana iskati rešitve, ki višajo kakovost njihovih izdelkov, nižajo lastno ceno in povečujejo fleksibilnost oz. prilagodljivost potrebam kupca. Trend vpeljave vitkih orodij v proizvodne procese sili uporabnike k iskanju novih, inovativnih principov izboljšave proizvodnih procesov. To dosegajo z vpeljavo različnih protokolov delovanja, spremljanja in reguliranja procesov in vlaganji v avtomatizacijo procesov. Eden izmed segmentov za avtomatizacijo je tudi očna kontrola izdelkov, ki se največkrat nadomešča z aplikacijami strojnega vida.

Najpogostejši namen aplikacij strojnega vida v industriji je optimizacija oz. zmanjševanje stroškov končne kontrole in reklamacij. Poleg slednjih obstaja še množica drugih pozitivnih učinkov, ki dodatno upravičujejo smiselnost investicije in so pogosto spregledani [3].

2. POZITIVNI UČINKI UPORABE STROJNEGA VIDA

2.1 Znižanje stroškov končne kontrole

Prihranki na račun zniževanja stroškov končne kontrole so v praksi najbolj izraziti in hitro dosegljivi. Ti se določijo na podlagi prihranka, ki ga uporabniku prinese zmanjšanje stroška očnega pregledovanja na račun avtomatizacije očne kontrole. Višji kot je taktni čas, več napak kot mora delavec prepoznati in dražja kot je delavna sila, večji je strošek očnega pregledovanja.

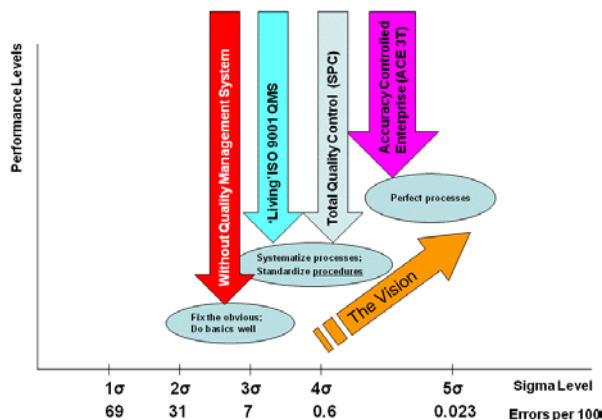
Na drugi strani ekonomske upravičenosti investicije sta nabavna cena optičnega sistema in operativni stroški. Zaznati je mogoče tesno povezavo cene optičnega sistema s številom optičnih glav tj. sestava kamere, objektiva in svetilnega vira. Povezava izhaja iz tega, da je potrebno za vsak dodaten »pogled kamere« zagotoviti ustrezno mehaniko, manipulacijo, opraviti inštalacijska in montažerska dela, izdelati aplikacijo ter poskrbeti za zagon in optimizacijo delovanja. Cena se od začetne investicije naprej dokaj linearno povečuje s številom optičnih glav.

Operativni stroški strojev izhajajo iz naslova porabe energentov in vzdrževalnih del. Priključne moči so <3kW, zato je strošek porabe električne energije zanemarljiv. Vzdrževalni stroški so v rangi <0,5 % nabavne vrednosti. Redna vzdrževalna dela so redka in enostavna (čiščenje, po potrebi avtomatska kalibracija, praznjenje trdih diskov).

Skupni stroški investicije se gibljejo od nekaj tisoč evrov za sisteme z enim pogledom, pa vse do nekaj sto tisoč evrov. Kljub na prvi pogled visokim cenam, dosegajo optični sistemi relativno kratke časa povračila investicije. Ti se gibljejo v skrajnih primerih od nekaj mesecev pa do treh let.

2.2 Eliminacija subjektivnih odločevalcev, posledično izboljšanje kakovosti izdelkov in znižanje stroškov reklamacij s strani kupca

Drugi večji razlog za upravičenost investicije v avtomatsko optično kontrolo je eliminacija subjektivnih odločevalcev. Ljudje delamo napake. Študije kažejo, da človek v dobro vodenem okolju z uporabo klasičnih metod nadzora kvalitete na sto enakih priložnosti za odločitev naredi do deset napak. Če poskrbimo za procese standardizacije, se vrednost lahko zniža na rang 0,5 napak na sto enakih priložnosti za odločitev. Do napak pri odločitvah pogosto prihaja na t.i. mejnih kosih. To so tisti izdelki, na katerih so napake na meji toleranc sprejemljivosti. Človeška napaka je odvisna od fizičnih dejavnikov, organizacijskih, osebnostnih, izkušenosti, šolanja, stresa in počutja [2].



Slika 2: Število napak v odvisnosti od standarda kakovosti

Dobavitelj, ki dobavi izdelek s kritično napako in ga kupec identificira, krije nastale stroške in dobavi nadomestno količino izdelkov. Stroški iz tega naslova so lahko visoki. Optični sistem v tem primeru deluje kot varovalka. Meje sprejemljivost za posamezni detekcijski parameter so nastavljene tako, da preprečijo uhajanje slabih izdelkov. Mejni kosi so navadno razporejeni med slabe. Izmet je višji kot pričakovano, a je del le-tega lažen. Strošek očnega pregledovanja lažnega izmeta je navadno nižji kot strošek reklamacij, kar dodatno upravičuje investicijo.

2.3 Realno-časovno spremljanje trendov proizvodnje, posredno izboljšanje tehnoloških procesov in zmanjšanje izmeta

Ločimo dva tipa integracij aplikacij strojnega vida na proizvodno linijo. Prvi predvideva postavitev stroja za optično kontrolo na konec linije, kjer se kontrolirajo končani izdelki. Ti stroji so navadno kompleksnejši, z večjim številom različnih tipov optičnih glav. Pri drugem tipu integracije so optični sistemi razporejeni med procesi. Slednji so običajno enostavnejši, z eno optično glavo in omogočajo zaznavanje napak takoj za tem, ko se zgodijo. Optični sistemi ob postavitvi takoj za procesom omogočajo hitro ukrepanje ob zaznanju večjega števila ponavljajočih napak oz. slabega trenda. Optični sistem opozori uporabnika hitreje in s tem zniža nastalo škodo, s tem preprečimo izdelavo večje serije slabih kosov ali nadgradnjo slabega polizdelka na procesih, ki sledijo. Višina prihranka iz naslova realno-časovnega spremljanja trenda proizvodnje je povezana s ceno samega izdelka in stroški procesov. Dražji kot je izdelek in/ali proces bolj je investicija smiselna.

Dodatno nam medfazne kontrole omogočajo pridobivanje povratne informacije o učinkovitosti posameznega proizvodnega procesa. Vsak optični sistem je merilna naprava, zato so ti sistemi lahko vir informacij za izboljšanje oz. optimizacijo procesa in zmanjšanje izmeta. Predvideva se, da bo v prihodnosti močan poudarek na razvoju programskih rešitev, ki bodo povezovale procesne informacijske sisteme (angl. MES) in sisteme strojnega vida, kar sovпада s trendom industrije 4.0.

Povratne informacije se lahko koristijo tudi za ustvarjanje povratne zanke v smislu avtomatske regulacije samega procesa. Omenjena funkcionalnost se danes najpogosteje uporablja pri aplikacijah manipulacije izdelkov, varjenj in sestavljanj izdelkov (robot-kamera), vendar še ni povsem izkoriščena.

2.4 Višja produktivnost

Taktni časi proizvodnih linij se med panogami razlikujejo. V prehrabni in embalažni industriji so ti časi najvišji in lahko dosegajo do nekaj deset izdelkov na sekundo. V avtomobilski industriji in proizvodnji gospodinjskih aparatov so ti časi od nekaj sekund pa do nekaj minut. Očna optična kontrola lahko predstavlja ozko grlo linije. Avtomatizacija in višanje taktnih časov te operacije vpliva na produktivnost.

Uporabljena oprema v industrijskih aplikacijah se neprestano izboljšuje. Zmogljivejši procesorji, kamere z višjim številom zajemov slik na sekundo, hitrejši komunikacijski protokoli, hitrejši algoritmi, vse to pozitivno vpliva na višanje produktivnosti. V splošnem lahko rečemo, da so sami optični sistemi hitri, saj se časi zajema slike in obdelave gibljejo od nekaj deset pa do nekaj sto mili sekund. Sistemi za manipulacijo izdelkov imajo navadno daljše taktne čase, ti se gibljejo od pol sekunde pa do nekaj sekund. Višje taktne čase navadno dosegamo z multiplikacijo manipulacijskih sistemov, kar faktorsko vpliva na ceno sistema. V kolikor točno poziciranje izdelka ni potrebno, manipulacija ne predstavlja ozkega grla.

Zmogljivosti opreme v današnjem času so take, da lahko en stroj za optično kontrolo kontrolira izdelke več linij, kar direktno vpliva na višjo produktivnost.

2.5 Manjša delovna površina in zmanjšanje medfaznih zalog

Stroji za optično kontrolo navadno zasedajo nekajkrat manjšo delavno površino, kot jo zasedejo delavna mesta za očno kontrolo. Ta lastnost se pogosto zanematja, lahko pa prihrani nekaj tisoč evrov letno.

Principi vitke proizvodnje učijo, da so vse zaloge strošek. V splošnem velja, da nižje kot so zaloge, bolj agilen, vitek je sistem in višji so prihranki. Zaradi kompaktnosti in delovanja v liniji imajo stroji za optično kontrolo nizko medfazno zalogo. Medfazne zaloge nastajajo v transportu med procesom pred strojem in procesom (navadno pakiranjem) za strojem. Ta je navadno linijski, zato so te zaloge nizke. Dodatne zaloge nastajajo v strojih, ki uporabljajo manipulacijske sisteme, ki z namenom ohranjajo taktni čas. Primer so vrtljive mize z gnezdi.

Ne glede na naštetu so medfazne zaloge nizke. Navadno govorimo o nekaj deset kosih in so zanemarljive v primerjavi z zalogami, ki nastajajo pri očni kontroli. Pri tem tipu delovanja se lahko med procesi hrani nekaj sto ali celo tisoč izdelkov.



Slika 1: Stroj za optično kontrolo komutatorjev zasega 9 krat manjšo površino kot delavna mesta za očno kontrolo

3. ZAKLJUČEK

Obstaja množica pozitivnih učinkov avtomatizirane optične kontrole, ki jih uporabniki na žalost še ne izkoriščajo povsem ali ne poznajo. Namen prispevka je seznaniti bralca o široki paleti priložnosti, ki lahko znatno vplivajo na izkoriščenost opreme, optimizacijo proizvodnje in kot rezultat nižjo lastno ceno izdelka.

Uporabniki hitro opazijo in/ali izkoristijo t.i. direktne učinke, ki so povezani z umestitvijo stroja v linjo. Tak primer so znižani stroški končne kontrole ali nižanje stroškov reklamacij. Druge, ki zahtevajo dodatno statistično analizo, optimiranje procesov, znižanje izmeta in predvsem večjo zavzetost, pa redko izkoristijo. Te je mogoče spodbuditi s filozofijo stalnega napredka, ki meri napredek in nagrajuje zaposlene z dodatki za inovativnost in izboljšave.

LITERATURA

1. http://www.visiononline.org/vision-resources-details.cfm/vision-resources/Cost-Justification-Strategies-for-Machine-Vision/content_id/1330
Cost Justification Strategies for Machine Vision by Nello Zuech, Contributing Editor - AIA
2. http://www.lifetime-reliability.com/cms/tutorials/reliability-engineering/human_error_rate_table_insights/
Mike Sondalini, Lifetime Reliability Solutions
3. <http://www.cognex.com/what-is/machine-vision/benefits/?id=17627&langtype=1033>

ROBUSTNOST INDUSTRIJSKIH SISTEMOV STROJNEGA VIDA

Jure Škrabar

Kolektor Orodjarna d.o.o. – PC Vision

E-pošta: jure.skrabar@kolektor.com

URL: <http://www.kolektorvision.com>

POVZETEK: Napredek na področju avtomatizacije močno vpliva na izgled industrijskih proizvodnih procesov. Napredek kaže predvsem na povečanih kapacitetah procesov brez padca kvalitete končnega izdelka. Industrijski strojni vid je tehnologija, ki omogoča spremljanje in nadzor nad kvaliteto izdelkov in posledično proizvodnih procesov, saj je zanesljiv, objektivni in v večini primerov v celoti avtomatiziran. V procesu izgradnje je za kvalitetno zasnovo in izvedbo sistema industrijskega strojnega vida pomembno sodelovanje obeh strani, tako naročnika kot dobavitelja. S strani naročnika je pomembna kvalitetna priprava zahtevnika, kot ga narekuje standard VDI/VDE 2632, s strani izvajalca pa specifikacija sistema in prevzemnih kriterijev, vse pa se usklajuje tudi s kupcem. Končni cilj vpletenih v poslovni odnos je kvalitetno izveden sistem industrijskega strojnega vida, ki bo v celotnem življenjskem ciklu izdelka sposoben pokriti variacije proizvodnih procesov ter zagotavljal objektivno oceno kvalitete izdelkov.

1. UVOD

Robustnost je ključna odlika kvalitetno zasnovanih industrijskih sistemov strojnega vida. Sistem mora ustrezati zahtevanim lastnostim v celotni življenjski dobi izdelka, kjer bo prihajalo do variacij vhodnega materiala, dimenzij, lastnosti površine, barv, staranja vgrajenih komponent ter s tem tudi pogojev osvetljevanja, okolja obratovanja in sprememb geometrije samega izdelka. Robusten industrijski sistem strojnega vida mora biti neobčutljiv na takšne spremembe, kar lahko dosežemo, če pričakovane variacije lastnosti izdelkov upoštevamo pri zasnovi industrijskega sistema strojnega vida.

V prispevku so predstavljene smernice za gradnjo sistemov z vidika standarda VDI/VDE 2632 ter potrebni koraki za uspešno zaključeno integracijo v industrijsko okolje. Naslovljene so različne faze izgradnje sistemov od uvodnega zahtevnika do definicije prevzemnih kriterijev.

2. PRIPRAVA ZAHTEVNKA ZA IZVEDBO INDUSTRIJSKEGA SISTEMA STROJNEGA VIDA

Definicija industrijskega sistema strojnega vida po standardu VDI/VDE 2632 2. del se deli na dva sklopa: 1) specifikacija zahtev naročnika in 2) specifikacija sistema s strani

izvajalca/dobavitelja [1]. Vsebine v nadaljevanju so do določene mere opisane tudi v standardu, vendar jih je smiselno izpostaviti.

2.1 Priprava zahtevnika za kontrolo

Specifikacija zahtev naročnika, pogosto imenovana tudi »zahtevnik za kontrolo« mora vsebovati vse informacije za kvalitetno zasnovo. Namen standarda je, da naročnik posreduje dobavitelju vse ključne informacije v pravilni obliki. Vsaka najmanjša podrobnost, ki se spregleda v tej fazi, ima lahko nepredvidljive negativne posledice na robustnost končne aplikacije. Vplivni dejavniki se delijo na nalogo, predmet kontrole, sceno oziroma okolje obratovanja, način integracije v proces, vmesnik človek-stroj in ostale dejavnike, ki jih ni možno umestiti v prej zapisane točke.

Z definicijo naloge se grobo definirajo področja in metode, s katerimi bo dobavitelj pristopil k izvedbi. Pomembne so vse informacije, ki jih lahko podamo, kot so tip meritve, namen preizkušanja z opisom naloge, na kakšen način se problematika rešuje v trenutnem stanju, predvidi se statistične karakteristike in način evalvacije meritev ter nenazadnje tudi časovne zahteve in zahteve po razpoložljivosti. Iz stališča dobavitelja so tu najbolj pomembne točke predvsem podroben opis naloge s specifikacijo napak in tolerancami letih.

2.1.1 Ključni parametri za kontrolo in pomen standardov

Identifikacija parametrov za kontrolo je ključna točka pri pripravi kvalitetnega zahtevnika za kontrolo, kjer je potrebno narediti selekcijo med parametri, ki so po eni strani pomembni s stališča zahtev naročnika in na katere je možen tehnološki vpliv ter parametri, ki so posredne narave. Primer posrednih meritev so lokacije izvrtin pri izdelavi v enem vpetju na CNC obdelovalnem stroju.

Pri dimenzijskih meritvah je ta korak dokaj enostaven, saj se večino informacij pridobi iz tehnične dokumentacije izdelka. Pomemben faktor pri definiciji meritve ter kasnejši oceni izvedljivosti so reference, na katerih temelji meritev. Dokler so reference lokalno deterministične, t.j. definirane in vidne v enakem pogledu kot merjeni parameter, so meritve v večini primerov relativno enostavne s stališča referenc. Ko pa se meritve opirajo na bazne koordinate, kar je pogosto predvsem pri večjih izdelkih z večimi značilkami, pa kompleksnost izvedbe meritev narašča.

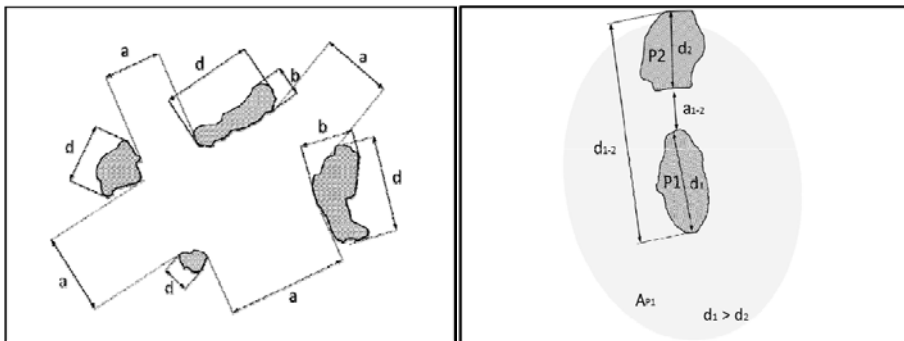
Površinske kontrole oz. identifikacija napak na površini izdelka je v industrijskem strojnem vidu izredno zahtevno področje aplikacij. Naročniki končnih izdelkov do neke mere še tolerirajo manjše oz. manj vidne poškodbe, če le-te ne vplivajo na funkcijo izdelka. V večini primerov pri površinskih napakah prihaja do variacij pojavnih oblik in lokacij napak v kombinaciji z ne-determinističnimi strukturiranimi površinami. Velik doprinos k objektivni oceni površinskih napak v večini primerov pride s strani t.i. ekspertov oz. ekspertnega znanja. Zaposleni v končnih kontrolah dnevno pregledajo nekaj deset do nekaj tisoč kosov (odvisno od tipa) in se lahko njihovo, sicer subjektivno znanje, če je na pravi način usmerjeno, izkoristi za pripravo objektivnih kriterijev. Realni primer dela v končni kontroli z namenom pridobitve objektivne ocene o kvaliteti mejnega kosa

(OK / NOK) v različnih izmenah. Dober kos jim je bil predstavljen kot slab in obratno, rezultati so prikazani v tabeli 1.

Tabela 1: Ocena kvalitete površine kosov v končni kontroli izdelkov

KOS	REALNO STANJE KOSA	PREDSTAVLJENO STANJE KOSA	IZMENA 1	IZMENA 2	IZMENA 3
1	OK	NOK	OK	OK	NOK
2	NOK	OK	OK	NOK	NOK
3	NOK	NOK	NOK	OK	NOK

Iz vplivnih dejavnikov na kvaliteto je potrebno izločiti subjektivno oceno, kar je tudi eden izmed razlogov za vpeljavo industrijskih sistemov strojnega vida v proizvodni proces. Pri pripravi objektivnih kriterijev se, če je le-to mogoče, v večji meri opiramo na zahteve naročnikov končnega izdelka in veljavne standarde. Kot primer iz avtomobilske industrije je Volkswagen standard 50097 oziroma ekvivalentni Daimler-Benz DBL 4949, ki sta napisana glede na avtomatizirano kontrolo z industrijskim strojnim vidom [3] in opisujeta dopustni velikostni razred ter porazdelitev lunckerjev na obdelanih površinah ulitkov za motorne sklope.



Slika 1: Definicija velikosti lunckerja ter najmanjše razdalje med njimi (levo), definicija eliptičnega območja združevanja večih manjših napak v eno (desno) po standardu DBL 4949 [2]

2.1.2 Okolje obratovanja industrijskega sistema strojnega vida

V kolikor okolje obratovanja ni upoštevano pri zasnovi industrijskega strojnega vida, lahko ima le-to izrazito negativen učinek na stabilno delovanje. Standard VDI/VDE 2632 2. del posebej izpostavlja vplivne faktorje okolice, kot so pozicioniranje, ozadje z refleksijami, vibracije in oscilacije, zunanja svetloba, ozračje s tujki, temperatura in elektromagnetno sevanje. V nadaljevanju sta izpostavljena dva ključna vplivna dejavnika.

Nihanje temperature okolice ter temperatura preizkušane kosa med obratovanjem ima neposredni vpliv na dimenzijske meritve. V kolikor je kos enostaven in je temperatura enakomerno porazdeljena po izdelku, je teoretično možno prispevek termičnih raztezkov in skrčkov linearno upoštevati pri izvedbi dimenzijskih meritev. Zaradi zaostalih notranjih napetosti in generalno kompleksnih geometrij realnih izdelkov pa so raztezki in skrčki nelinearni, kar močno oteži oziroma onemogoči implementacijo korekcij glede na temperaturo.

Konstantna osvetlitev izdelkov v življenjski dobi sistema je ena ključnih predpostavk, na katerih so grajeni robustni sistemi industrijskega strojnega vida. Predvsem je pomembna pri zajemu informacij v barvnem spektru. Zavedati se je potrebno razlike med fotometrično in radiometrično percepcijo barv. Sistemi strojnega vida niso sposobni zaznavati barv in izvajati primerjave z barvnimi lestvicami in kartami direktno, saj so le-te neodvisne od osvetljenosti površine. Zaznavanje in primerjava barv je izvedljiva posredno preko primerjalnih metod in kalibracij. V sistemih strojnega vida so merjeni energijski nivoji različnih valovnih dolžin oziroma radiometrični energijski nivoji, na katere ima vpliv tudi osvetlitev kosa in okolice, ki je v življenjski dobi sistema podvrženo variacijam [4].

2.1.3 Pozicioniranje izdelkov za kontrolo

Pozicioniranje izdelkov ima pomembno vlogo v robustnih sistemih. Najprej mora biti leto ponovljivo, predvsem v aplikacijah z uporabo 2D metod, kjer so definirane reference v ravnini pogleda ne pa njihova lokacija v globini. Upoštevati je potrebno točkovni izvor pogleda pri uporabi klasične optike ter povezane optične deformacije. Pri ozkih tolerančnih območjih lahko najmanjši nagib kosa povzroči nepravilno zaznavo značilke, čemur se lahko ognemo samo z dodatnimi zaznavami referenčne površine.

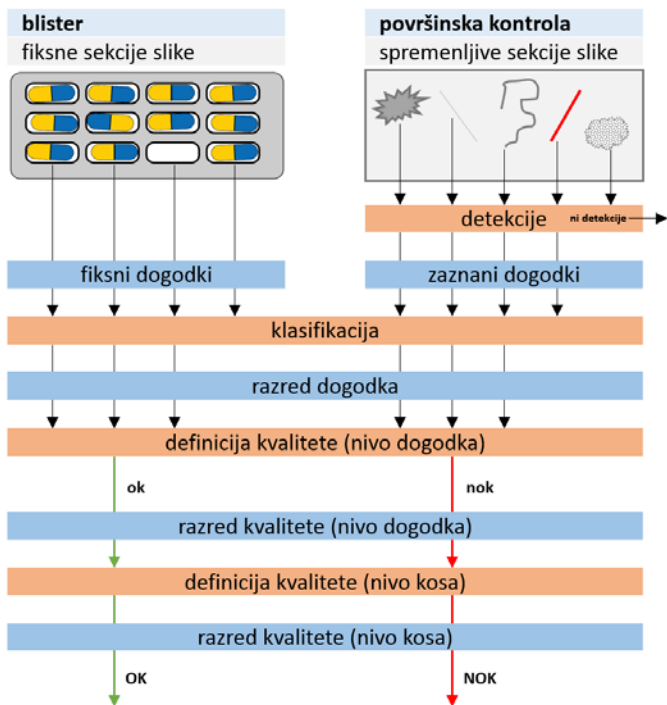
3. DEFINICIJA PREVZEMNIH KRITERIJEV

Kvantitativna ocena sposobnosti merilnega procesa (ne-klasifikacijska) je v industriji že vzpostavljena preko različnih standardov (DIN 1319, JCGM 100), katerih ključni parameter za oceno je merilna negotovost. Validacijski proces poteka preko t.i. MSA kriterija (ang. Measurement System Analysis), s katerim statistično ocenjujemo tudi izvor variacij meritev. Klasifikacijska sposobnost merilnega procesa se ozira na atributivne spremenljivke [2].

Način prevzemanja s prevzemnim preizkusom mora biti definiran znotraj zahtevnika za izvedbo meritev ter specifikacij s strani dobavitelja [1]. Izvedba prevzemnega preizkusa na vzorčnih kosih ali na dejanski serijski proizvodnji je stvar dogovora med naročnikom in izvajalcem sistema industrijskega strojnega vida. Industrijski sistem strojnega vida v dogovorjeni meri pravilno zazna nepravilnosti na kosih tekom prevzemnega preizkusa, robusten sistem pa v celotni življenjski dobi stroja.

3.1 Nadzor kvalitete z atributivnimi parametri

Na predmetu kontrole so v večini primerov pomembni samo manjši segmenti, za katere je potrebno definirati atributivne parametre (tip zaznane nepravilnosti). Na sliki 2 je predstavljen mehanizem klasifikacije zaznanih nepravilnosti za dva različna primera, kjer je prvi primer blister s fiksnimi sekcijami slike. V primerjavi s površinsko kontrolo (drugi primer) le-ta ne potrebuje dodatnega koraka »detekcije«, saj so področja zanimanja vnaprej definirana (»dogodek«).



Slika 2: Proces nadzora kvalitete za atributivne parametre, kot ga navaja standard VDI/VDE 2632 3. del [2]

V zadnjem koraku (slika 2) se definira kvaliteta kosa, ki vsebuje definicije kvalitete vseh dogodkov na predmetu kontrole [2]. Zavedati se je potrebno, da obstaja možnost, da se pojavi nova ali drugačna pojavna oblika površinske napake in jo sistem ne prepozna (zgrešen dogodek). Izločanje slabih kosov torej ni 100 %, a se temu lahko približamo na sprejemljivo raven za končnega kupca izdelkov. Potreben je kompromis med stopnjo lažnega izmeta dobrih izdelkov ter prepustnostjo slabih izdelkov skozi proces kontrole.

3.2 Prezemni proces z referenčnimi kosi

Referenčni kosi morajo biti določeni v uvodnih fazah projekta oziroma že znotraj zahtevnika za kontrolo. V praksi se sposobnost sistema ocenjuje na dovolj velikem naboru prevzemnih kosov, da lahko s statističnimi metoda ocenimo delovanje. Z večanjem vzorca

se sicer večja zanesljivost statističnih rezultatov, a se s tem tudi večja vloženo delo v pripravo in evalvacijo kosov. Vseh pojavnih oblik nepravilnosti na kosu z vzorci ni mogoče pokriti.

3.3 Prezemni proces med serijsko proizvodnjo

Prednost prevzemnega preizkusa med serijsko proizvodnjo je v ocenjevanju sposobnosti v pogojih realne proizvodnje. Zavedati se je potrebno, da se določene nepravilnosti na izdelkih pojavljajo naključno ali v posebnih pogojih, lahko pa nastanejo nepravilnosti v proizvodnem procesu. Zastopanost vzorcev z napakami znotraj časovnega obdobja prevzemnega preizkusa je zato lahko majhna. Pomembna je sledljivost vsakega kosa, saj je potrebna ekspertna ocena kvalitete vsakega kontroliranega kosa. Prevzemni proces med serijsko proizvodnjo se v praksi ne izvaja, razen v izrednih primerih, kjer ni mogoče drugače definirati prevzemnih pogojev. Tipični primer so aplikacije za vodenje procesov.

4. ZAKLJUČEK

Standardizacija procesa zasnove s strani naročnika in dobavitelja ter končne validacije sistemov je bila vzpostavljena z namenom kvalitetnega pretoka informacij ob pravem času. Iz sicer navidezno nepovezanih informacij, na katere nas opozarja standard, je možno zagotoviti robustno delovanje sistema v celotnem življenjskem obdobju. Robustni industrijski sistemi strojnega vida v praksi obratujejo dlje od predvidenih okvirov za povračilo investicije in dolgoročno ne izkažejo večjih napak, ki bi vplivale na končno odločitev o kvaliteti izdelkov.

Pri prvi integraciji sistemov v industrijsko okolje, kjer tehnologija še ni v uporabi se pogosto naleti na težave povečanega izmeta izdelkov in nestabilnost proizvodnega procesa. V kolikor se pridobljene informacije izkoristijo na pravi način, se lahko vpliva na celoten proizvodni proces ter s tem večja produktivnost, manjša izmet in stroške proizvodnje.

4. LITERATURA

1. VDI/VDE/VDMA 2632 – part 2 (2015), Machine vision – Guideline for the preparation of a requirement specification and a system specification
2. VDI/VDE/VDMA 2632 – part 3 (2016), Machine vision/industrial image processing – Acceptance test of classifying machine vision systems
3. DBL 4949 (2014), Standardporenklassen für Gussteile, Mercedes-Benz Werknorm
4. B. Dawson (2016), Advanced Color Machine Vision and Applications, *Vision Conference Boston, MA*

TRENDI NA PODROČJU STROJNEGA VIDA

Samo Šela

Kolektor Orodjarna d. o. o., PE Vision

E-pošta: samo.sela@kolektor.com

URL: <http://www.kolektorvision.com>

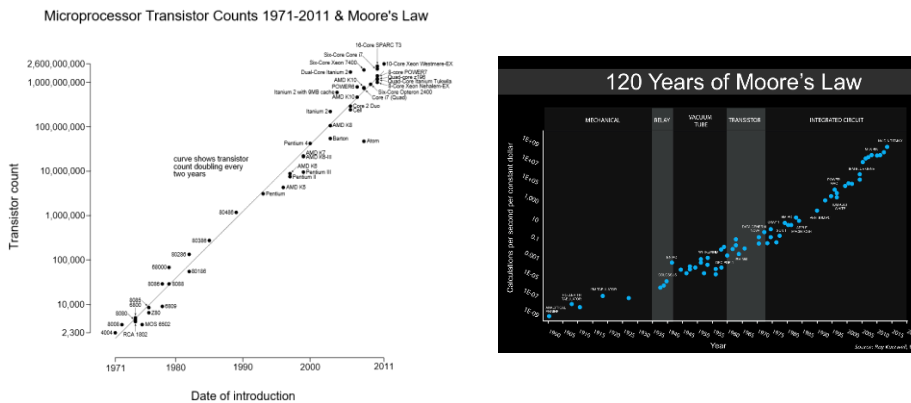
POVZETEK: *Sodeč po Gartnerjevem ciklu popularnosti novih tehnologij so trenutno najbolj aktualne teme s področja strojnega vida: nastajanje različnih platform interneta stvari, pametni roboti, strojno oziroma globoko učenje. Skladno s tem je mogoče opaziti, da med tradicionalne igralce na področju industrijskega strojnega vida čedalje bolj vstopajo nizkocenovne vgradne naprave, ki nadomeščajo drage pametne kamere; napredek na področju 3D tehnologij se kaže v vse bolj uspešnih aplikativnih primerih robotskega pobiranja neurejenih izdelkov, s pospešenim razvojem metod globokega učenja pa je vse hitrejši tudi prodor multispektralnih in hiperspektralnih kamer v industrijsko rabo za analizo sestave različnih materialov. V podjetju Kolektor se zavedamo, kateri bodo potrebni dejavniki za uspešnost ponudnikov rešitev strojnega vida v četrti industrijski revoluciji. Odgovore na vprašanja bomo iskali znotraj projekta GOSTOP (Gradniki, Orodja in Sistemi za Tovarne Prihodnosti), ki je del evropskega programa za raziskave in razvoj Obzorje 2020.*

1 UVOD

Leta 2010 se je skupina neodvisnih strokovnjakov s področja znanosti in tehnologije, ki svetuje predsedniku Združenih držav Amerike, dotaknila napredka na področju programske opreme [1]. Trditev, da je napredek na področju algoritmov prekosil Moorov zakon (slika 1), ki govori o podvajanju števila komponent na integriranem vezju na vsaki dve leti, je velikokrat vzeta iz konteksta. V resnici sta napredek na področju strojne in programske opreme močno prepletena. Dober primer takšnega napredka je razvoj principov in metod za globoko učenje. Razvoj učinkovitih algoritmov za učenje na zmogljivih grafičnih karticah je hkrati prinesel razvoj namenske strojne opreme, ki je specializirana prav za to problemsko domeno.

Z napredkom in trendi na področju strojnega vida kot ključne omogočitvene tehnologije industrije 4.0 se je ukvarjal Skvarč v delu [3]. Izgubi industrijske dodane vrednosti v vrednostni verigi je sledil odgovor Evrope z obsežnim investicijskim programom katere del je prav tako projekt GOSTOP. Internet stvari (angl., IoT – *internet of things*) je eden izmed znanilcev prihoda četrte industrijske revolucije. Vsesplošna povezljivost med realnim in virtualnim svetom bo omogočala izjemno fleksibilnost proizvodnje v ne tako

oddaljeni prihodnosti. Na valu vsesplošne povezljivosti bo sledil razvoj novih poslovnih modelov. Skvarč osvetli trende na področju strojnega vida skozi tri smiselno povezana področja: optične in električne gradnike, programsko opremo in uporabniško izkušnjo ter nove in napredne primere uporabe.



Slika 1: Graf Moorovega zakon, ki prikazuje podvajanje števila komponent integriranega vezja na vsaki dve leti (levo), in prirejena verzija, ki prikazuje število kalkulacij na sekundo pri konstantni vrednosti ameriškega dolarja, katere avtor je futurist in eden izmed vodilnih znanstvenikov na področju umetne inteligence pri podjetju Google, Ray Kurzweil (desno). Zadnjih nekaj točk na grafu pripada grafičnim procesnim enotam podjetja Nvidia [2].

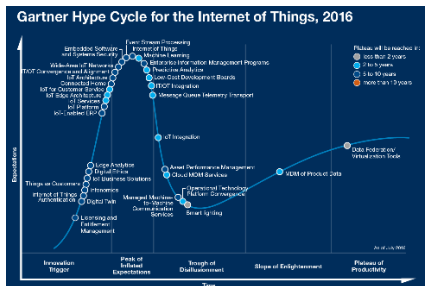
2 TRENDI

Gartnerjev cikel popularnosti novih tehnologij opisuje različne faze razvoja posamezne tehnologije [4, 5]. Kljub medijski pozornosti ima prihod nove tehnologije na trg le redko že pravo aplikativno vrednost. Ko tehnologija doseže vrh pričakovanja z medijskimi zgodbami o uspehu in neuspehu, sledi upad popularnosti, ki je bolj kot ne faza spoznanja, da ni vse zlato, kar se svetli. Temu sledi nadaljnji razvoj tehnologije, ki ga preživijo le resnično uspešne, ki na koncu preidejo v širšo praktično uporabo.

Računalniški oziroma strojni vid je močno prepleten z nekaterimi porajajočimi tehnologijami, ki jih lahko vidimo na grafih na sliki 2. Med tistimi, ki so še v zgodnji fazi, opazimo med drugim brezpilotna letala, naprave IoT, pametne robote, strojno učenje in avtonomna vozila.

V naslednjih desetih letih lahko pričakujemo nadaljevanje povečevanja računske moči, raznovrstne generatorje neskončne množice podatkov, nepredstavljen napredek na področju globokega učenja. Doba pametnih naprav bo povsem spremenila področje tehnologije strojnega vida. V ospredje bodo prišla podjetja s pametnimi napravami za zajem vizualnih in drugih podatkov in izkoriščanje teh za izboljševanje lastnih procesov

ali nudenje storitev zainteresiranim kupcem. Porajajoče tehnologije bodo tako na valu interneta stvari na novo definirale koncepte strojnega vida kot ga poznamo danes.



Slika 2: Gartnerjev cikel popularnosti novih tehnologij v letu 2016 (levo), študija cikla popularnosti interneta stvari v letu 2016 (desno) [4, 5].

2.1 Vgradne naprave

Prodor interneta stvari je prinesel tudi pospešen razvoj raznovrstnih cenovno dostopnih vgradnih naprav. Napredek na tem področju se kaže v povezljivosti in zmogljivosti. Na eni strani imamo preproste večfunkcijske senzorje, kot je L-TEK-ov FireFly, na drugi strani zmogljive pametne kamere, kot je Roseekov Woodpecker.

Pomemben faktor pri razvoju novih senzorjev in kamer ter njihove programske opreme so večkrat zanemarjeni uporabniški vmesniki. Uporabniki smo danes vajeni dobro izpiljenih aplikacij na mobilnih telefonih in spletu. Uporabniški vmesniki bodo morali biti lični, preprosti, hkrati pa omogočati intuitiven način za hitro doseganje zelenih rezultatov.

Tradicionalni sistemi strojnega vida so sestavljeni iz para kamera-računalnik. Med seboj morajo biti povezani s hitrimi povezavami z nizkimi latencami. Večanje zmogljivosti pametnih senzorjev bo omogočilo, da se bo vsaj delna, če ne celotna, obdelava in analiza vršila že na sami vgradni napravi. Skladno s tem je ključno, da proizvajalci programske opreme poskrbijo, da lahko njihova programska koda teče na različnih procesorskih arhitekturah (x86, x64, ARM).

Nenazadnje pa je pomembno razumevanje razvoja programske opreme. Kar se je zdelo še pred petimi leti nemogoče, lahko danes praktično opravimo s pametnim mobilnim telefonom.

2.2 Odprta koda

Velik trend je razvoj različnih odprtokodnih programskih rešitev, ki segajo vse od operacijskih sistemov (Ubuntu, Debian, CentOS, Fedora), knjižnic za računalniški vid in robotiko (OpenCV, SimpleCV, ImageMagick, CCV, PCL - Point Cloud Library, ROS -

Robot Operating System) do orodij za strojno in globoko učenje (Weka, Orange, scikit-learn, Tensorflow, Theano, Caffe).

Odprtokodni projekti na eni strani omogočajo lažji vstop novim igralcem na trg z nišnimi produkti, na drugi strani pa omogočajo obstoječim ponudnikom hitrejše in cenejše širitve na še ne osvojene trge. Odprtokodno skupnost in produkte med drugim podpirajo največji igralci v IT sektorju. Gre za podjetja kot so Google, Intel, IBM in Microsoft. Razširjena uporaba odprtokodnih rešitev hkrati zagotavlja, da so to izpiljeni in robustni programski paketi.

2.3 Strojno in globoko učenje

Čedalje hitrejši razvoj metod globokega učenja pomeni odmik od klasičnega pristopa načrtovanja kompleksnih značilnic in uporabe ene izmed standardnih metod strojnega učenja (npr. metode podpornih vektorjev). Na valu zmogljivih grafičnih kartic je v porastu uporaba konvolucijskih nevronske mreže. Te najdemo v mnogih različnih aplikacijah kot so prepoznavna obrazov, zaznavanje in klasifikacija objektov v slikah, nadzor vožnje avtonomnih vozil, ali bolj industrijsko specifičnih kot prepoznavna in klasifikacija napak pri proizvodnji jekla [6]. Že danes lahko na trgu najdemo programske pakete, ki izkoriščajo moč globokega učenja (perClass, Halcon, ViDi Suite, Squeeze brains Retina).

Da so storitve strojnega in globokega učenja na področju računalniškega vida v oblaku prisotne med nami že danes, priča ponudba velikih podjetij preko različnih programskih vmesnikov. Med temi lahko najdemo Amazon (*Machine Learning Rekognition*), Microsoft (*Face API*, *Emotion API*, *Computer Vision API*, *Content moderator*), IBM (*Visual Recognition*) in Google (*Cloud Vision Api*). Podrobnosti o posameznih paketih navajamo v tabeli 1.

Tabela 1: Programska orodja za računalniški vid v oblaku

<i>Programski vmesnik</i>		<i>Opis</i>
Amazon	Machine Learning Rekognition	Servis za analizo slik omogoča prepoznavo objektov, scene, obrazov in slik. Mogoča je tudi primerjava različnih obrazov med seboj.
Microsoft	Face API	Servis omogoča prepoznavo in primerjavo podobnih obrazov, organizacijo ljudi v skupine na podlagi vizualne podobnosti, identifikacijo prej naučenih oseb v novih slikah.
Microsoft	Emotion API	Servis omogoča analizo obrazov na podlagi vrste različnih čustev.
Microsoft	Computer Vision API	Servis omogoča pridobivanje bogatih informacij iz slik za namen kategorizacije in procesiranja vizualnih podatkov. Med drugim omogoča zaščito uporabnikov pred nezaželeno vizualno vsebino.
Microsoft	Content moderator	Servis omogoča avtomatično moderiranje besedila, slik ali video vsebin.

Google Cloud Vision Api	Servis omogoča razumevanje vsebine slik na podlagi zmogljivih modelov strojnega učenja. Mogoča je klasifikacija slik v tisoče različnih kategorij, prepoznava objektov in obrazov v slikah, iskanje in prepoznavanje besedila v slikah.
IBM Visual Recognition	Servis omogoča semantično razumevanje vsebine slik, iskanje človeških obrazov, ocenjevanje starosti in spola, iskanje podobnih slik v podatkovnih zbirkah. Sistem je mogoče naučiti nove koncepte, npr. identifikacija oblačila ali prepoznavanje pokvarjenega sadja v trgovini.

2.4 Poslovni modeli

Napredek na področju programske opreme in razširjenost sistemov strojnega vida v industrijskih okoljih sta ključna faktorja, da je danes mogoče izkoristiti pridobljene podatke za namen učečih se sistemov strojnega vida. V prihodnjih letih bo zanimivo opazovati kako se bodo tradicionalni igralci na področju strojnega vida prilagodili na nove razmere na trgu. Vseprisotnost množice kamer in senzorjev ponuja priložnosti za razvoj novih poslovnih modelov, ki temeljijo predvsem na razvoju novih programskih platform.

Prve korake so na tem področju naredili velikani kot so Amazon, Microsoft, IBM in Google. V prihodnosti bodo sledili tradicionalni in novi ponudniki rešitev strojnega vida kot storitve v oblaku. Množica cenovno dostopnih senzorjev ali prilagoditev obstoječih sistemov bo posredovala podatke na analizo v oblak.

Tipičen sistem optične kontrole zahteva temeljito načrtovanje postavitve kamer in principov osvetljevanja. Prenos kompleksnega in natančno premišljenega sistema bodo v prihodnosti lahko nadomestili preprostejši sistemi, ki se bodo zanašali na napredne storitve strojnega vida v oblaku na podlagi kompleksnih modelov globokega učenja. Ti bodo po zmogljivosti sorodni ali celo prekašali zmogljivost človeškega vida.

Problem pridobivanja podatkov za potrebe nadzorovanega učenja lahko že danes rešujemo s pomočjo platforme *LabelMe*, ki jo je mogoče pogoniti na platformi *Amazon Mechanical Turk* za najem delovne sile. Tako lahko problem anotacije podatkovnih zbirk preprosto prenesemo na množico zunanjih sodelavcev. Prav tako je danes mogoče izkoristiti platforme kot je *Kaggle*, da pridemo do kompleksnih rešitev na področju strojnega in globokega učenja, ki bi sicer zahtevale zaposlitev redke in drage delovne sile. Andrew Ng., vodja pri kitajskem podjetju *Baidu Research*, je dejal, da je danes največji problem pomanjkanje strokovnjakov s področje umetne inteligence.

Podjetja kot so Google, Facebook, Twitter, so že zdavnaj ugotovila, da so podatki ključ do uspeha. Ob prehodu panoge strojnega vida v storitveno in nastanek samo učečih se sistemov bo bolj pomembno lastništvo podatkov kot kakšne metode ima podjetje na voljo.

3 ZAKLJUČEK

Napredek na področju razvoja strojne in programske opreme je nosilec razvoja področja strojnega vida. Na trg vstopajo nizkocenovne vgradne naprave, ki bodo v prihodnosti nadomeščale drage pametne kamere uveljavljenih proizvajalcev. Zmogljivost nekaterih že danes zadošča za srednje zahtevna opravila strojnega vida. Pri tem je ključno, da ponudniki orodja in rešitev razvijajo v smeri prenosljivih programskih platform.

Panoga bo na valu napredka prisiljena razvijati nove poslovne modele. Razširjenost sistemov strojnega vida že danes omogoča zbiranje gore podatkov za razvoj učečih se sistemov. Prvim ponudnikom računalniškega vida v oblaku se bodo v prihodnosti pridružili bolj specializirani servisi strojnega vida.

Pohod napredka v smeri popolne avtomatizacije in robotizacije v različnih industrijskih panogah bo povzročil tudi globoke in daljnosežne družbene spremembe. Oxfordska študija napoveduje izginotje 47% delovnih mest v Združenih državah Amerike v naslednjih 25 letih [7]. Tako kot je družba preživela prihod mehanizacije, masovne produkcije ter dobo računalništva in avtomatizacije, bo morala preživeti neločljivo povezanost kibernetičnega in fizičnega sveta prepletenege z novimi poslovnimi praksami.

4 LITERATURA

1. <https://www.cis.upenn.edu/~mkearns/papers/nitrd.pdf>
Report to the president and congress, Designing a digital future: federally funded research and development in networking and information technology, str. 71.
2. https://en.wikipedia.org/wiki/Moore%27s_law
Moorov zakon
3. J. Skvarč (2016), Strojni vid – ključna tehnologija industrije 4.0, *ROSUS 2016, Računalniška obdelava slik in njena uporaba v Sloveniji 2016*, vol. 11, str. 3-8.
4. <http://www.gartner.com/newsroom/id/3412017>
Gartner's 2016 Hype Cycle for Emerging Technologies Identifies Three Key Trends That Organizations Must Track to Gain Competitive Advantage
5. <http://www.gartner.com/smarterwithgartner/7-technologies-underpin-the-hype-cycle-for-the-internet-of-things-2016/>
7 Technologies Underpin the Hype Cycle for the Internet of Things, 2016
6. Lopez in sodelavci (2016), High-speed inspection system finds defects in steel, *Vision Systems Design*, vol.21, str. 24-27.
<http://www.vision-systems.com/articles/print/volume-21/issue-11/features/high-speed-inspection-system-finds-defects-in-steel.html>

7. Carl B. Frey, Michael A. Osborne (2013), The Future of Employment: How susceptible are jobs to computerisation?,
http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

INDUSTRIJSKE APLIKACIJE

RAČUNALNIŠKA OBDELAVA SLIKE V PAPIRNIŠTVU IN TISKARSTVU

Aleš Hladnik

Katedra za informacijsko in grafično tehnologijo
Oddelek za tekstilstvo, grafiko in oblikovanje
Naravoslovnotehniška fakulteta, Univerza v Ljubljani
E-pošta: ales.hladnik@ntf.uni-lj.si
URL: <https://sites.google.com/site/aleshladnik/>

POVZETEK: V prispevku sta predstavljena dva primera uporabe računalniške obdelave slik za vrednotenje kakovosti papirnega substrata. Prvi opisuje določitev topografije papirja oziroma njegove gladkosti/hrapavosti z laserskim profilometrom in metodo teksturne analize GLCM. Metoda je uporabna tudi za določanje neenakomernosti odtisa. Drugi primer prikazuje ločitev formacije od markiranja, dveh tipičnih oblik vzorcev prisotnih v papirnem listu, s pomočjo hitre Fourierove transformacije in filtriranja v frekvenčni domeni.

1. UVOD

Papir, vse do nedavnega daleč najbolj razširjeni nosilec informacij, dandanes vse težje tekmuje s sodobnimi digitalnimi mediji in tehnologijami, kot so svetovni splet in družbena omrežja, e-bralniki, pametni telefoni in računalništvo v oblaku. Kljub temu pa si številnih vsakdanjih izdelkov brez tega starodavnega in vsestransko uporabnega naravnega materiala še vedno ne moremo predstavljati. V embalaži, higienskih (papirnati robčki, toaletni papir), tiskovnih (časopisi, revije, knjige) in nekaterih specialnih vrstah papirja ostaja kot surovina nepogrešljiv.

Postopki računalniške obdelave digitalne slike (ang. *Digital image processing*) se, podobno kot na mnogih drugih področjih znanosti in tehnike, vse močneje uveljavljajo tudi v procesih industrijske proizvodnje papirja in njegove dodelave, npr. v tisku, zlasti pri kontroli kakovosti [1]. Izmed številnih aplikacij (tabela 1) bosta v prispevku nekoliko podrobneje predstavljena dva zanimivejša primera uporabe omenjenih postopkov in orodij.

Tabela 1: Primeri postopkov slikovnega procesiranja in analize v papirništvu in tisku [2]

<i>Primer uporabe</i>	<i>Področje</i>
Določanje nečistoč (velikost, oblika, usmerjenost, ipd.) v surovini za	Papirništvo

izdelavo papirja – celulozna vlaknina, recikliran papir – in v papirju	
Merjenje kontaktnega kota med papirjem in tekočino (npr. vodo, etanolom)	Papirništvo
Merjenje debeline premazne mase na površini papirja	Papirništvo
Določanje topografije – gladkosti oz. hrapavosti – površine papirja	Papirništvo
Detekcija, ločevanje in vrednotenje periodičnih in aperiodičnih struktur v papirnem listu	Papirništvo
Vrednotenje kakovosti natisnjenih predmetov (npr. črk) na podlagi določitve deleža pokritosti površine papirja s tiskarsko barvo	Tiskarstvo
Analiza nazobčanosti robov natisnjenih predmetov, razlivanja črnila in neenakomernosti odtisa	Tiskarstvo
Določanje števila manjkajočih točk pri odtisih, izdelanih v tehniki globokega tiska	Tiskarstvo

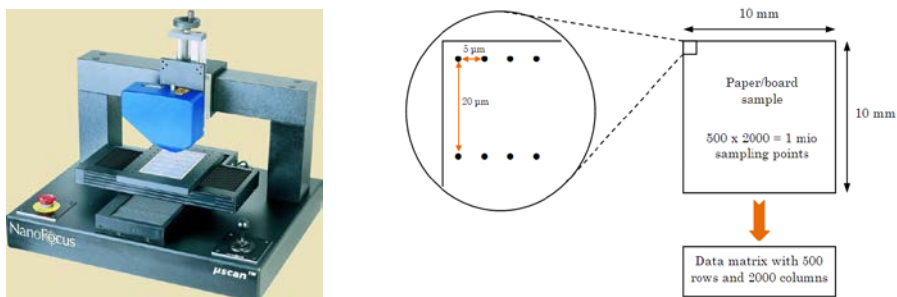
2. VREDNOTENJE TOPOGRAFIJE PAPIRJA Z LASERSKIM PROFILOMETROM IN METODO GLCM

Hrapavost (oziroma njeno nasprotje – gladkost) papirja odločilno vpliva na njegov videz in sijaj, igra pa tudi pomembno vlogo pri različnih postopkih dodelave, zlasti v tisku. Tako lahko na primer pri eni od klasičnih tiskarskih tehnik, globokem tisku, zaradi neposrednega stika med trdo tiskovno formo in papirjem prevelika hrapavost le-tega pri prenosu tiskarske barve povzroči nezaželen pojav manjkajočih natisnjenih pik (ang. *missing dots*) na površini papirja.

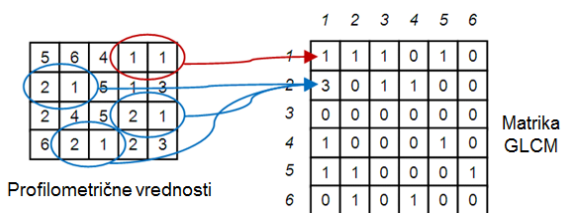
Tradicionalne metode za določanje hrapavosti/gladkosti papirja, ki se še vedno rutinsko uporabljajo v papirniških laboratorijih, temeljijo na merjenju pretoka zraka – njegovega volumna ali pretečenega časa – kar naj bi bilo merilo za povprečno oddaljenost med površino substrata in merilno glavo instrumenta. Te posredne metode pa so precej nenatančne in ne ponujajo nikakršnih podatkov o lokalnem spreminjanju topografije površine merjenega vzorca. Tovrstno informacijo pa je možno pridobiti z različnimi mikroskopskimi (npr. konfokalni laserski vrstični mikroskop – CLSM, mikroskop na atomsko silo – AFM) ali pa profilometričnimi metodami. V eni naših raziskav smo s pomočjo meritev, opravljenih z laserskim profilometrom, na podlagi teksturne analize proučevanih vzorcev papirja z metodo GLCM (ang. *Gray Level Cooccurrence Matrix*) ovrednotili njihovo topografijo.

2.1 GLCM

Konfokalni točkovni senzor laserskega vrstičnega 3D profilometra [3] je za vsak kvadratni centimeter vzorca papirja ali kartona general 500 x 2000 = 1 milijon podatkovnih točk, to je vrednosti oddaljenosti senzorja od površine vzorca (slika 1). Te meritve so predstavljale vhodne vrednosti, iz katerih smo v programu MATLAB za vsak vzorec izračunali t.i. matriko vezanih verjetnosti sivin (GLCM) [4, 5]. Gre za tabelo, ki beleži število pojavitev različnih medsebojnih kombinacij vrednosti (= "sivin"), pridobljenih s profilometrom. Praviloma se spremljajo pari vrednosti, pri čemer se proučevani vrednosti nahajata na določeni medsebojni razdalji (oddaljenost 1, 2, ... do največ 8 podatkovnih točk) in smeri (kot med vrednostima je običajno 0, 45, 90 ali 135°). Princip izračunavanja za pare horizontalno-sosednjih vrednostih je predstavljen na sliki 2.



Slika 1: Laserski vrstični profilometer s konfokalnim točkovnim senzorjem (levo) in postopek vzorčenja (desno).



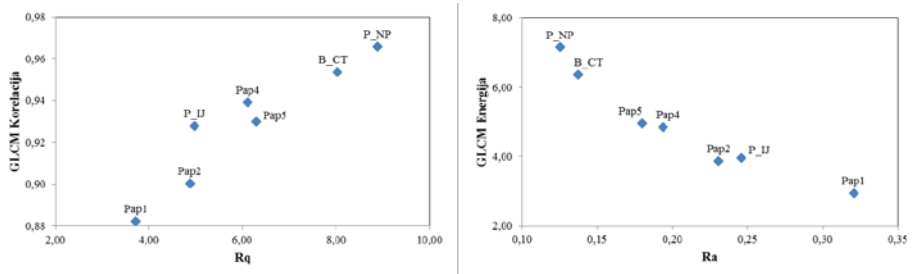
Slika 2: Izračun matrike vezanih verjetnosti sivin (GLCM) za hipotetičen primer vrednosti, dobljenih z laserskim profilometrom.

Na podlagi tako dobljene matrike je možno izračunati razne statistične mere oz. značilke (ang. *texture measures*). Haralick je v svojem izvornem članku o GLCM [6] predlagal 14 značilke, od katerih pa se jih danes v praksi uporablja le manjše število. Tabela 2 prikazuje izračunane vrednosti štirih GLCM značilk za sedem proučevanih vzorcev papirja in kartona. Primerjalno so navedene tudi vrednosti ISO [7] topografskih parametrov R_a in R_q , ki podajata aritmetično povprečno hrapavost in RMS hrapavost

profila površine. Iz tabele in slike 3 je razvidno, da zlasti značilki *korelacija* in *energija* zelo dobro sovpadata z obema ISO parametroma: Pearsonov korelacijski koeficient *Rq-korelacija* znaša 0,94, *Ra-energija* pa -0,96. Soodvisnost teh rezultatov z meritvami, dobljenimi s posredno "papirniško" metodo po Bendsenu, ki temelji na merjenju volumna pretečenega zraka, je bila po pričakovanju slabša.

Tabela 2: Vrednosti GLCM teksturnih značilk, ISO topografskih parametrov in metode po Bendsenu za proučevane vzorce papirja in kartona

Koda vzorca	GLCM teksturna značilka				ISO topografski parameter		Hrapavost po Bendsenu
	Kontrast	Korelacija	Energija	Homogenost	Rq	Ra	
Pap1	0,119	0,882	0,321	0,941	3,72	2,95	79
Pap2	0,161	0,900	0,231	0,919	4,87	3,88	314
P_IJ	0,121	0,928	0,246	0,939	4,98	3,96	205
Pap4	0,148	0,939	0,194	0,926	6,11	4,85	262
Pap5	0,179	0,930	0,180	0,911	6,30	4,96	272
B_CT	0,185	0,954	0,137	0,908	8,03	6,36	1100
P_NP	0,166	0,966	0,125	0,917	8,89	7,15	810



Slika 3: Soodvisnost med GLCM značilko *korelacija* in ISO parametrom *Rq* (levo) ter med GLCM značilko *energija* in ISO parametrom *Ra* (desno)

2.2 Neenakomernost odtisa – mottling

Omeniti velja, da je opisana metoda uporabna tudi za določanje neenakomernosti odtisa – homogenosti površine papirnega substrata, natisnjene z eno samo barvo (t.i. *mottling*). Gre za enega najpogostejših nezaželenih pojavov, s katerim se srečujejo tiskarji tako pri klasičnih kot pri digitalnih tehnikah tiska. Vzroke za njegov nastanek je praviloma potrebno iskati že v neoptimalni izdelavi samega osnovnega papirja/kartona, pogosto pa tudi pri procesih njegovega premazovanja in sušenja. Te napake na nepotiskani površini najprej niso vidne, opazimo jih šele, ko papir potiskamo. Pojav lahko ovrednotimo na različne načine, eden možnih je tudi analiza teksture z metodo

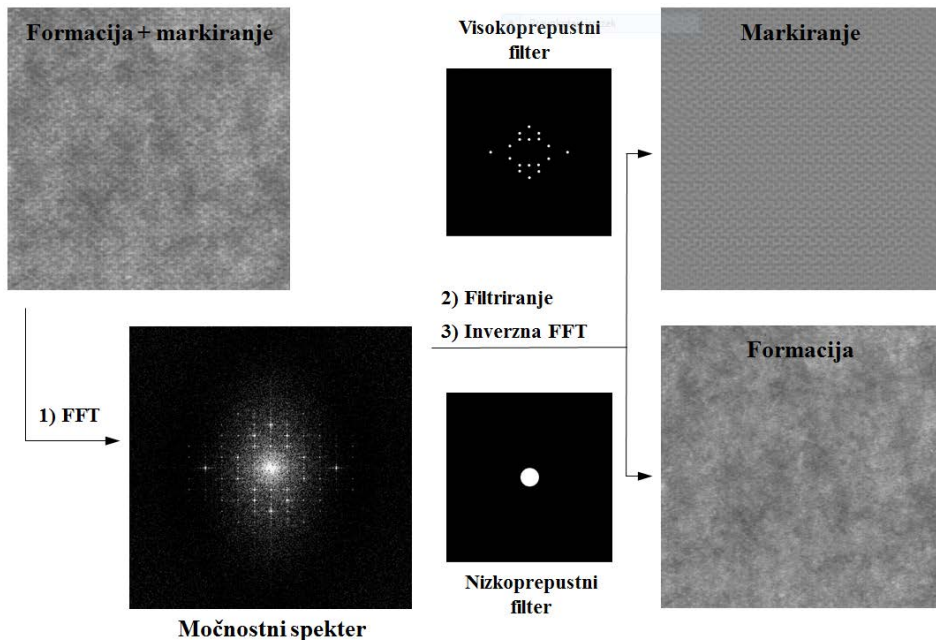
GLCM, s katero lahko uspešno kvantificiramo lokalno spreminjanje barve odtisa oziroma svetilnosti točk (pikslov) pripadajoče sivinske slike.

3. HITRA FOURIEROVA TRANSFORMACIJA IN FILTRIRANJE SLIKE PAPIRJA V FREKVENČNI DOMENI

Cilj druge študije je bila detekcija, medsebojna ločitev in kvantitativno ovrednotenje dveh vrst vzorcev oziroma neenakomernosti, ki jih lahko najdemo v listu papirja po njegovi izdelavi. Zaradi bolj ali manj naključnega nalaganja celuloznih vlaken v začetnem delu proizvodnje papirja pridobi papir značilno *formacijo* oziroma *pregled* ("oblačnost") z značilno aperiodično, stohastično strukturo. Poleg tega pa pride papir v nadaljnjem procesu izdelave v neposreden stik s papirniškim sitom sintetičnega izvora, čigar periodična struktura se vanj šibkeje ali močneje vtisne – t.i. *markiranje*. Prisotnost ene ali obeh vrst neenakomernosti, če sta preveč izraziti, močno zniža kakovost izdelanega papirja in s tem vrednost končnega izdelka.

Da lahko ta dva tipa struktur prepoznamo kot slikovni šum nizkih oziroma visokih frekvenc, zaradi česar lahko za učinkovito reševanje tovrstnih problemov uporabimo hitro Fourierovo transformacijo (*Fast Fourier Transform*, FFT) na digitalni sliki vzorca papirja, je prepoznal že Tanson v 90-ih letih prejšnjega stoletja [8, 9]. Naš pristop temelji na njegovem delu in sestoji iz naslednjih glavnih korakov [10]:

- Zajem slike s ploskim skenerjem v presevnem načinu in njena pretvorba v 8-bitno sivinsko sliko.
- Pretvorba slike iz prostorske v frekvenčno domeno z uporabo FFT in generiranje močnostnega spektra. Ta prikazuje kvadrirane amplitudne vrednosti frekvenčnega spektra v obliki 2D diagrama: vrednosti blizu njegovega središča ustrezajo nizkofrekvenčnim območjem, vrednosti daleč od središča pa visokofrekvenčnim območjem originalne slike v prostorski domeni.
- Po potrebi povečanje kontrasta z namenom lažjega generiranja filtra.
- Filtriranje slike v frekvenčni domeni z uporabo ustreznega nizko - oziroma visokoprepustnega filtrirnega okna.
- Pretvorba modificirane slike nazaj v prostorsko domeno.



Slika 4: Prikaz ločevanja pojavov markiranja in formacije v papirnem listu s pomočjo FFT in filtriranja slike v frekvenčni domeni.

Osnovno idejo filtriranja slike v frekvenčni domeni prikazuje slika 4. Na podlagi informacij iz močnostnega spektra smo izdelali ustrezne visoko- oziroma nizkoprepustne filtre, s katerimi smo bodisi prepustili zgolj pravilne strukture v papirnem listu in na ta način omogočili vizualizacijo bodisi izključno markiranja bodisi zgolj nepravilnih, stohastičnih struktur, ki predstavljajo formacijo papirja. Kompleksnejši primeri bodo predstavljeni na predavanju.

V nadaljevanju lahko izvedemo slikovno analizo dobljenih struktur; za kvantitativno vrednotenje formacije lahko, med drugim, uporabimo tudi metodo GLCM.

3. ZAKLJUČEK

Navkljub visoki stopnji avtomatizacije proizvodnih procesov v papirništvu in tisku je laboratorijsko preverjanje surovin in končnih izdelkov še vedno pomemben vidik zagotavljanja kakovosti izdelkov iz papirja in tiskovin. V prihodnosti pričakujemo vse večjo integracijo v prispevku prikazanih postopkov slikovnega procesiranja, kakor tudi metod računalniškega vida in strojnega učenja v industrijske sisteme nadzora kakovosti.

LITERATURA

1. A. Hladnik, T. Muck (2010), *Obdelava digitalnih slik v grafiki – Osnove*, Ljubljana: Naravoslovnotehniška fakulteta.
2. <https://sites.google.com/site/obdelavadigitalnihslik/>
Spletno mesto avtorja o obdelavi digitalnih slik na področju papirništva in tiska
3. <http://www.nanofocus.com/produkte.html?&L=1>
3D laserski profilometer s konfokalnim točkovnim senzorjem proizvajalca Nanofocus
4. <http://www.fp.ucalgary.ca/mhallbey/tutorial.htm>
Opis metode GLCM
5. A. Hladnik, M. Lazar (2011), Paper and board surface roughness characterization using laser profilometry and gray level cooccurrence matrix, *Nordic Pulp and Paper Research Journal*, 26(1), str. 99-105.
6. R. M. Haralick, K. Shanmugam, I. Dinstein (1973), Textural features for image classification, *IEEE Trans. Syst. Man Cybern.*, 3, str. 610–621.
7. BS EN ISO Standard (2000): 4287:2000.
8. S. J. T'Anson (1995), Identification of periodic marks in paper and board by image analysis using two-dimensional fast Fourier transforms - Part 1: The basics, *Tappi Journal*, 78(3), str. 113.
9. S. J. T'Anson (1995), Identification of periodic marks in paper and board by image analysis using two-dimensional fast Fourier transforms - Part 2: Forming and press section marks, *Tappi Journal*, 78(7), str. 97.
10. A. Hladnik, T. Muck, M. Stanić, M. Černič (2012), Fast Fourier transform in papermaking and printing: two application examples, *Acta Polytechnica Hungarica*, 9(5), str. 155-166.

OPTIČNA DETEKCIJA NAPAK NA VISOKO-REFLEKTIVNIH STRUŽENIH POVRŠINAH

Teja Jereb, Urban Pavlovčič, Matija Jezeršek

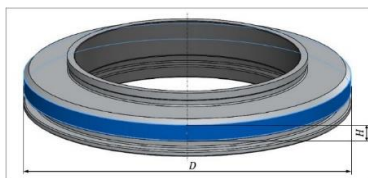
Laboratorij za lasersko tehniko
Fakulteta za strojništvo
e-pošta: jereb.teja@outlook.com

POVZETEK: V članku je predstavljen razvoj laboratorijskega sistema za optično detekcijo površinskih napak na visoko reflektivnih, cilindričnih površinah. Izrazit kontrast med napako in merjeno površino smo dosegli z metodo temnega polja osvetlitve z belo svetlobo. Izdelan je bil program za detekcijo in karakterizacijo površinskih napak. Program izvede svetlostno poravnavo slik, segmentacijo, izločitev drobnih napak in izmeri dimenzije detektiranih napak. Razvit in preizkušen laboratorijski sistem izpolnjuje zahteve, saj je na 128 referenčnih kosih zaznal vse napake, katerih velikost presega 0,6 mm. Koncept je primeren za nadaljnji razvoj avtomatskega sistema za detekcijo napak na struženih nepresevnih površinah.

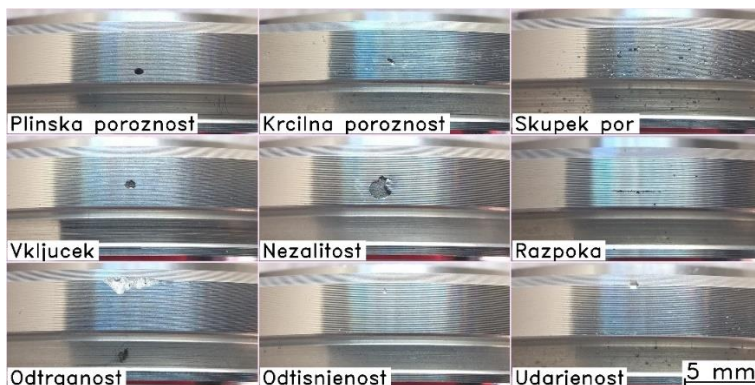
1. UVOD

Optične metode merjenja se vse pogosteje uporabljajo za nadzor kakovosti izdelkov, saj se njihova prednost odlikuje predvsem na brezkontaktnem principu delovanja, zato je z njimi mogoče izvajanje 100 % kontrole izdelkov. Izbira optičnih metod je ustrezna predvsem za doseganje večjih natančnosti in vnaprej predvidenih točnosti [1].

Merjeni objekt predstavlja tesnilni element, izdelan s postopkom visokotlačnega litja aluminija, ki je v nadaljevanju postružen na končne dimenzije, kot je prikazano na sliki 1. Za zagotavljanje tesnilnega učinka na kontaktnih površinah mora biti površina po mehanski obdelavi fine hrapavosti, kar se odraža tudi v visoki reflektivnosti. Zaradi samega postopka izdelave se na kontaktnih površinah po procesu struženja pojavijo površinske napake, ki povzročajo netesnost. Tipične napake, ki nastanejo pri tlačnem litju so prikazane na sliki 2.

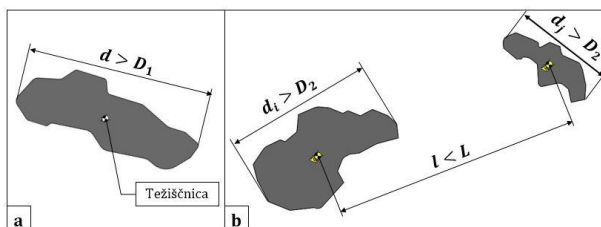


Slika 1: Oblika merjenega objekta. Z modro barvo je označena merjena površina premera $D = 99,7$ mm in višine $H = 4,90$ mm.



Slika 2: Površinske napake, ki nastanejo pri tlačnem litju.

Optični merilni sistem mora prepoznati površinske napake in upoštevati dimenzijsko neustreznost napak, pri čemer mora dosegati optično ločljivost vsaj 0,06 mm. Po končani kontroli posameznega izdelka mora sistem prikazati diskretno odločitev ustrezen/neustrezen s svetlobno signalizacijo, pri čemer pogoj za slab izdelek prikazuje slika 3. Če je najdaljša izmerjena diagonala d večja od D_1 , je izdelek označen kot slab (slika 3a). Prav tako je izdelek slab, če sta najdaljši izmerjeni diagonali i -te in j -te napake večji od D_2 in se nahajata znotraj razdalje L (slika 3b).



Slika 3: Pogoji za neustreznost površinskih napak iz velikostnega razreda PK4. (a) Maksimalna velikost napake D_1 . (b) Minimalna razdalja med dvema napakama L , katere dimenzije presegajo D_2 .

Pri postopku visokotlačnega litja aluminija se pojavljajo površinske napake različnih velikosti, oblik in orientacij. Zato je v prvi vrsti potrebno napake kar se da dobro okarakterizirati [2].

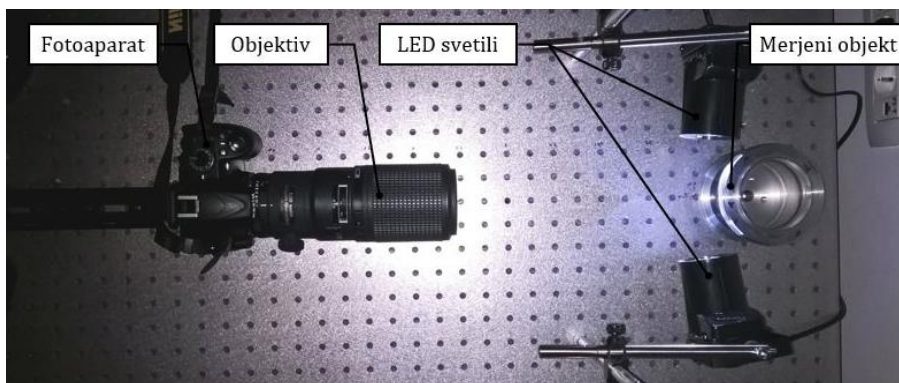
Pogosto uporabljeni metodi za detekcijo površinskih napak na nepresevnih materialih, sta osvetljevanje s svetlim (angl.: *bright field illumination*) [3] in temnim poljem (angl.: *dark field illumination*) [3–6]. Sledi še izbira kamere, s katero vplivamo na velikost merilnega območja, ločljivost in na hitrost detekcije napak.

Za analizo slike je treba izvesti programsko obdelavo. Tekstura površine lahko predstavlja manjši delež motenj, zato so se avtorji del najpogosteje posluževali Gaussovega filtra [7]. Nato so izvajali segmentacijo slike z binarnimi transformacijami [3, 8]. Šuma v obliki

sledi vzorcev mehanske obdelave so se avtorji znebili z uporabo morfoloških operacij in algoritmov za zaznavo kontur [3]. Za slednje so najpogosteje uporabljali LaPlaceov [7] in Sobelov [9] filter, s katerimi so ločili površinske napake od ozadja slike.

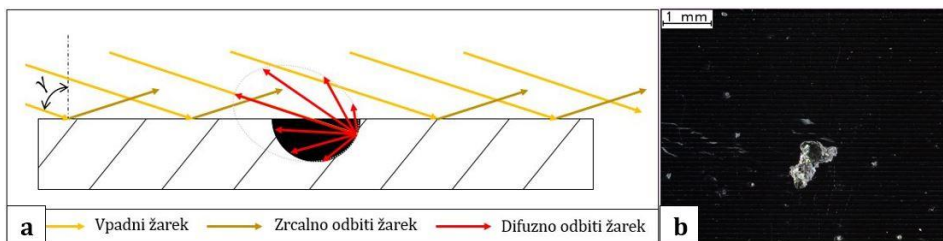
2. EKSPERIMENTALNI SISTEM

Ekperimentalni sistem prikazuje slika 4. Sestoji iz dveh LED svetil, ki pod vpadnim kotom skoraj 90° osvetljujeata merjeni objekt, ki ga slikamo z digitalnim fotoaparatom.



Slika 2: Ekperimentalni sistem.

Osvetlitev ima najpomembnejšo vlogo pri zagotavljanju visokega kontrasta med merjeno površino in napako. Treba je poskrbeti predvsem za homogenost osvetlitve, s čimer se doseže enak intenzitetni profil čez celotno merjeno površino. V našem primeru se poslužujemo metode temnega polja osvetlitve od spredaj z belima LED reflektorjema. Pri tem svetloba pada na merjeno površino pod kotom blizu 90° glede na optično os kamere. Večina svetlobe se s površine odbije stran od detektorja slike, zato je površina temnih barv. Ko vpadni žarek zadane napako, pride do difuznega odboja, zaradi česar je napaka na sliki svetlejša (slika 5).



Slika 3: (a) princip temnega polja osvetlitve. (b) tipična slika površinske napake.

Za slikanje merjene površine smo uporabili zrcalno-refleksni fotoaparat Nikon D3100, makro objektiv s fiksno goriščno razdaljo AF MICRO NIKKOR $f = 200 \text{ mm } 1:4 \text{ D}$ in dve LED-svetili s kolimirano svetlobo (angl.: *spotlight*) proizvajalca Lixada.

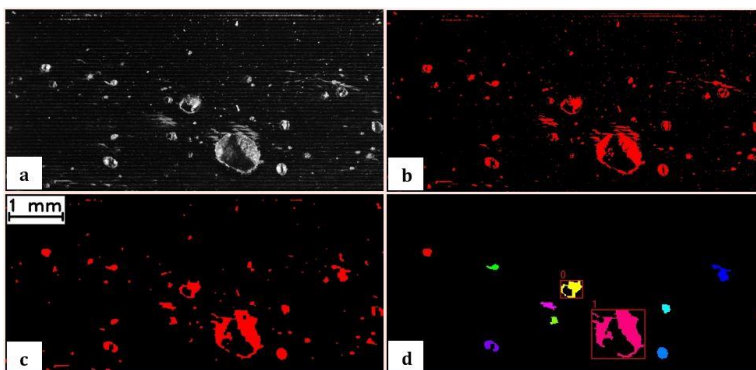
4. PROGRAMSKA OPREMA

Program za detekcijo površinskih napak je bil izdelan v grafičnem programskem jeziku LabVIEW 2015. Vhod v program predstavlja vnos barvne slike, ki jo pretvorimo v sivinsko (slika 6a) in ji določimo območje interesa. Sledi izračun histograma, s katerim določimo vrednost enostranskega praga pretvorbe v binarno sliko (slika 6b). Opazimo lahko, da je na posnetku še vedno prisoten šum v obliki ozkih črt, ki so posledica struženja, majhnih delcev in prask. Zato v naslednjem koraku sledi zaporedje treh morfoloških operacij za odpravo omenjenih drobnih napak (slika 6c). Uporabili smo kombinacijo metod erozije in dilatacije. Erozija deluje na principu krčenja, dilatacija pa na principu razširitev delcev za vnešeno število operacij. Preglednica 1 prikazuje zaporedje korakov morfoloških operacij in funkcije *zapolnitev lukenj*. Pomembno je, da je skupno število operacij erozije in dilatacije v končni fazi enako, saj bi v nasprotnem primeru vplivali na velikost posamezne napake.

Preglednica 1: Zaporedje uporabe morfoloških operacij in funkcije zapolnitev lukenj.

korak	morfološka operacija	število iteracij
1	erozija	2
2	zapolnitev lukenj	/
3	dilatacija	10
4	erozija	8
5	zapolnitev lukenj	/

Za izločitev napak, katerih najdaljše diagonale znašajo manj kot D_2 , smo uporabili morfološki filter, ki izvede 11 erozij. Ker je optična ločljivost sistema (velikost slikovne točke) enaka 0,005 mm, pomeni, da odstrani vse napake, katerih krajša dimenzija je krajša od D_2 . Na koncu smo uporabili še funkcijo *štetje objektov* (LabVIEW, *IMAQ count objects 2*) [10], s katero se zazna napake, katerih minimalna diagonalna znaša 0,6 mm, jih oštevilči in očrta z rdečim kvadratom, kot je prikazano na sliki 6d.



Slika 4: Obdelava slike: (a) določanje območja interesa, (b) binarizacija slike, (c) rezultat morfoloških operacij in (d) končni rezultat.

Na koncu algoritem uvrsti izmerjeno površino med ustrezne ali neustrezne po pravilih, ki so prikazana na sliki 3.

5. REZULTATI

S sistemom smo izmerili 128 vzorcev, ki so bili predhodno vizualno pregledani in izmerjeni s strani eksperta in razvrščeni glede na tip površinske napake. Ugotovili smo (preglednica 2), da je program v sedmih primerih naredil napako prve vrste (ustrezen vzorec je označil kot neustreznega). Napake druge vrste (neustrezen vzorec je označil kot ustreznega) sistem ni naredil.

Preglednica 2: Rezultat testiranja sistema na referenčnih vzorcih. Napako 1. vrste imenujemo tudi *tveganje dobavitelja*, napako 2. vrste pa *tveganje kupca*.

vzorci	število	odstotek
vsi	128	100,0 %
preverjeno neustrezni	50	39,1 %
izmerjeno neustrezni	57	44,5 %
napaka 1. vrste	7	5,5 %
napaka 2. vrste	0	0,0 %

Rezultati meritev so prikazani v preglednici 3, kjer je razvidno, da so programsko izmerjene dimenzije napak v povprečju za 0,2 mm večje. V vseh sedmih primerih je sistem označil ustrezen vzorec kot neustreznega, kar pripelje do 5,7 % napake. Pri analizi vseh napačnih odločitev sistema smo ugotovili, da je program podal napačne odločitve v primerih dveh tipov napak *skupek por* in *poroznost*.

Preglednica 3: Primerjava ročno in programsko izmerjenih vzorcev, kjer je sistem ustrezne vzorce označil kot neustrezne.

vrsta napake	Vzorec	ROČNO IZMERJENO		PROGRAMSKO IZMERJENO		opomba
		napaka 1 [mm]	napaka 2 [mm]	napaka 1 [mm]	napaka 2 [mm]	
skupek por	1	0,63	0,41	0,66	0,64	Praske
	2	0,59	0,58	0,90	0,85	Praske
	3	0,69	0,45	1,18	0,83	Nečistoče
poroznost	1	0,61	/	1,23	/	Praska
	2	0,90	0,57	0,80	0,61	
	3	0,60	0,54	1,03	/	Praske
	4	0,99	/	1,11	/	

Po naknadni analizi napačno izmerjenih površin smo ugotovili, da je bil v večini primerov vzrok napačnih odločitev poškodovanost površine s praskami različnih oblik in dimenzij ter kontaminacija z nečistočami. Program prasko ali nečistočo in površinsko napako zaradi osvetlitve združi v eno samo napako in posledično ima združena napaka najdaljšo

diagonalno daljšo od D_2 ali celo od D_1 . Poudariti je treba, da so se omenjene napake pojavile med transportom iz podjetja do laboratorija, torej po pregledu s strani eksperta. Predvidevamo, da bi v primeru takojšnje kontrole vzorcev sistem samo v dveh primerih označil ustrezen vzorec kot neustrezen. V takem primeru bi program za konfiguracijo temnega polja osvetlitve deloval z napako 1,6 %.

6. ZAKLJUČKI

Razvit je bil laboratorijski sistem za optično detekcijo površinskih napak na cilindričnih, visoko reflektivnih površinah. Prototip je dimenzioniran na zaznavnost napake minimalne prečne dimenzije 0,06 mm. Uporabili smo metodo temnega polja osvetlitve, s katero smo dobili dober kontrast med temno podlago in svetlimi napakami. Razvita programska oprema omogoča robustno detekcijo napak, določitev njihovih dimenzij in določitev ustreznosti merjenega objekta.

Rezultati testiranja kažejo, da je metoda temnega polja osvetlitve predvsem primerna za detekcijo plitkih napak, saj v tem primeru sistem ni podal nobene napačne odločitve. Napačen rezultat se je pojavil v primeru globljih napak, kot je na primer poroznost, zato je sistem v 5,7 % označil ustrezen vzorec kot neustreznega. Do primera, ko bi sistem neustrezen vzorec označil kot ustreznega, ni prišlo. Z nadaljnjo analizo posnetkov smo ugotovili, da so vzrok za napačno odločitev sistema poškodbe v obliki prask in nečistoče. S primernim rokovanjem bi lahko napako sistema zmanjšali na 1,5 %.

Na osnovi rezultatov testiranja menimo, da je razvit koncept primeren za nadaljnji razvoj avtomatskega sistema za detekcijo napak na struženih nepresevnih površinah.

LITERATURA

- [1] A. T. Jeffrey: *Optical Measurement Techniques*. Predavanja pri predmetu ME637 – Particle Transport, Deposition and Removal II, Ogljed: 28.12.2016.
- [2] V. Gruden: *Lasersko podprto ocenjevanje kakovosti ulitkov v realnem času: Doktorska disertacija*, Univerza v Ljubljani, Fakulteta za Strojništvo, 2010.
- [3] N. Bonnot, R. Seulin, F. Merienne: *Machine vision system for surface inspection on brushed industrial parts*. V: *Proceedings of the SPIE*, zvezek 5303, 2004, str. 64 – 72.
- [4] T. Giesko, A. Mazarkiewicz, A. Zbrowski: *Advanced Mechatronic System for In-Line Automated Optical Inspection of Metal Parts*. **11**(2010) str. 22 – 38.
- [5] L. C. Jiang, Z. Zhang, I. Nakamura, in sod.: *Developing a new automatic vision defect inspection system for curved surfaces with highly specular reflection*. *International Journal of Innovative Computing, Information and Control* **8**:7(B) (2012) str. 5121 – 5136.
- [6] G. Rosati, G. Boschetti, A. Biondi, A. Rossi: *Real-time defect detection on highly-reflective curved surfaces*. V: *Proceedings of the OPTIMESS2007 Workshop*. Leuven, Belgium, 2007, str. 115 – 122.

- [7] S. J. Świllo, M. Perzyk: *Automatic inspection of surface defects in die casting after machining*. Archives of Foundry Engineering **11**(2011) str. 231 - 236.
- [8] S. Ho Choi, J. Pil Yun, *in sod.: Real-time Defects Detection Algorithm for High-Speed Steel Bar in Coil*. International Journal of Computer, Electrical , Automation, Control and Information Engineering **1**:1 (2007) str. 22 – 26.
- [9] X. Zhang, Y. Ding, Y. Lv, *in sod.: A vision inspection system fort he surface defects of strongly reflected metal based on multi-classs SVM*. Expert Systems with Applications **38**(2011) str. 5930 – 5939.
- [10] National Instruments: *IMAQ Count Objects 2* VI. Dostopno na: https://zone.ni.com/reference/en-XX/help/370281M-01/imaqvision/imaq_count_objects_2/, Oglad: 14.12.2016.

PREGLED PROGRAMSKIH ORODIJ ZA GLOBOKO UČENJE Z VIDIKA UPORABE V INDUSTRIJSKIH APLIKACIJAH

Domen Tabernik, Danijel Skočaj

Univerza v Ljubljani

Fakulteta za računalništvo in informatiko

E-pošta: domen.tabernik@fri.uni-lj.si

URL: <http://www.vicos.si>

POVZETEK: *Globoko učenje je prineslo revolucionarne spremembe na področju računalniškega vida in si utira svojo pot tudi na področje industrijskega strojnega vida. V tem članku predstavljamo šest najbolj poznanih orodij za delo z globokimi arhitekturami: Caffè, Torch, Theano, MatConvNet, TensorFlow in Keras. Predstavili bomo njihove glavne značilnosti tako z vidika razvoja kot integracije v industrijske aplikacije.*

1. UVOD

V zadnjih nekaj letih smo bili na področju računalniškega vida priča veliki spremembi v paradigmi reševanja problemov. S povečanjem računskih zmogljivosti in razpoložljivosti ogromnega števila slik ter z napredkom pri razvoju pristopov, ki temeljijo na principu globokega učenja, so se globoke arhitekture uveljavile kot najuspešnejša predstavitev vizualne informacije, ki omogoča uspešno reševanje različnih problemov kot so detekcija, razpoznavanje in kategorizacija predmetov in njihovih delov, ali semantična segmentacija.

Industrija je tradicionalno previdnejša pri sprejemanju novih tehnologij. Kljub temu je opazen trend prodiranja globokega učenja tudi na področje industrijskih aplikacij kot je nadzor kvalitete izdelkov. Avtomatsko učenje značilnic in klasifikatorjev na podlagi učnih primerov se namreč zdi v veliko primerih primernejše kot klasično programiranje namenskih detektorjev za vsak problem posebej. Zato lahko pričakujemo, da bo vse več rešitev na področju industrijskega strojnega vida temeljilo na globokih arhitekturah. Pri tem se postavlja vprašanje katera orodja za delo z globokimi arhitekturami so najprimernejša za razvoj takšnih rešitev in integracijo le-teh v sisteme strojnega vida.

V tem članku bomo predstavili pregled šestih najpomembnejših obstoječih programskih orodij za globoko učenje ter opisali njihove prednosti in slabosti. Orodja bomo analizirali predvsem z vidika njihove uporabnosti v aplikacijah strojnega vida kot so nadzor kvalitete, detekcija napak oz. klasifikacija površin industrijskih izdelkov in polizdelkov ipd. Povzeli bomo njihove lastnosti v fazi razvoja aplikacij kot tudi primernost za integracijo razvitih rešitev v celovite sisteme strojnega vida.

2. ORODJA

V tem poglavju so opisana sledeča orodja za globoko učenje: Caffe, Torch, Theano, MatConvNet, TensorFlow in Keras. Tabela 1 povzema nekatere glavne značilnosti teh orodij: katere programske jezike in operacijske sisteme podpirajo ter pod kakšno licenco so izdani. V nadaljevanju poglavja so ta orodja predstavljena bolj podrobno.

Tabela 1: Povzetek glavnih značilnosti šestih orodij za globoko učenje.

Orodje	Programski jeziki	Operacijski sistem	Licenca
Caffe [1]	C/C++ in CUDA jedro, C++, Python, Matlab vmesnik	Primarno Linux, Windows slabše podprt	BSD 2-Clause
Torch [2]	C/C++ in CUDA jedro, Lua vmesnik, Python	Primarno Linux, Windows slabše podprt	BSD 3-Clause
Theano [3]	Python, C in CUDA jedro	Primarno Linux, Windows slabše podprt	BSD
MatConvNet [4]	Matlab z Mex/C oz. CUDA jedri	Linux, Windows	Permissi ve BSD
TensorFlow [5]	C/C++ in CUDA jedro, Python/C++ vmesnik	Primarno Linux, MacOS, Windows	Apache 2.0
Keras [6]	Python vmesnik nad orodjem Theano ali TensorFlow	Odvisno od izbire jedra (Theano ali TensorFlow)	MIT

2.2 Caffe

Orodje Caffe [1] je osredotočeno na globoko učenje aplicirano na slikovne zbirke. V ta namen ima odlično podporo za modele, ki izhajajo iz konvolucijskih mrež (ConvNet) in zanje vsebuje ustrezne najnovejše module oz. nivoje kot so BatchNorm, Dropout, ReLU, PReLU itd. V zadnjem času se izboljšuje tudi podpora za ne-konvolucijske mreže, kot so rekurzivne mreže (RNN) z nivoji LSTM.

Jedro orodja Caffe se razvija v programskem jeziku C/C++, pri tem pa ima večina nivojev implementacijo tudi v programskem jeziku CUDA, kar omogoča izvajanje na grafičnih karticah na sistemih z enim in več grafičnimi procesorji. Večina nivojev ima implementacijo CUDA realizirano tudi s pomočjo knjižnice CuDNN, kar omogoča optimalno izvajanje mrež na grafičnih karticah nVidia. Vmesnik do jedra je implementiran v jezikih C++, Python in Matlab. Uporaba vmesnika Python in Matlab omogoča hiter in prototipen razvoj metod, vendar le z uporabo že pripravljenih nivojev. Nove nivoje je sicer mogoče implementirati v jeziku Python ali Matlab, vendar je zaradi učinkovitega izvajanja na CPU in GPU potrebno nove nivoje implementirati ročno v jezikih C/C++ in CUDA. Orodje Caffe tudi nima neposrednih orodij za vizualizacijo, vendar je to mogoče nadomestiti z vizualizacijskimi orodji iz knjižnic programskih jezikov Python in Matlab.

Glavni nosilec razvoja orodja Caffe je skupina Berkeley Artificial Intelligence Research (BAIR), obstaja pa tudi širša skupnost raziskovalcev, ki pomaga pri razvoju. Orodje

Caffe je eno izmed glavnih orodij raziskovalcev pri razvoju novih metod globokega učenja, zaradi česar je podpora za vzdrževanje in razvoj stabilna. Kot glavno orodje raziskovalcev ima prednost v hitrem dostopu do najnovejših, najaktualnejših metod in ogromnem naboru pred-naučeni modelov, kot so AlexNet, VGG, ResNet, GoogLeNet in različnih pristopov kot so R-CNN, CPM, DCNN itd. Velika večina teh modelov je zbranih na uradni strani pod imenom Caffe Model Zoo [7].

Z vidika integracije ima Caffe dokaj enostaven proces prevajanja z uporabo orodja CMake. Pri tem ima sicer nekaj pomembnih odvisnosti od tretjih knjižnic: HDF5, protobuf, glog, gflags, hdf5, ATLAS/OpenBLAS, Boost in CUDA. Ker je Caffe primarno razvit za operacijski sistem Linux je proces prevajanja na tem sistemu dokaj enostaven, medtem ko za operacijski sistem Windows obstaja podpora le iz ločene veje GIT in ne zagotavlja najnovejših popravkov in nadgradenj.

Glavni problem orodja Caffe z vidika integracije je uporaba “globalnih” knjižnic kot so glog in gflags, ki skrbita za napake in jih vračata kot izhod iz programa. V primeru uporabe orodja Caffe kot ločene knjižnice to predstavlja velik problem. Odprava te napake bi zahtevala premik velikega dela kode, kar bi povzročilo težave pri vzdrževanju kode ob dodajanju najnovejših popravkov.

2.3 Torch

Programsko orodje Torch [2] je splošno namenska knjižnica tako za globoko učenje, kot tudi za strojno učenje. Poglavitna lastnost orodja je v implementaciji globokega učenja z osnovnimi tenzorskimi operacijami, le-te pa so učinkovito implementirane bodisi v programskem jeziku C, bodisi v jeziku CUDA/CuDNN ali OpenCL za izvajanje na grafičnih karticah. Preko vmesnika je nato mogoče implementirati oz. uporabiti različne tipe globokega učenja, od konvolucijskih mrež (ConvNet) do rekurzivnih mrež (RNN) in modelov LSTM, omogoča tudi podporo za modele kot so variacijski avtoenkoderji in modelov RBM (ang. Restricted Boltzman Machine). Orodje ima tudi velik nabor splošno namenskih metod strojnega učenja, kot so K-Means, SVM, PCA, LDA itd. Orodje vsebuje veliko različnih najaktualnejših modelov in pristopov, vendar pa je z vidika konvolucijskih mrež teh modelov malenkost manj kot v primerjavi z orodjem Caffe. Obstaja tudi veliko pred-naučeni modelov, obenem pa je mogoče z uporabo neuradnih pretvornikov v Torch iz orodja Caffe pretvoriti večino modelov, ki uporabljajo le standardne nivoje.

Primarni vmesnik do orodja Torch je implementiran v programskem jeziku Lua, obstaja pa tudi poenostavljena podpora za MATLAB. Omeniti velja tudi PyTorch, ki se v zadnjem času aktivno razvija in omogoča vmesnik do orodja Torch preko programskega jezika Python. Hiter, prototipen razvoj je z jezikom Lua dokaj enostaven. Vmesnik Lua vsebuje tudi knjižnice za vizualizacijo preko ogrodja Qt ali knjižnice GNUPlot, pri čemer vmesnik priskrbi knjižnice za izris, vizualizacijo modela pa je podobno kot pri Caffe potrebno narediti ročno.

Torch je glavno orodje za raziskovalce iz skupine Facebook-AI (FAIR), zato je podpora pri vzdrževanju in razvoju kode zelo dobra, obenem pa obstaja tudi širša skupnost raziskovalcev, ki pri tem pomaga. Primarno se orodje Torch razvija za operacijski sistem

Linux, medtem ko obstaja za operacijski sistem Windows neuradna podpora s strani skupnosti. Z vidika integracije je Torch zelo kompleksno orodje z veliko paketi zato ima posledično veliko število odvisnosti do tretjih knjižnic. Ž osnovni paket vsebuje nekaj pomembnih odvisnosti (OpenBLAS, libjpeg, imagemagick, zeromq, graphicsmagick, xquartz), pri tem pa se za proces prevajanja ne uporablja avtomatskih orodij ampak običajne skripte. Za uporabo v tretjih aplikacijah predstavlja težavo integracija velikega nabora paketov, prav tako pa je prilagajanje procesa prevajanja lahko kompleksno opravilo.

2.4 Theano

Theano [3] je splošno namenska knjižnica za simbolično reševanje matematičnih optimizacijskih problemov. Pri tem je knjižnica idealna za probleme modelirane z računskim grafom. Glavna lastnost je prevajanje iz simbolične definicije problema neposredno v kodo C, kar omogoča generiranje močno optimizirane kode. Obenem omogoča prevajanje neposredno v kodo GPU (tudi CuDNN), s čimer postane knjižnica primerna za obdelavo ogromnih količin podatkov iz globokih mrež. Model računskega grafa je tudi primeren za implementacijo poljubnih nevronske mreže, vključno s ConvNet, še bolj primeren pa je za rekurzivne mreže (RNN). Kot splošno namenska knjižnica Theano nima eksplicitno pripravljenih nivojev/mrež za globoko učenje in jih je potrebno ročno ustvariti s pomočjo vmesnika. Prav tako ne obstaja množica obstoječih pred-naučenih modelov kot pri orodjih Caffe in Torch. Obstaja pa neuradno orodje za pretvarjanje modelov iz orodja Caffe v orodje Theano s čimer je mogoče uporabiti nekatere pred-naučene modele iz orodja Caffe.

Theano vsebuje vmesnik implementiran v programskem jeziku Python, ki omogoča hiter in prototipen razvoj. Podobno kot orodje Caffe ne vsebuje knjižnic za vizualizacijo zato je potrebno uporabiti tretje knjižnice, kot so matplotlib. Razvoj orodja Theano podpira Montreal Institute for Learning Algorithms z Univerze v Montrealu.

Theano se razvija primarno za operacijski sistem Linux, obstaja pa tudi neuradna podpora za operacijski sistem Windows, ki pa je lahko manj stabilna. Z vidika integracije orodje Theano ne vsebuje veliko odvisnih knjižnic. Glavne knjižnice so BLAS in Python knjižnice (NumPy, SciPy), kar močno poenostavi proces prevajanja. Glavni problem je vprašljiva modularnost, saj je API močno integriran v Python. Prav tako Theano nima eksplicitnega aplikacijsko programskega vmesnika za C/C++, kar onemogoča uporabo pred-naučenih modelov v produkcijskem okolju preko C/C++ vmesnika.

2.5 MatConvNet

MatConvNet [4] je implementacija konvolucijskih mrež (ConvNet) kot programski paket (ang. toolbox) za programski jezik MATLAB. Konvolucijske mreže so implementirane v kodi C in CUDA (tudi z uporabo knjižnice CuDNN) in imajo le vmesnik MEX za MATLAB. Orodje podpira le konvolucijske in enostavne polno povezane mreže. Pri tem ima podporo za nekatere najbolj popularne nivoje, kot so

BatchNorm, ReLU, Dropout in ROI Pooling, medtem ko je podpora za ostale tipe mrež, kot so rekurzivne mreže, pomanjkljiva. Orodje podpira večino popularnih pred-naučenih mrež, kot so AlexNet, VGG, GoogLeNet in ResNet, manjka pa podpora za ostale najaktualnejše modele kot so semantična segmentacija itd. MatConvNet ne vsebuje orodij za vizualizacijo, vendar je preko vmesnika mogoče narediti vizualizacijo z različnimi funkcijami v MATLABu. Prav tako je mogoče hiter in prototipen razvoj s pomočjo okolja MATLAB.

MatConvNet je podprt s strani skupine z Univerze v Oxfordu. Pri tem skupina sčasoma vključuje bolj pomembne modele in nivoje globokega učenja, vendar zaradi obširnosti področja lahko pokriva le najbolj razširjene modele.

Proces prevajanja je pri MatConvNet relativno preprost, vendar ne omogoča enostavne integracije v tretje aplikacije, saj je vmesnik omejen le na okolje MATLAB. Celotna koda je tudi zgrajena na uporabi vmesnika MEX, zato je praktična uporaba zunaj okolja MATLAB nemogoča.

2.6 TensorFlow

TensorFlow [5] je podobno kot Theano orodje za reševanje matematičnih problemov definiranih z računskimi grafi. Podobno kot Theano lahko orodje samodejno izračuna odvode in gradiente na podlagi matematične definicije s čimer omogoča enostavno in hitro prototipno razvijanje novih metod. Za razliko od orodja Theano ima pripravljenih večino pomembnih nivojev za globoko učenje (konvolucijski nivoji, max pooling, ReLU, dropout, BatchNorm, itd). Ker je osnovan na računskih grafih, ima tudi dobro podporo za rekurzivne mreže (RNN). Poleg osnovnih funkcij ima neuradno podporo za še večji nabor metod globokega učenja, kot so metode Monte Carlo, variacijske metode, CRF, itd. Za orodje TensorFlow obstaja nekaj pripravljenih modelov, kot so ResNet, Inception, avtoenkoderji, itd. Nekaj najbolj popularnih mrež je tudi pred-naučenih in pripravljenih za uporabo (VGG, ResNet), vendar je večina teh mrež razdrobljenih po različnih spletnih straneh in niso enostavno zbrani na enem mestu kot na Caffe Model Zoo. Obstaja tudi neuradno orodje za pretvarjanje modelov s standardnimi nivoji iz orodja Caffe v model za orodje TensorFlow. Ker raziskovalci ne uporabljajo pogosto orodja TensorFlow kot primarno orodje za razvoj, večina state-of-the-art modelov pride na orodje TensorFlow z manjšo zamudo.

TensorFlow ima jedro napisano v C/C++, ter v CUDA (tudi s knjižnico CuDNN) za GPU, medtem ko ima vmesnik implementiran za programska jezika Python in C++. Zaradi enostavne uporabe vmesnika Python omogoča hiter in enostaven prototipni razvoj metod. K temu pripomore tudi že obstoječa knjižnica za vizualizacijo (TensorBoard), ki omogoča hiter pregled nad delovanjem mreže.

Poglavitna prednost je tudi v odlični podpori za porazdeljeno procesiranje. Orodje TensorFlow podpira tako porazdelitev na več enot GPU znotraj enega računalnika, kot tudi enot porazdeljenih po več računalnikih. Pri tem se lahko računski model razdeli vertikalno (vsak nivo je svojanaprava) ali horizontalno (vsaka naprava je odgovorna za

določen paket primerov). Slabost orodja TensorFlow je predvsem v zapletenosti uporabe, ter v počasni implementaciji računskega grafa kadar je implementiran le kot model v jeziku Python.

Orodje TensorFlow uradno podpira korporacija Google zato ima zelo dobro urejeno podporo, v zadnjem času pa se izboljšuje tudi dokumentacija, ki je bila v preteklosti pomanjkljiva. Z vidika integracije ima TensorFlow primaren proces prevajanja preko orodja Python PIP, kjer pa lahko pride do konflikta pri odvisnosti od ostalih Pythonovih knjižnic, ki jih TensorFlow pripelje zraven. Temu se je mogoče izogniti z neodvisno namestitvijo, bodisi v ločene datoteke (virtualen), bodisi s slikami Docker. V osnovi orodje TensorFlow zahteva le manjše število odvisnih tretjih knjižnic: Python, six, NumPy, SciPy, protobuf, werkzeug, wheel, Eigen (interno prevajanje). Omogoča pa proces prevajanja preko različnih orodij za avtomatizirano prevajanje, uradno preko orodij Python PIP, bazel in conda, in neuradno preko CMake. Uradna podpora obstaja za operacijske sisteme Linux, MacOS in Windows, pri čemer je za sistem Windows zahtevan Python 3.5. Obsežen postopek prevajanja lahko zaplete integracijo v tretje aplikacije, če pa se integrirata le C++ vmesnik in jedro, je to mogoče narediti enostavno preko orodja bazel, ki prevede TensorFlow v deljeno knjižnico (ang. shared library), odpade pa tudi večina odvisnosti od knjižnic Pythona.

2.7 Keras

V nasprotju z ostalimi orodji je Keras [6] le programski ovoj nad obstoječim orodjem, ki podpira tenzorske operacije preko računskega grafa. Keras omogoča uporabo orodja Theano ali TensorFlow kot glavnega jedra, pri tem pa z enostavnejšim vmesnikom Python močno poenostavi razvoj. Keras ima pripravljene obstoječe nivoje, kot so konvolucijski nivo, polno povezani nivo, rekurzivni nivo, nivoji batch normalization, ReLU, PReLU, itd. Pri tem omogoča zelo enostavno združevanje nivojev za uporabo v globokem učenju in je primeren za uporabo pri navadnih konvolucijskih mrežah kot tudi rekurzivnih mrežah. Keras omogoča tudi enostavno vizualizacijo, poleg tega pa omogoča izpis podatkov v obliki primerni za TensorBoard, s čimer se odpre vizualizacija z orodji TensorFlow.

Orodje podpira tudi nekaj popularnih že naučenih modelov, kot so VGG, ResNet in Inception, obstajajo pa tudi neuradna orodja, s katerimi je mogoče pretvoriti modele iz orodja Caffe v modele primerne za orodje Keras. Glede na dokumentacijo ne obstaja uradni postopek za uporabo pred-naučenih modelov iz orodja TensorFlow neposredno v orodju Keras.

Ker predstavlja Keras le dodaten vmesnik v jeziku Python ni primeren za integracijo v tretje aplikacije. Če pa se uporablja orodje TensorFlow, je možno modele narejene v okolju Keras izvoziti neposredno v modele za orodje TensorFlow, s čimer je mogoče model uporabiti v vmesniku C++ do orodja TensorFlow.

3. ZAKLJUČEK

Izbira ustreznega orodja bo predvsem odvisna od specifičnih zahtev bodisi z raziskovalnega bodisi z integracijskega vidika industrijskih aplikacij. V primeru le raziskovalne uporabe je zelo primerno orodje MatConvNet, saj omogoča hiter proces prevajanja in enostavno uporabo v programskem okolju MATLAB. Slabost je v zelo omejenem dostopu do najaktualnejših modelov in pristopov, ter v nezmožnosti integracije v industrijske aplikacije. Podoben problem velja tudi za orodje Theano, ki ima sicer zaradi splošne namenskosti večjo uporabnost, vendar ima zelo omejen nabor pred-naučenih modelov in nezmožnosti integracije v industrijske aplikacije.

V primeru zahteve po uporabi najnovejših pristopov so primerna orodja Caffe, Torch in TensorFlow, pri čemer ima orodje Caffe največji nabor najnovejših modelov za konvolucijske mreže, saj večina raziskovalcev razvija prav v tem orodju. Slabost orodja Caffe je v omejenosti na konvolucijske nivoje, v zahtevi po ročni implementaciji novih nivojev v jezikih C++ in CUDA, ter v pomanjkljivi modularni zasnovi, ki otežuje integracijo. V primeru, da so najaktualnejši modeli narejeni s standardnimi nivoji, je vse modele mogoče z manjšim naporom pretvoriti iz modelov orodja Caffe v modele za orodje Torch ali TensorFlow, zato sta obe orodji primerni z raziskovalnega vidika. Glavna prednost orodja Torch je ogromna zbirka funkcionalnosti, ki se razteza od običajnih globokih mrež do generativnih mrež (ang. variational autoencoders), in do splošnih funkcij strojnega učenja (SVM, PCA, LDA, itd.). Slabost orodja Torch je predvsem kompleksen postopek integracije v industrijske aplikacije, pa tudi manj znani jezik vmesnika (Lua).

V nasprotju z orodji Caffe in Torch ima TensorFlow najmanjšo podporo za prednaučene najaktualnejše modele, ker obstajajo orodja za pretvarjanje modelov iz orodja Caffe v modele za orodje TensorFlow, ostane glavna slabost le pri modelih, ki ne uporabljajo standardnih nivojev kot so npr. Faster R-CNN. Poglavitna prednost orodja TensorFlow je v že-vgrajenih vizualizacijskih orodjih, v dokaj veliki zbirki pomembnih nivojev globokega učenja ter v enostavni uporabi vmesnika Python tudi s knjižnico Keras. Prav tako omogoča orodje TensorFlow enostavno uporabo modelov preko vmesnika C++, ima podporo na različnih strojnih arhitekturah (Android, GPU, CPU itd), ter edini omogoča porazdeljeno učenje preko različnih računalnikov in različnih grafičnih kartic. Tudi z vidika integracije v ločene aplikacije ima orodje TensorFlow še najboljšo podporo za integracijo z neposrednim vmesnikom preko jezika C++ in možnostjo prevajanja v ločeno knjižnico, kar močno olajša integracijo v aplikacije in kasnejše vzdrževanje aplikacije.

LITERATURA

1. Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, "Caffe: Convolutional Architecture for Fast Feature Embedding", *arXiv preprint arXiv:1408.5093*, 2014, Online: <http://caffe.berkeleyvision.org>

2. Torch, a scientific computing framework for LuaJIT, Online: <http://torch.ch>
3. Theano Development Team, "Theano: A Python framework for fast computation of mathematical expressions", *arXiv e-prints, abs/1605.02688*, 2016, Online: <http://arxiv.org/abs/1605.02688>, Online: <http://deeplearning.net/software/theano/>
4. A. Vedaldi and K. Lenc, "MatConvNet - Convolutional Neural Networks for MATLAB", In *Proceeding of the ACM Int. Conf. on Multimedia*, 2015, Online: <http://www.vlfeat.org/matconvnet/>
5. Tang, Yuan. "TF.Learn: TensorFlow's High-level Module for Distributed Machine Learning." *arXiv preprint arXiv:1612.04251* (2016), Online: <https://www.tensorflow.org>
6. F. Chollet, "Keras.", In *GitHub*, 2015, Online: <https://github.com/fchollet/keras>
7. Caffe Model Zoo, Online: <https://github.com/BVLC/caffe/wiki/Model-Zoo>

NANOEYE – PROGRAM ZA SEMI-AVTOMATIČNO ANALIZO SLIK

Janez Langus, Tomaž Šuštar

C3M d.o.o., Tehnološki park 21, 1000 Ljubljana

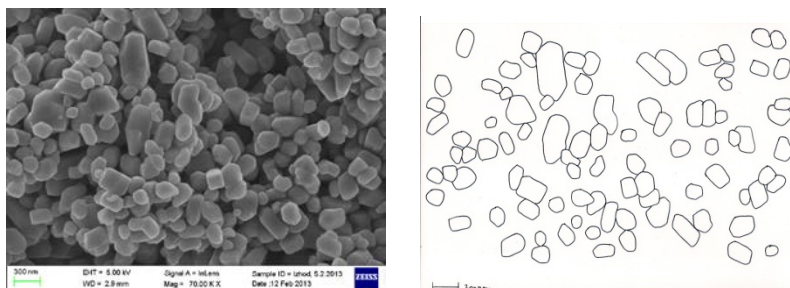
E-pošta: info@c3m.si

URL: <http://www.c3m.si>

POVZETEK: V prispevku predstavljamo program NanoEye, ki smo ga razvili v sodelovanju s podjetjem Cinkarna Celje d.d. Program je prvenstveno namenjen semi-avtomatični analizi mikroskopskih slik nano delcev titanovega dioksida, vendar zaradi modularne zasnove omogoča hitro modifikacijo za analizo drugačnega tipa materiala. Predstavili bomo nekaj primerov do sedaj razvitih modulov za analizo mikroskopskih slik praškastih materialov, metalnih domen in bioloških tkiv.

1. UVOD

C3M je visoko tehnološko podjetje, ki se že od leta 1992 ukvarja z razvojem specializiranih programskih rešitev za slovenska in tuja podjetja. Prvenstveno se ukvarjamo z razvojem specializiranih simulacijskih rešitev po metodi končnih elementov, ki jim moramo za uspešno rešitev problema povezovati z drugimi metodami (metoda diskretnih elementov, metoda končnih volumnov, optimizacijske zanke...). Zaradi sistematičnega pristopa k reševanju problemov smo razvili znanje in orodja, ki nam omogočajo učinkovit razvoj aplikacij na področjih modeliranja materialov, preoblikovanja kovin, biomehanike, vesoljskih aplikacij, kemijskega inženiringa, farmacevtskih aplikacij in analize slik [1].



Slika 1: Primer mikroskopske slike vzorca nano delcev titanovega dioksida (levo) in pripadajočega ročnega obrisovanja oblik delcev (desno).

S strani Cinkarne Celje d.d. smo bili postavljeni pred zanimiv izziv analize oblike nano delcev [2] na podlagi slike, zajete na elektronskem mikroskopu (slika 1). Do tedaj razvita metodologija je temeljila na ročnem obrisovanju delcev in zaradi dolgotrajnosti ni bila primerna za praktično uporabo v namen kontrole kvalitete titanovega dioksida.

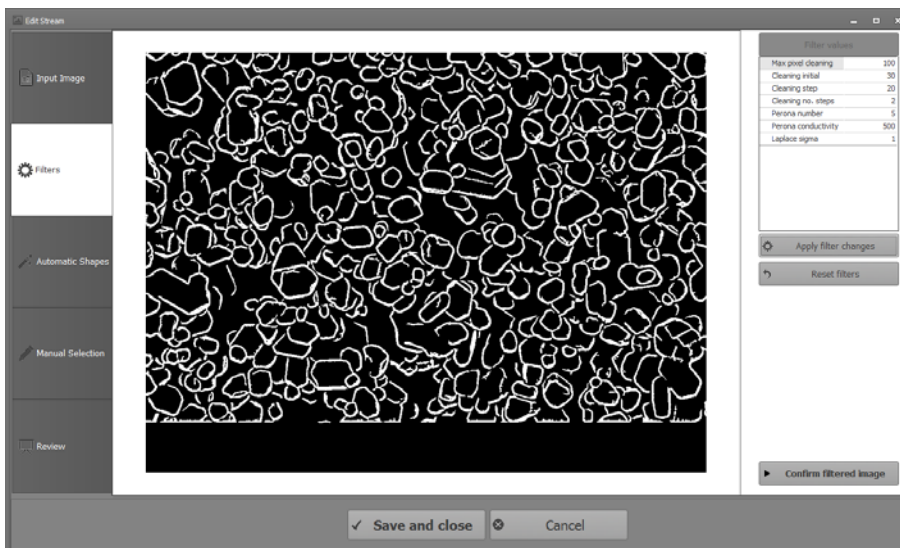
Glavni izziv pri avtomatizaciji detekcije oblik nano delcev je prekrivanje (slika 1, levo), ki je posledica aglomeracije vzorca. Z natančnejšo in dolgotrajnejšo pripravo vzorca se da doseči koplanarno distribucijo nano delcev pred slikanjem, vendar ta priprava zahteva čas visoko usposobljenega laboratorijskega osebja.

2. PROGRAM NANOEYE

V sodelovanju z razvojnim oddelkom podjetja Cinkarna Celje d.d. smo sočasno razvijali numerično rešitev za avtomatično določevanje oblik nano delcev na podlagi mikroskopske slike in grafični uporabniški vmesnik, ki je prilagojen delovnemu procesu od zajema slike do prikaza rezultatov. V nadaljevanju bomo na kratko predstavili potek semi-avtomatične analize mikroskopskih slik s programom NanoEye.

2.1 Filtracija slike - določevanje robov

Po določitvi protokola slikanja (predpisane nastavitve elektronskega mikroskopa) smo razvili specializirano kombinacijo obstoječih filtrov in namensko razvitih filtrov podjetja C3M, ki stabilno deluje na večjem naboru referenčnih slik nano praškastega titanovega dioksida. Končni rezultat uporabniku prikazemo na zaslonu, kjer mu z nastavitvami vrednosti parametrov posameznih filtrov damo možnost modifikacije procesa filtracije.

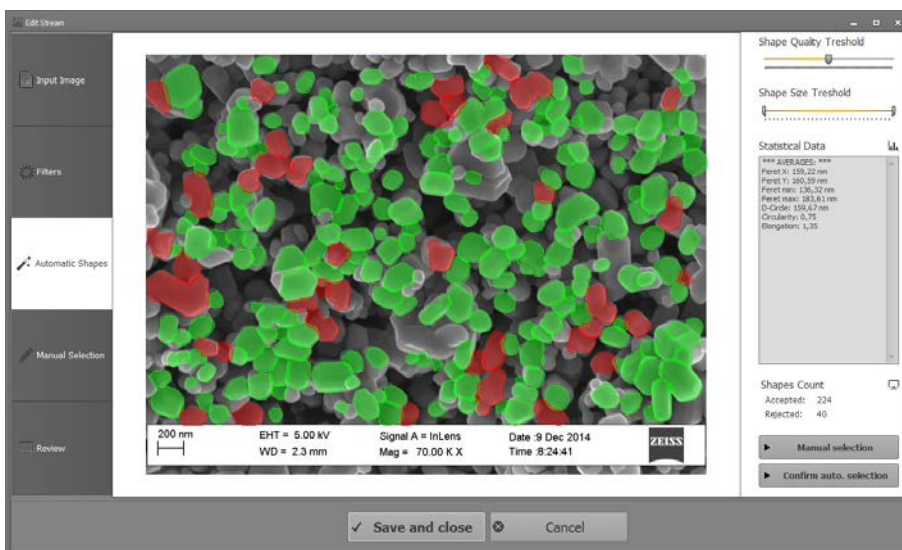


Slika 2: Rezultat filtracije mikroskopske slike

2.2 Iskanje oblike delcev

Kljub razviti kombinaciji filtrov so izluščeni robovi nano delcev le v redkih primerih zaključene krivulje, kar predstavlja naslednji izziv pri določevanju oblik delcev. Razvili smo namensko metodo, ki simulira napihovanje mehurčkov znotraj praznih (črnih) polj filtrirane slike. Razširjanje »opne« mehurčka je omejeno z navidezno elastičnostjo, ki zaustavi razširjanje mehurčka v primeru manjših prekinitvev detektiranega roba.

Iz nabora napihnjenih mehurčkov izluščimo oblike, ki so glede na geometrijske kriterije (velikost, pravilnost oblike,...) najverjetnejši kandidati za pravilno prepoznane delce in jih na zaslonu obarvamo zeleno (slika 3). Ostale prepoznane oblike obarvamo rdeče in jih ne vključimo v statistično analizo.



Slika 3: Avtomatično izbrane oblike nano delcev

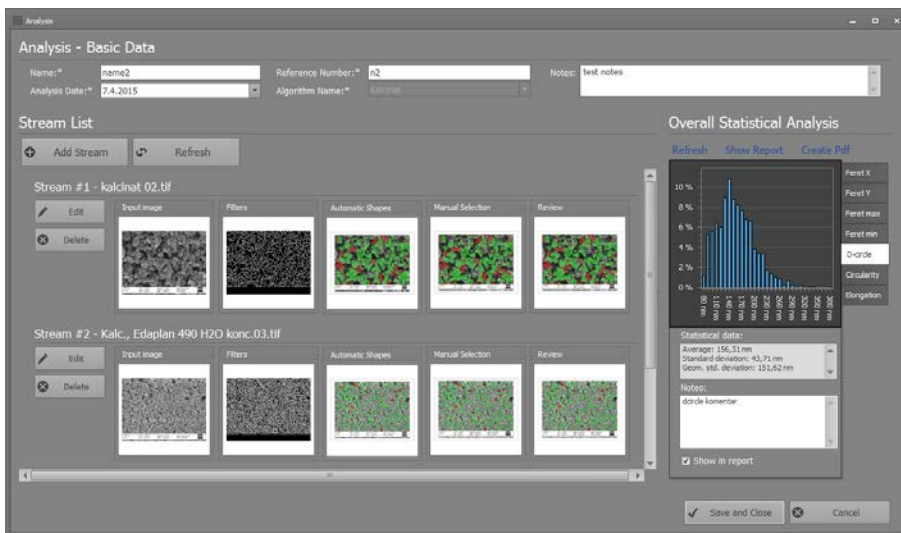
V primeru, ko se uporabnik ne strinja z avtomatičnim izborom oblik nano delcev, lahko s klikom na obliko spremeni njen status (zeleno obliko spremeni v rdečo in jo s tem izloči iz statistične analize in obratno, rdečo obliko spremeni v zeleno in jo vključi v statistično analizo).

2.3 Izboljševanje kvalitete statistike

Ker je število prepoznanih delcev na posamezni sliki relativno majhno (okoli 100 pri 70.000-kratni povečavi) je za kvalitetnejšo statistično analizo priporočljivo narediti prepoznavo nano delcev na večjem številu slik določenega vzorca zajetih na različnih mestih.

V programu NanoEye je analiza posameznega vzorca organizirana po načelu vzporednih tokov (stream), ki se na koncu združijo v skupno statistično analizo. Vsak slikovni tok je razdeljen v pet delovnih korakov (uvoz slike, filtracija, avtomatična detekcija oblik,

ročni popravek izbora in pregled končnega rezultata). Ko uporabnik na končnem koraku potrdi pravilnost podatkov za trenutni slikovni tok, se podatki iz tega toka prenesejo v skupni nabor detektiranih oblik za določen vzorec.



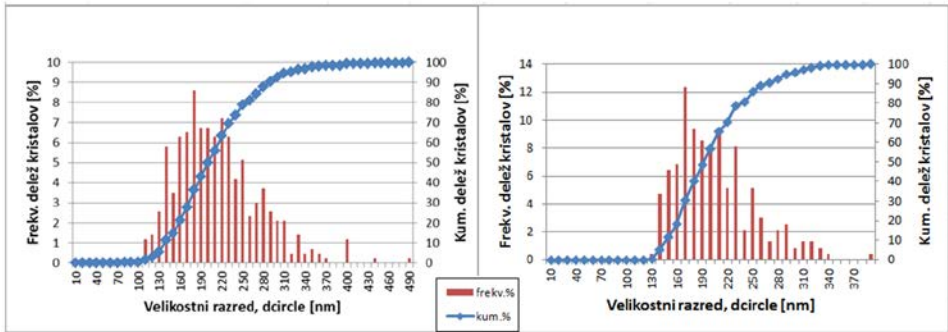
Slika 4: Vzoredni podatkovni tokovi v programu NanoEye

Skupni seznam detektiranih oblik je osnova za statistično analizo nano delcev. Naročnik programa NanoEye je želel dobiti povprečno vrednost, standardno deviacijo in porazdelitveno krivuljo za sedem geometrijskih parametrov ki se tipično uporabljajo za opisovanje praškastih delcev. V primeru drugačnega tipa materiala lahko program NanoEye prilagodimo tako, da izpisuje statistiko poljubnih parametrov detektiranih oblik.

Rezultate statistične analize skupnega nabora detektiranih oblik uporabnik z enim klikom izvozi v poročilo v obliki PDF. Za namene naprednejših statističnih analiz se v tekstovno datoteko izvozi tudi seznam geometrijskih parametrov za vse detektirane oblike.

2.4 Primerjava avtomatične in ročne detekcije delcev

Pravilnost delovanja avtomatične detekcije nano delcev smo tekom razvoja primerjali z rezultati ročnega očitovanja. Kombinacija predpisanega protokola slikanja vzorca, filtracije slike, procesa napihovanja in filtracije dobljenih oblik glede na geometrijske lastnosti privede do dobrega ujemanja med ročno in avtomatično metodo (slika 5).



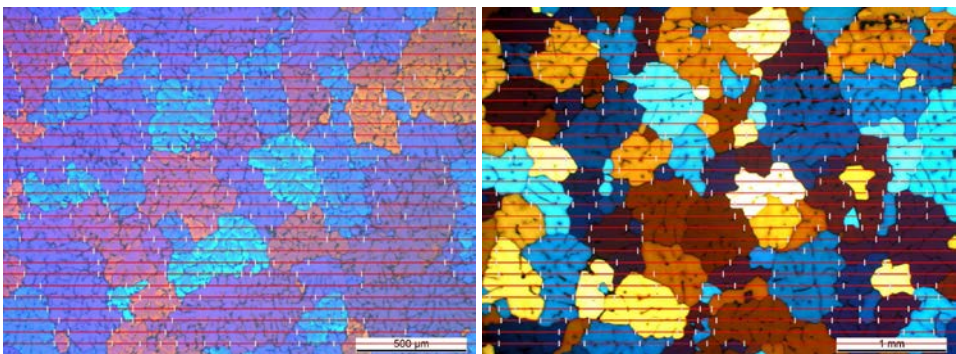
Slika 5: Primerjava rezultatov ročnega (levo) in avtomatičnega štetja (desno) na primeru parametra D_{circle} (ekvivalentni radij kroga z enako površino)

3. PRIMERI UPORABE

V stiku s partnerji iz industrije in raziskovalnih institucij smo ugotovili, da obstaja širok nabor problemov, pri katerih se uporabniki trenutno poslužujejo zamudne ročne analize slik. Za nekaj primerov smo razvili numerične module za avtomatizacijo procesa.

3.1 Detekcija metalnih domen

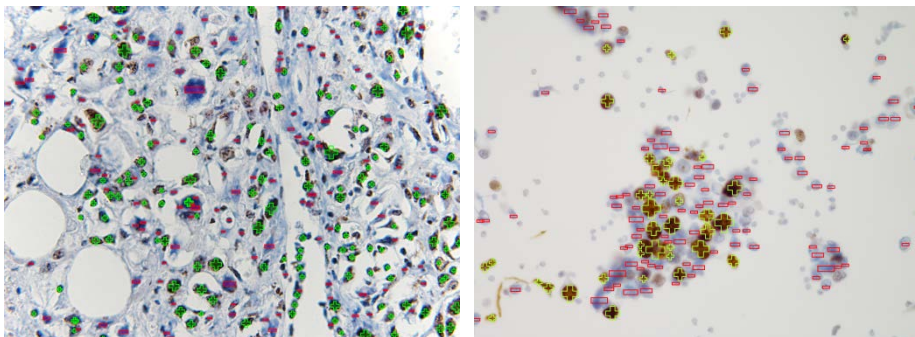
V kovinski industriji je metalografija zelo uveljavljena metoda za določanje kvalitete preiskovanega vzorca. Na sliki 6 je prikazan primer mikroskopske slike površine aluminijeve zlitine. Razvili smo avtomatizirano metodo, ki po metodi ASTM E112 določi povprečno presečno velikost domene. Primerjava z ročno določenimi povprečnimi velikostmi domen je pokazala dobro delovanje razvite avtomatične analize slike.



Slika 6: Rezultat avtomatizirane metode ASTM E112 za primer površine dveh vzorcev aluminijeve zlitine

3.2 Biološki vzorci

Razvili smo numerično strategijo za identifikacijo pozitivno in negativno obarvanih celičnih jeder (slika 7), ki raziskovalcem na področju onkologije in citopatologije pomaga pri določevanju deleža pozitivno obarvanih celic v določenem vzorcu. To razmerje je indikator stopnje obolenosti določenega vzorca in je eden glavnih kazalcev pri spremljanju učinkovitosti metod zdravljenja rakavih obolenj.



Slika 7: Detekcija obarvanosti celičnih jeder v tkivu (levo) in v emulziji (desno).

4. ZAKLJUČEK

V programu NanoEye implementirani proces dela z mikroskopskimi slikami je možno aplicirati na številna področja, zato smo pri načrtovanju uporabniškega vmesnika velik poudarek namenili modularni zasnovi. Z zamenjavo numeričnega modula (proces filtracije in strategija iskanja oblik na filtrirani sliki) lahko strankam ponudimo orodje za analizo mikroskopskih slik, ki je prilagojeno njihovemu specifičnemu problemu.

Modularna zasnova numeričnega modula nam omogoča, da razvite module uporabimo tudi ločeno od programa NanoEye v popolnoma avtomatiziranih postopkih (in-line nadzor kakovost npr.).

LITERATURA

1. <http://www.c3m.si/#solutions>
Informacije o razvojnih numeričnih orodjih podjetja C3M.
2. <http://www.cinkarna.si/si/izdelki/ultrafini-titanov-dioksid>
Informacije o produktu titanov dioksid

DETEKCIJA VELIKEGA ŠTEVILA KATEGORIJ PROMETNIH ZNAKOV

Peter Uršič, Domen Tabernik, Rok Mandeljc, Danijel Skočaj

Univerza v Ljubljani

Fakulteta za računalništvo in informatiko

E-pošta: danijel.skocaj@fri.uni-lj.si

URL: <http://www.vicos.si>

POVZETEK: *Detekcija prometnih znakov je dobro raziskano področje znotraj skupnosti računalniškega vida. Navkljub temu se velika večina obstoječih pristopov osredotoča na detekcijo relativno majhnega števila kategorij prometnih znakov. V tem članku apliciramo metodo globokega učenja Faster R-CNN na detekcijo velikega števila kategorij prometnih znakov. V ta namen smo ustvarili zbirko slik s 160 različnimi kategorijami prometnih znakov zajetih na cestah različnih občin v Sloveniji. Na tej zbirki pokažemo odlične preliminarne rezultate z izbrano metodo globokega učenja ter nakažemo na slabosti in potencialne izboljšave izbrane metode za to specifično aplikacijo.*

1. UVOD

Detekcija in razpoznavanje prometnih znakov sta precej pogosto naslovljena problema na področju računalniškega vida. Namen detekcije je na slikah poiskati lokacije in velikosti prometnih znakov, medtem ko je cilj razpoznavanja klasificirati te znake v eno izmed kategorij. Zelo veliko pristopov se ukvarja zgolj z detekcijo [3,4,5] ali samo z razpoznavanjem prometnih znakov [6,7,8], nekaj pa jih naslavlja tudi oba problema hkrati [9,10,11].

Detekcija prometnih znakov se zdi ključna tehnologija v asistenčnih sistemih za pomoč voznikom ali v avtonomnih vozilih, zato je za te namene na voljo več rešitev [12,13]. Rezultati, ki jih avtomatski sistemi za detekcijo takšnih prometnih znakov dosegajo so odlični. Bi lahko torej rekli, da je problem detekcije prometnih znakov rešen? Odgovor je negativen. Ti sistemi namreč delujejo zelo uspešno na omejem številu (okrog 50 ali celo manj) najpomembnejših kategorij z majhno variabilnostjo izgleda in nespremenljivo obliko. Obstoječi sistemi pa niso primerni za detekcijo velikega števila kategorij prometnih znakov, vključno z zelo spremenljivimi prometnimi znaki.

Prav to pa je problem, ki ga naslavljam v tem članku. Cilj je zgraditi učinkovit sistem za detekcijo velikega števila prometnih znakov. V ta namen smo v sodelovanju s podjetjem DFG Consulting d.o.o. zgradili bazo prometnih znakov s 160 kategorijami. Pri

tem je 123 kategorij standardnih, z nespremenljivim razmerjem velikosti stranic in relativno majhno do srednjo variabilnostjo vsebine, medtem, ko je 37 razredov precej bolj zahtevnih, saj se razmerja velikosti stranic prometnih znakov spreminjajo, ravno tako je lahko zelo različna tudi sama vsebina prometnega znaka, kot prikazuje slika 1. Vse večje javno objavljene zbirke slik, na katerih raziskovalci evalvirajo svoje algoritme, vsebujejo izključno znake iz prve podmnožice standardnih znakov [3,8,14,15,16]. Naš cilj je torej s pomočjo te zbirke izdelati učinkovit sistem za detekcijo velikega števila kategorij prometnih znakov s konkretno aplikacijo avtomatizacije procesa za vzdrževanje evidence vertikalne prometne signalizacije. V tem članku bomo predstavili osnovno metodologijo za reševanje tega problema in preliminarne rezultate.



Slika 1: Zbirka slik (levo); standardni razredi (zgoraj) in zahtevni razredi (spodaj). Tipična slika ter variabilnost nekaterih standardnih in zahtevnih razredov (desno).

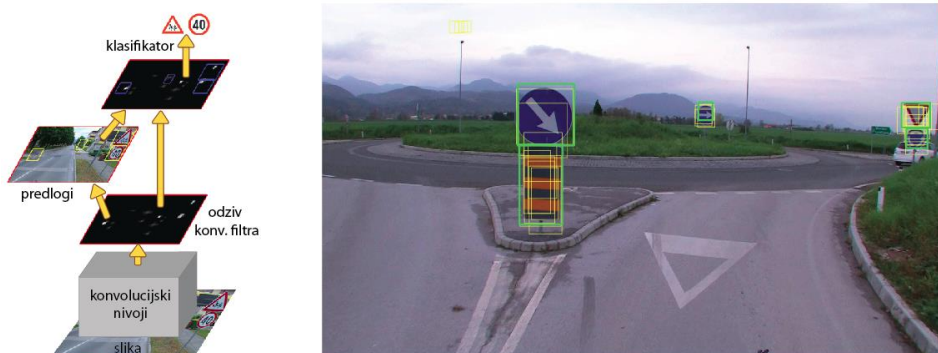
2. METODA

Problem detekcije in razpoznavanja prometnih znakov rešujemo z metodo, ki oba podproblema naslavlja s poenoteno globoko mrežo Faster R-CNN [2]. Pristop Faster R-CNN združuje dve glavni komponenti. Prva komponenta je globoka, polno-

konvolucijska nevronska mreža za detekcijo predlogov, oz. potencialno zanimivih področij v vhodni sliki (t.i. Region Proposal Network, RPN). Poleg očitanih pravokotnikov, ki omejujejo detektirana področja, mreža za vsako področje izračuna tudi oceno zanimivosti. Druga komponenta je uveljavljena mreža Fast R-CNN [1], ki podana območja v sliki klasificira v končne razrede, oziroma jih zavrže kot del ozadja. Mreža Fast R-CNN za vsa predlagana področja hkrati izračuna odzive konvolucijskih nivojev, s čimer dosež visoko računsko učinkovitost. Slednja je še nadaljnje izboljšana s souporabo (deljenjem) konvolucijskih nivojev med mrežo RPN in Fast R-CNN, kot je ponazorjeno na sliki 2, levo.

Učenje mreže Faster R-CNN poteka znotraj poenotenega optimizacijskega ogrodja, in sicer v štirih korakih, znotraj katerih na ciljni zbirki izmenično učimo na zbirki ImageNet pred-naučeni mreži RPN in Fast R-CNN. Pri tem obe mreži najprej učimo ločeno, v zadnjih dveh korakih pa združimo njune konvolucijske nivoje in učimo samo zgornje nivoje obeh mrež. Razpoznavanje prometnih znakov torej poteka tako, da se na sliki najprej poiščejo predlogi potencialno zanimivih regij (označeni z rumenimi pravokotniki na sliki 2, desno), nato pa se na teh regijah izvede razpoznavanje, kar po odstranjevanju lokalnih nemaksimumov privede do končnega rezultata, na vzorčni sliki označenega z zelenimi pravokotniki.

Za učenje globokih mrež potrebujemo število vzorcev, ki presega velikost razpoložljive zbirke vzorcev prometne signalizacije. Zato smo uporabili postopek sintetične izgradnje dodatnih učnih primerov, pri čemer smo za osnovo uporabili segmentirane primere znakov iz učne množice. Tako dobljene predloge smo nato popačili z različnimi geometrijskimi preslikavami (afine preslikave, spremembe v skali) ter popačitvami izgleda (svetlost in kontrast, megljenje, spremenljiva osvetlitev). Iz vsake predloge smo na ta način pridobili dodatnih dvesto primerov ter jih vstavili v slike, za katere smo predhodno ugotovili, da ne vsebujejo prometnih znakov.



Slika 2: Shema poenotene mreže za detekcijo in razpoznavanje Faster R-CNN (levo). Primer detekcije predlogov regij in klasifikacije predmetov (desno).

3. EKSPERIMENTALNI REZULTATI

Model za detekcijo prometnih znakov smo naučili in evalvirali na zbirki slik prikazani na sliki 1. Zbirka vsebuje 6165 slik velikosti 1920x1080 slikovnih elementov, ki jih je podjetje DFG Consulting posnelo na cestah v šestih slovenskih občinah. Slike smo razdelili na učno (4680 slik) in testno množico(1485 slik), pri čemer smo poskrbeli, da se noben primerek prometnega znaka ne pojavi hkrati v obeh množicah. Na vsaki sliki je bil vsaj en prometni znak, lahko pa jih je bilo tudi več. Tako smo imeli v učni množici 8.332, v testni množici pa 3.193 primerkov. V vsaki kategoriji je bilo vsaj 20 primerov prometnega znaka. Zbirka vsebuje 160 kategorij, ki smo jih nadalje razdelili na 123 kategorij standardnih znakov in 37 kategorij zahtevnejših znakov. V učni fazi smo model naučili na celotni učni množici vseh kategorij, v testni fazi pa smo evalvirali tudi obe podmnožici posebej.

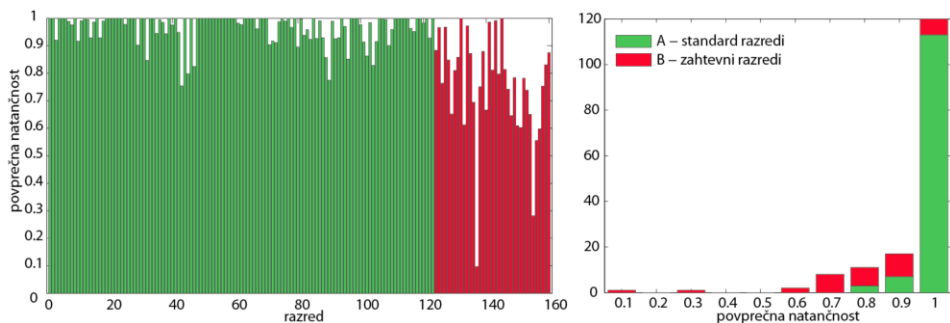
Uspešnost smo merili s srednjo povprečno natančnostjo (ang. mean average precision, mAP), ki izraža srednjo vrednost povprečnih natančnosti po vseh razredih. Povprečna natančnost za en razred se izračuna kot integral natančnosti po vseh priklicih, pri čemer nam natančnost (ang. precision), pove kakšen delež detektiranih znakov je pravilen, medtem ko priklic (ang. recall) predstavlja delež znakov, ki smo jih uspeli detektirati.

Rezultati so prikazani v tabeli 1, ki prikazuje na testnih slikah izmerjene vrednosti srednjih povprečnih natančnosti za različne podmnožice kategorij prometnih znakov in pri različnih omejitvah minimalne velikosti prometnih znakov (30, 50 in 100 slikovnih elementov). Posamezne vrednosti povprečnih natančnosti so prikazane na sliki 3, levo, njihov histogram pa na sliki 3, desno, pri čemer so rezultati za podmnožico standardnih znakov obarvani zeleno, rezultati zahtevnih razredov prometnih znakov pa z rdečo, podani pa so za minimalno velikost znakov 30 slikovnih elementov.

Rezultati so zelo obetavni. Na celotni množici razviti postopek preseže pri obravnavi vseh prometnih znakov z velikostjo vsaj 30 slikovnih elementov srednjo povprečno natančnost 92%, pri čemer le-ta doseže na podmnožici standardnih prometnih znakov skoraj 97%. Slabše se v tem primeru odreže na podmnožici zahtevnih prometnih znakov, kjer sodeže 76% mAP. Predvsem se metoda slabše obnese na detekciji manjših prometnih znakov, kar je razvidno iz izboljšave rezultatov v primerih, ko majhne prometne znake ne obravnavamo. Čase omejimo na prometne znake večje od 100 slikovnih elementov so rezultati že zelo dobri.

Tabela 1: Rezultati detekcije za vse razrede skupaj ter za dve podmnožici izraženi s srednjo povprečno natančnostjo (mAP) dobljeno za različne minimalne velikosti prometnih znakov.

razredi	30px	50px	100px
vseh 160 razredov	92,04	95,66	98,39
podmnožica A (standardnih 123 razredov)	96,85	98,60	99,46
podmnožica B (zahtevnih 37 razredov)	76,05	85,90	94,82



Slika 3: Povprečna natančnost posameznih razredov (levo) in histogram povprečnih natančnosti vseh razredov (desno).

Na sliki 4 je prikazanih nekaj primerov detekcij. Zeleni pravokotniki ponazarjajo uspešne detekcije zelo različnih kategorij prometnih znakov. Vidimo, da je velika večina detekcij uspešnih in sicer za različne kategorije prometnih znakov od enostavnih znakov za omejitev hitrosti do kompleksnih usmerjevalnih tabel. Prikazane pa so tudi nekatere napake. Rdeči pravokotniki ponazarjajo napačne detekcije (ang. false positives). Na srednji spodnji sliki je bil detektiran znak za nevarnost s hrbtni strani zaradi izrazite oblike in podobnosti izgleda z znakom za splošno nevarnost. Podobno je bilo na desni spodnji sliki s hrbtni strani detektirano ogledalo. Na isti sliki je bila napačno razpoznana kategorija dopolnilne table. Dopolnilne table so še posebej problematične, ker so različne kategorije po obliki zelo podobne, razlikujejo se v glavnem po vsebini. Na zgornji desni sliki ni bila detektirana usmerjevalna tabla - najbrž zaradi premajhne velikosti. Je pa takih napak relativno malo, kot dokazujejo kvantitativni rezultati v tabeli 1. Z nadgradnjo razvitega modela ter povečanjem števila učnih primerov bomo lahko te rezultate še izboljšali.



Slika 4: Nekaj primerov detekcije prometnih znakov. Zeleni očitani pravokotniki označujejo uspešne detekcije, rdeči napačne detekcije, vijolični pa nedetektirane prometne znake.

4. ZAKLJUČEK

V članku smo obravnavali problem detekcije velikega števila kategorij prometnih znakov. V ta namen smo zgradili bazo s 160 kategorijami prometnih znakov. V nasprotju s sorodnimi podatkovnimi zbirkami smo v našo poleg 123 kategorij prometnih znakov brez velike variabilnosti oblike ali vsebine vključili tudi 37 kategorij z velikimi spremembami znotraj kategorije, bodisi po obliki bodisi po vsebini ali videzu. Za metodo detekcije smo nato uporabili detektor Faster R-CNN [1], ki je osnovan na metodah globokega učenja in v poenotenem okvirju generira predloge zanimivih regij na sliki in jih nato tudi klasificira. Z eksperimentalno evalvacijo smo pokazali, da metoda Faster R-CNN omogoča detekcijo 160 kategorij znakov z 92% srednjo povprečno natančnostjo (mAP), pri obravnavi prometnih znakov velikosti 30 slikovnih elementov in več, oz. celo 98,4%, če se omejimo samo na znake z velikostjo vsaj 100 slikovnih elementov. Pri tem je metoda veliko bolj uspešna na podмноžici 23 kategorij standardnih znakov, medtem ko na podмноžici 37 kategorij zahtevnejših znakov dosega nekoliko slabše rezultate. Z vidika avtomatizacije procesa za vzdrževanje evidence vertikalne prometne signalizacije se lahko zadovoljimo z detekcijo le večjih znakov, saj je dovolj le ena detekcija znaka na sliki zajeti tik pred prometnim znakom, ko je le-ta na sliki največji .

Navkljub odličnim preliminarnim rezultatom obstaja še veliko možnost za izboljšave. Za prilagoditev algoritma detekciji manjših prometnih znakov bi bila potrebna povečava ločljivosti slik, bodisi že ob učenju, bodisi le ob detekciji. Na drugi strani pa bi lahko izboljšavo dosegli tudi s popravki globoke mreže, ki premočno zmanjša ločljivost značilnic na višjih nivojih. V prihodnje nameravamo tudi izboljšati zbirko slik, jo povečati na vsaj 200 kategorij ter povečati število primerov na kategorijo. Seveda pa bi veljalo zbirko obogatiti še s prometnimi znaki zajetimi v neugodnih vremenskih pogojih, kot so megla ter dežali sneg, ter tako še povečati zahtevnost detekcije in jo še približati uporabi tudi v najbolj zahtevnih realnih pogojih.

LITERATURA

1. R. Girshick. "Fast R-CNN." In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1440-1448. 2015.
2. R. Shaoqing, K. He, R. Girshick, and J. Sun. "Faster R-CNN: Towards real-time object detection with region proposal networks." In *Advances in neural information processing systems*, pp. 91-99. 2015.
3. S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel. "Detection of traffic signs in real-world images: The german traffic sign detection benchmark." In *IJCNN*, pages 1–8, 2013.
4. J. Lillo-Castellano, I. Mora-Jimenez, C. Figueroa Pozuelo, and J. Rojo-Alvarez.

- "Traffic sign segmentation and classification using statistical learning methods." In *Neurocomputing*, 153:286–299, 2015.
5. G. Overett and L. Petersson. "Large scale sign detection using hog feature variants." In *Intelligent Vehicles Symposium*, pages 326–331, 2011.
 6. D. Ciresan, U. Meier, J. Masci, and J. Schmidhuber. "A committee of neural networks for traffic sign classification." In *IJCNN*, pages 1918–1921, 2011.
 7. M. Haloi. "A novel pLSAbased traffic signs classification system". *CoRR*, abs/1503.06643, 2015.
 8. J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel. "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition." In *Neural Networks*, 32:323–332, 2012.
 9. A. Ellahyani, M. E. Aansari, and I. E. Jaafari. "Traffic sign detection and recognition using features combination and random forests." In *IJACSA*, 7(1):6861– 6931, 2016.
 10. J. Greenhalgh and M. Mirmehdi. "Recognizing textbased traffic signs." In *Transactions on Intelligent Transportation Systems*, 16(3):1360–1369, 2015.
 11. Y. Zhu, C. Zhang, D. Zhou, X. Wang, X. Bai, and W. Liu. "Traffic sign detection and recognition using fully convolutional network guided proposals." In *Neurocomputing*, 214:758–766, 2016.
 12. G. Stein, O. Shachar, Y. Taieb, and U. Wolfovitz. "Detecting and recognizing traffic signs." *Patent*, US 8064643, 2014.
 13. L.-P. Becker. "Method and device for traffic sign recognition." *Patent*, US 2010/0283855 A1, 2014.
 14. R. Timofte, K. Zimmermann, and L. V. Gool. "Multiview traffic sign detection, recognition, and 3D localisation." In *WACV*, pages 1–8, 2009.
 15. V. Vukotić, J. Krapac, and S. Šegvič. "Convolutional neural networks for croatian traffic signs recognition." In *CCVW*, pages 15–20, 2014.
 16. F. Larsson and M. Felsberg. SCIA, "Using Fourier Descriptors and Spatial Models for Traffic Sign Recognition", In *SCIA*, pages 238–249. 2011.

MEDICINSKE IN BIOMEDICINSKE APLIKACIJE

OCENA NAPAKE PRI ANALIZI OBJEKTOV Z GEOMETRIJSKO MORFOMETRIJO

Franc Janžekovič¹, Ahmad Mohmoudi², Tina Klenovšek¹

¹Oddelek za biologijo

Fakulteta za naravoslovje in matematiko, Univerza v Mariboru

E-pošta: franc.janzekovic@um.si

²Department of Biology

Faculty of Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

POVZETEK: *Analiza variabilnosti bioloških objektov s postopki geometrijske morfometrije je v biologiji, medicini, forenziki in drugih vedah uporabna metoda za zaznavanje majhnih razlik v obliki in velikosti med njimi. Pri vrednotenju razlik med objekti moramo oceniti vpliv napak, ki nastajajo v posameznih korakih analize. Do napak prihaja pri zajemanju slikovnih vzorcev – napaka fotografiranja, pri izvajanju meritev – napaka meritve ter glede na število proučevanih vzorcev – napaka vzorca. Napake vrednotimo z analizo variance večkratnega fotografiranja, večkratnega izvajanja meritev ter analizo občutljivosti.*

1. UVOD

Morfometrija je kvantitativna metoda, katere namen je izmeriti velikost in obliko bioloških objektov oz. organizmov in njihovih struktur. Geometrijska morfometrija je skupina postopkov, ki obliko oz. geometrijo objektov opišejo s pomočjo oslonilnih točk v dvo- ali trirazsežnem prostoru, in nato informacije o relativnih prostorskih razmerjih med točkami ohranijo skozi celotno analizo. Omogočajo tudi ločeno analizo velikosti in oblike objektov in prikaz variabilnosti oblike objektov med ali znotraj skupin ter v odvisnosti od drugih spremenljivk. V biologiji se geometrijska morfometrija uporablja za opisovanje in statistično analizo variabilnosti v obliki organizmov med ali znotraj vzorcev in za analizo sprememb v obliki organizmov ali posameznih struktur zaradi ontogenetskih, evolucijskih, ekoloških in drugih dejavnikov. Poleg biologije je metoda uveljavljena v antropologiji, paleontologiji, medicini in sorodnih vedah [1], vendar se uporablja tudi za analizo nebioloških objektov.

V postopku priprave objektov, zajemanja meritev ter statistične analize na končni rezultat in sprejemanje zaključkov vplivajo vsaj trije viri šuma. Zajemanje fotografij je eden izmed ključnih pripravljalnih postopkov, napake, ki pri tem nastajajo, imenujemo *napaka fotografiranja*. Do napak prihaja tudi v fazi zajemanja meritev, zato govorimo o *napaki meritve*. Tretji sklop napak je posledica statističnega vrednotenja, njen vir pa je nezadostna velikost vzorca, zato govorimo o *napaki vzorca*. Vse tri sklope napak lahko

ovrednotimo, foto napako s ponovnim fotografiranjem, napako merjenja s ponovnim merjenjem ter napako vzorčenja z analizo občutljivosti (oz. analizo različno velikih vzorcev) (Tabela 1).

Namen našega prispevka je osvetliti vire napak pri analizi objektov z geometrijsko morfometrijo na primeru bioloških vzorcev: i – ocena napake fotografiranja, ii – ocena napake meritve ter iii – ocena napake velikosti vzorca. Postopke za vrednotenje in testiranje vseh treh vrst napak predstavljamo na primeru analize velikosti in oblike ventralne strani lobanje poljske voluharice. Osnove uporabe geometrijske morfometrije so pojasnjene v Priročniku [2].

Tabela 1: Viri in postopki vrednotenja napak v postopku analize z geometrijsko morfometrijo.

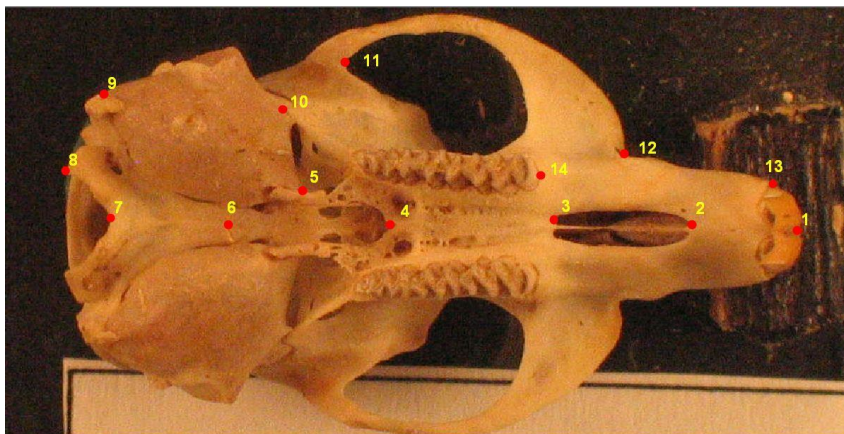
<i>NAPAKA</i>	<i>VIR NAPAKE</i>	<i>OCENITEV NAPAKE</i>	<i>ANALIZA</i>
Napaka fotografiranja	Zajemanje fotografij		Analiza variance med posameznimi serijami fotografiranja.
Napaka meritve	Izvedba meritev	Trikratna ponovitev meritev na isti fotografiji.	Analiza variance med posameznimi serijami meritev.
Napaka vzorca	Velikost vzorca	Zaporedna delitev osnovnega vzorca na manjše vzorce. Analiza občutljivosti.	Primerjava varianc med posameznimi vzorci.

2. NAPAKA FOTOGRAFIRANJA

Ventralna stran lobanje je bolj ali manj ploska struktura, kar nam omogoča, da sicer trirazsežen objekt analiziramo v dvorazsežnem prostoru, torej na fotografiji. Fotografije objektov so pogosto osnova za izvajanje meritev z določanjem točk. Vsi objekti vključeni v analizo morajo biti fotografirani pod enakimi ali čim bolj podobnimi pogoji. Fotograf zagotovi pravilno lego objekta na način, da objekt postavi v optično os in ga orientira tako, da je stran, ki nas zanima, pravokotna glede na optično os. Razdalja med fotoaparatom in objektom mora biti zadostna, da se izognemo napaki zaradi ukrivljenosti slikovnega polja. Napake najpogosteje izvirajo iz nepravilnega položaja objekta. Če obravnavana ploskev objekta ni v pravokotnem položaju glede na optično os, je torej objekt nagnjen. Če je objekt (lobanja) nagnjen po mediani ravnini (levo - desno), na preverjanju fotografij to opazimo kot nesimetričnost med levo in desno polovico sicer simetrične strukture. Če pa je objekt nagnjen po frontalni ravnini (naprej – nazaj), pa to težko ocenimo. V morfometrični praksi se navadno srečujemo z obema vidikoma nagiba, tako medianim kot frontalnim.

Napako fotografiranja ovrednotimo s ponovljenim fotografiranjem istih objektov v treh zaporednih serijah. Pri tem fotografiranje izvaja ista oseba v treh zaporednih dneh, da

dobimo fotografije treh časovnih serij istih objektov. Izvedemo meritve na vseh fotografijah in jih statistično ovrednotimo. Izmerjene vrednosti so odvisne spremenljivke, neodvisna spremenljivka pa je časovna serija fotografiranja. V našem primeru smo izmerili 13 oslonilnih točk na ventralni strani lobanje 31 osebkov poljske voluharice (slika 1).



Slika 1: Ventralna stran lobanje poljske voluharice s 13 oslonilnimi točkami.

Vrednosti meritev vključimo v Prokrustovo analizo variance (ANOVA) [5], pri čemer nas zanima razmerje med varianco znotraj skupin in med skupinami, formula 1 [3]. Rezultat izrazimo kot delež napake meritve - %ME (Measurement Error), formula 2 [3] (S^2_{among} – varianca med skupinami, S^2_{within} – varianca znotraj skupin, MS_{among} – povprečje kvadratov med skupinami, MS_{within} – povprečje kvadratov znotraj skupin, m – število ponovitev serij).

$$S^2_{among} = \frac{MS_{among} - MS_{within}}{m} \quad (1)$$

$$\%ME = \frac{S^2_{within}}{S^2_{within} + S^2_{among}} \cdot 100 \quad (2)$$

V geometrijski morfometriji velikost objektov (t.i. centroidna velikost) in obliko objektov (t.i. Prokrustove koordinate) obravnavamo ločeno, zato tudi izračun napak izvedemo ločeno za oba vidika variabilnosti. Na primeru 31 lobanj poljskih voluharic smo izvedli postopek Prokrustove ANOVA in dobili vrednosti analize variance ločeno za velikost (tabela 2) in obliko (tabela 3). Izračun deleža napake po formulah 1 in 2 daje naslednje vrednosti $\%ME_{velikost} = 0,22$ in $\%ME_{oblika} = 13,37$. Kritična vrednost za napako velikosti je do 1% za obliko pa do 30%. Izračunani vrednosti sta v obeh primerih znatno manjši od kritičnih vrednosti $\%ME_{velikost}: 0,22 < 1$; $\%ME_{oblika}: 13,37 < 30$. Iz navedenega strnemo, da je napaka fotografiranja v okviru dopustnih vrednosti in zato bistveno ne vpliva na rezultate in zaključke.

Tabela 2. Vrednosti Prokrustove analize variance pri oceni napake fotografiranja za centroidno velikost objekta (SS - vsota kvadratov, MS - povprečje kvadratov, df - stopinje prostosti).

<i>EFFECT</i>	<i>SS</i>	<i>MS</i>	<i>DF</i>
Individual / among	125073,14	4466,90	28
Residual / within	184,49	3,35	55

Tabela 3. Vrednosti Prokrustove analize variance pri oceni napake fotografiranja za obliko objekta (SS - vsota kvadratov, MS - povprečje kvadratov, df - stopinje prostosti).

<i>EFFECT</i>	<i>SS</i>	<i>MS</i>	<i>DF</i>
Individual / among	0,0291614	0,0000473	616
Residual / within	0,0028472	0,0000024	1210

3. NAPAKA MERITVE

Napako meritve ovrednotimo s ponovnim izvajanjem meritev na istih objektih v treh ponovitvah. Pri tem meritve izvaja ista oseba v treh zaporednih dneh. Dobimo meritve istih objektov v treh časovnih serijah. Meritve statistično ovrednotimo z vključitvijo v Prokrustovo ANOVA [5], pri čemer nas zanima razmerje med varianco znotraj skupin in med skupinami. Izračun napake izpeljemo s pomočjo formul 1 in 2 [3]. Izmerjene vrednosti so odvisne spremenljivke, neodvisna spremenljivka pa je časovna serija meritev.

Na fotografijah 30 lobanj poljskih voluharic smo na ventralni strani lobanje določili 13 oslonilnih točk, na vseh lobanjah isto število točk na istih strukturah (slika 1). Izvedli smo postopek Prokrustove ANOVA in dobili vrednosti analize variance, ki jih predstavljamo ločeno za velikost (tabela 4) in obliko (tabela 5). Izračun deleža napake po formulah 1 in 2 daje naslednje vrednosti $\%ME_{\text{velikost}} = 0,11$ in $\%ME_{\text{oblika}} = 7,90$. Kritična vrednost za napako velikosti je do 1% in za obliko do 30% tako kot pri izračunu foto napake. Izračunani vrednosti sta v obeh primerih znatno manjši od kritičnih vrednosti, zato ugotovitev strnemo, da je napaka meritve v okviru dopustnih vrednosti in zato bistveno ne vpliva na rezultate in zaključke.

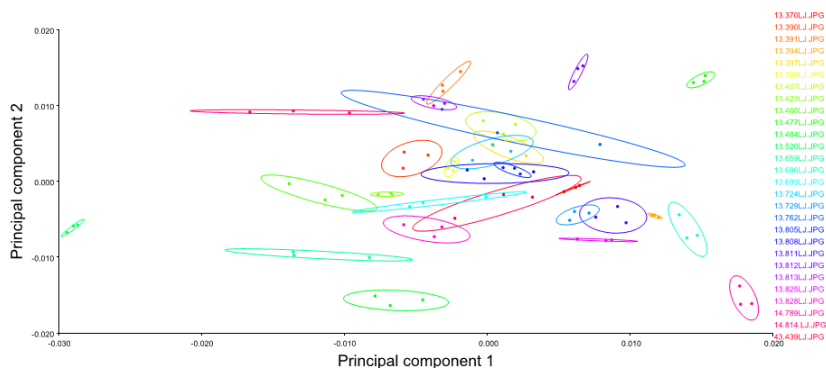
Obe do sedaj obravnavani vrsti napake, tako *napako fotografiranja* in *napako meritve*, lahko prikažemo tudi grafično s pomočjo projekcije večkrat fotografiranih ali merjenih istih osebkov na prvi dve glavni komponenti. V primeru *napake fotografiranja* gre za ponovitev meritev (določanja točk) na istem objektu fotografiranem v treh časovnih serijah. V primeru *napake meritve* gre za trikratno ponovitev meritve (določanja točk) na istem objektu na isti fotografiji. Na sliki 2 je ilustracija *napake meritve*. Pike predstavljajo porazdelitev osebkov izmerjenih v treh časovnih serijah. Isti osebki so povezani z elipso. Manjši razpon elipse pomeni manjše razlike med meritvami. Večji razpon elipse pomeni večje razhajanje med tremi meritvami istega objekta.

Tabela 4. Vrednosti Prokrustove analize variance pri oceni napake meritve za centroidno velikost objekta (SS - vsota kvadratov, MS - povprečje kvadratov, df - stopinje prostosti).

<i>EFFECT</i>	<i>SS</i>	<i>MS</i>	<i>DF</i>
Individual / among	126907,00	4700,27	27
Residual / within	92,50	1,65	56

Tabela 5. Vrednosti Prokrustove analize variance pri oceni napake meritve za obliko objekta (SS - vsota kvadratov, MS - povprečje kvadratov, df - stopinje prostosti).

<i>EFFECT</i>	<i>SS</i>	<i>MS</i>	<i>DF</i>
Individual / among	0,03124	0,000053	594
Residual / within	0,0018	0,000015	1232



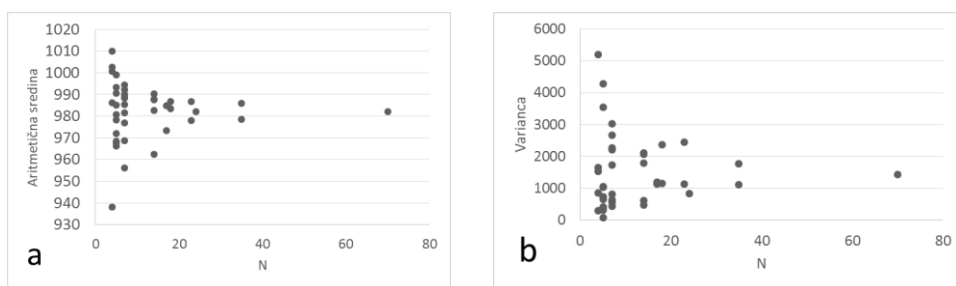
Slika 2: Projekcija 28 osebkov, na prvi dve glavni komponenti, izmerjenih v treh časovnih serijah. Trojica meritev na vsakem osebku je povezana z elipso in isto obarvana.

3. NAPAKA VZORCA

Eno od osnovnih vprašanj in težav v sodobni multivariatni morfometriji je vprašanje koliko enot (osebkov) naj vsebuje vzorec, da bodo izračuni in sklepi verodostojni. Odgovor ni niti enostaven niti enoznačen, saj je odvisen od statističnih metod, s katerimi ga analiziramo. V ta namen je razvitih več metod za oceno velikosti vzorca. V tem prispevku predstavljamo *analizo občutljivosti* (Sensitivity Analysis), ki smo jo povzeli po Carini-ju in sod. [4]. Naša analiza temelji na vzorcu 70 lobanj poljske voluharice. Analizirali smo statistična parametra aritmetično sredino in varianco za spremenljivki velikost (centroidna velikost) in obliko ventralne strani lobanje (Prokrustove

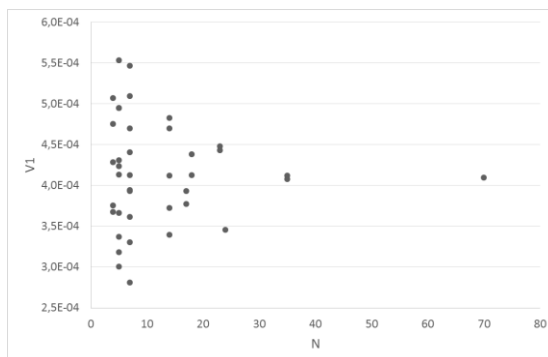
koordinate) s pomočjo geometrijske morfometrije. Celoten nabor objektov ($n=70$) smo naključno razdelili v manjše podvzorce, pri tem pa spremljali, kako zmanjševanje velikosti vzorcev vpliva na njihove statistične parametre. Podvzorce smo oblikovali z deljenjem osnovnega vzorca na 2, 3, 4, 5, 10 in 15 podvzorcev, pri tem smo oblikovali vzorce (s): $s_1 = 1/1$; $s_2 = 1/2$; $s_3 = 1/3$; $s_4 = 1/4$; $s_5 = 1/5$; $s_6 = 1/10$ in $s_7 = 1/15$ od 70 enot.

Na centroidno velikost lobanje zmanjševanja velikosti vzorca nima izrazitega vpliva. Grafična predstavitev variabilnosti aritmetične sredine izkazuje izrazito odstopanje šele med vzorci s po 15 enotami in manj (slika 3a). Medtem vizualizacija variabilnosti variance med vzorci izkazuje, da je že izrazita variabilnost dosežena pri vzorcih s po 23 enotami, z zmanjševanjem velikosti vzorcev pa variabilnost narašča (slika 3b).



Slika 3: Analiza občutljivosti za centroidno velikost: pike predstavljajo variabilnost v aritmetični sredini (a) in varianci (b) posameznih sukcesivno deljenih vzorcev.

Variabilnost v obliki lobanje posameznega vzorca smo izračunali kot vsoto varianc posameznih Prokrustovih koordinat (VI po [4]). Sukcesivno deljenje vzorcev z manjšanjem števila enot (osebkov) močno vpliva na variabilnost med vzorci (dejansko na varianco posamezne skupine). Pri prvi delitvi na dva vzorca s po 35 enotami skoraj ni razlike v varianci med vzorcema (slika 4).



Slika 4: Analiza občutljivosti za obliko objekta: razpršenost pik predstavljajo variabilnost v varianci posameznih sukcesivno deljenih vzorcev.

Pri delitvi na tri vzorce s po 23 oz. 24 osebkov pa se močno poveča variabilnost med vzorci. Nadaljnje delitve na večjo število vzorcev z manjšim številom enot pa samo povečuje razlika med njimi. Iz navedenih rezultatov lahko sklepamo, da moramo za verodostojni prikaz variabilnosti v obliki lobanje glodavcev v analizo vključiti vzorce z vsaj 30 osebki.

4. ZAKLJUČEK

Geometrijska morfometrija je v biologiji, medicini, forenziki in drugih vedah med drugim cenjena tudi zaradi možnosti zaznavanja majhnih razlik v obliki in velikosti med strukturami oz. objekti analize. Prav pri natančnih analizah in kadar so razlike med objekti majhne moramo upoštevati in oceniti vpliv napak na rezultate raziskav. V prispevku smo predstavili vrednotenje treh virov napak: napaka fotografiranja, napaka meritve in napaka vzorca.

LITERATURA

1. D. C. Adams, F. J. Rohlf, D. E. Slice (2013), A field comes of age: geometric morphometrics in the 21st century, *Hystrix, the Italian Journal of Mammalogy*, 24(1): 1–8.
2. T. Klenovšek (2014), Priročnik za uporabo geometrijske morfometrije v biologiji, Fakulteta za naravoslovje in matematiko, Univerza v Mariboru, Maribor.
3. S. M. Yezerinac, S. C. Loughheed, P. Handford (1992), Measurement error and morphometric studies: statistical power and observer experience, *Systematic Biology*, 41: 471–482.
4. A. Cardini, K. Seetah, G. Barker (2015), How many specimens do I need? Sampling error in geometric morphometrics: testing the sensitivity of means and variances in simple randomized selection experiments, *Zoomorphology*, 134(2): 149–163.
5. C.P. Klingenberg (2011), MorphoJ: an integrated software package for geometric morphometrics, *Mol. Ecol.-Res.*, 11: 353–357.

ŠTUDENTSKE APLIKACIJE

OBOGATENA RESNIČNOST GIBANJA UPORABNIKA V REALNEM ČASU NA MOBILNI NAPRAVI S POMOČJO PODATKOV GLOBINSKEGA SENZORJA

Nejc Škerjanc, Peter Peer

Fakulteta za računalništvo in informatiko, Univerza v Ljubljani

E-pošta: nejc.skerjanc01@gmail.com, peter.peer@fri.uni-lj.si

POVZETEK: *Ljudje težimo k vedno bolj natančni rekonstrukciji realnega sveta. Uporaba samo RGB podatkovnega toka ne daje zadovoljivih rezultatov pri detekciji in rekonstrukciji človeškega telesa. Z vpeljavo globinskega podatkovnega toka, ki nam ga posreduje globinski senzor, lahko rekonstruiramo okolico.*

Algoritem z uporabo sinhronega RGB in globinskega podatkovnega toka izvaja detekcijo s prepletom obeh tokov. S poznavanjem fizikalnih zakonitosti človeškega telesa lahko detektiramo posamezne dele človeškega telesa. Na podlagi detekcije sledi izris animiranega lika (z uporabo slikovnega atlasa), ki kar najbolj prekriva detektiranega človeka in človeško telo.

Kvalitativna in kvantitativna analiza je pokazala, da je z uporabo sinhronega RGB in globinskega podatkovnega toka detekcija človeškega telesa bolj natančna in učinkovita, kot samo uporaba posameznega podatkovnega toka.

1. UVOD

Ljudje težimo k vedno bolj natančni rekonstrukciji realnega sveta.

Globinski senzorji so naprave, ki omogočajo prenašanja 3D informacije iz realnega sveta v virtualni prostor. Bliskovit razvoj so doživeli v zadnjih letih zaradi cenovne dostopnosti širši skupini uporabnikov. Mobilne naprave so dosegle zadosten nivo zmogljivosti, da omogočajo procesiranje 3D informacije v realnem času. S pomočjo 3D informacij, ki nam jih posreduje globinski senzor, lahko rekonstruiramo okolico.

Do leta 2011 so se globinski senzorji uporabljali predvsem v industriji, avtomobilizmu, vojski in znanosti. Leta 2011 je Microsoft predstavil barvno globinski vizualni sistem Kinect. S tem je tehnologija zajema globinskih podatkov postala dostopna širši množici uporabnikov. V prihodnosti pričakujemo, da bodo postopoma vse mobilne naprave vsebovale globinske senzorje.

Mobilnost v navezavi z globinskim senzorjem prinaša nove možnosti uporabe mobilnih naprav. Ena izmed možnosti uporabe tovrstne tehnologije je tudi industrija zabavnih aplikacij, katerih prvotni namen je zabava. Uporabniška izkušnja bo dosegla nov nivo

(iskanje objektov, vzorcev, izvajanje kretenj ipd). Ob koncu leta 2016 je podjetje Lenovo v partnerstvu s podjetjem Google na trg poslalo prvi telefon z vgrajenim globinskim senzorjem (model Phab 2 Pro). Sistem za zaznavo globine se imenuje Tango [1].

V nalogi smo z uporabo Structure senzorja za iOS platformo [2] želeli izboljšati uporabniško izkušnjo detekcije človeškega telesa, kot jo ponuja detekcija samo z RGB podatkovnim tokom. Želeli smo doseči, da poze človeškega telesa ne bodo definirane vnaprej. Na podlagi detekcije smo prikazali animirani lik (z uporabo slikovnega atlasa), ki je kar najbolj prekril detektirano človeško telo. Prekrivanje deluje ne glede na pozicijo (odklon od trupa) rok in nog. Izboljšati smo želeli tudi delovanje v slabših svetlobnih pogojih (v primerjavi z detekcijo samo z uporabo RGB podatkovnega toka).

Primerjavo smo naredili z aplikacijo Cartoob [6]. Ta je sicer lastniška in detajli delovanja niso objavljeni, kljub vsemu pa vemo, da deluje le na osnovi RGB podatkovnega toka ter da je algoritem sestavljen iz naslednjih korakov:

1. Iskanje obraza in področja, kjer se nahaja človek.
2. Na področju človeka izračunamo HOG značilke.
3. Iz učne množice poiščemo najbolj verjetno pozo telesa glede na HOG značilke. Učna množica vsebuje naslednje poze človeškega telesa (27 poz): roka ob telesu, pravokotno na telo ali dvignjena, nogi na tleh, leva noga privzdignjena, desna noga privzdignjena.
4. Skaliranje virtualnega karakterja glede na zaznano velikost obraza v točki 1.
5. Upodabljanje karakterja.
6. Odpravljanje trepetanja lika z računanjem razlike slikovnih elementov med posameznimi slikami.

Detekcijo človeškega telesa s sinhrono uporabo RGB in globinskega toka smo zasnovali v 6 korakih:

1. Detekcija obraza.
2. Določitev območja zgornjih okončin.
3. Binarizacija območja zgornjih okončin in izvajanje podatkovne redukcije.
4. Postopek tanjšanja.
5. Detekcija zgornjih okončin.
6. Detekcija spodnjih okončin.

Razlaga in opis posameznih korakov našega algoritma se nahaja v naslednjem poglavju.

2. ALGORITEM ISKANJA ČLOVEŠKEGA TELESA

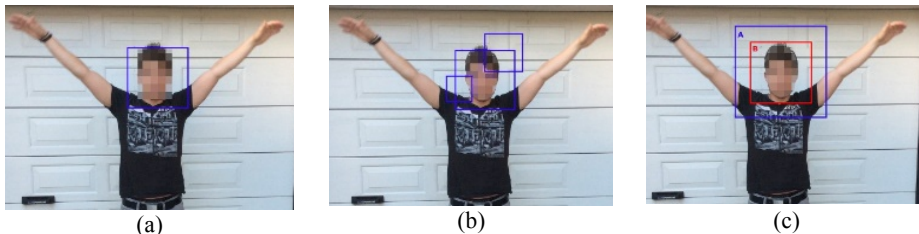
2.1 Detekcija obraza

Na RGB podatkovnem toku in z uporabo HAAR kaskadnega klasifikatorja poskušamo pridobiti lokacijo in velikost obraza [3]. Uspešno iskanje obraza je obvezen predpogoj za nadaljnje izvajanje algoritma.

Učinkovito iskanje obraza na celotni RGB sliki se je izkazalo za potratno operacijo, ki ni primerna za izvajanje v realnem času. Pri manjši natančnosti iskanja in slabši ambientalni svetlobi algoritem ni našel koordinat obraza.

Naredili smo optimizacijo iskanja obraza v dveh korakih:

- Iskanje z večjo natančnostjo smo naredili na enakih koordinatah in v 2 krat večjem območju od območja obraza v prejšnji sliki. Če v tem območju ne najdemo obraza, ga pričnemo iskati na območju celotne slike z zmanjšano natančnostjo (slika 1a).
- Iskanje obraza z veliko natančnostjo pogosto povzroči veliko lažno pozitivnih primerov, ki so običajno zelo majhne velikosti (slika 1b). Vsako potencialno območje, ki je znotraj drugega potencialnega območja, se izbriše. Zatem algoritem s seznama potencialnih kandidatov izbere tistega, ki ima največjo površino. Ostale kandidate izbriše (slika 1c).



Slika 1: Detekcija obraza: (a) območje iskanja z večjo natančnostjo, (b) najdena potencialna območja obraza in (c) izbira najboljšega kandidata. Zaradi boljše vizualizacije so prikazani odrezani deli originalnih slik.

2.2 Določitev območja zgornjih okončin

Določitev območja zgornjih okončin se izvaja ob prehodu z RGB podatkovnega toka na globinski tok podatkov. Kot izhodišče vzamemo najden obraz iz prejšnjega koraka. Območje povečamo po vnaprej znanih skalarjih v vseh smereh 2-dimenzionalne slike:

- v smeri navzgor za skalar 2,5,
- v smeri na levo in desno za skalar 3,5,
- v smeri navzdol za skalar 3.

Vrednosti skalarjev smo določili na podlagi renesančne predstavitve razmerij telesa.

RGB in globinski podatkovni tok imata različni ločljivosti vhodnih podatkov. Razširjeno območje zapišemo glede na celotno RGB sliko in ga prenesemo na globinski podatkovni tok.

2.3 Binarizacija območja zgornjih okončin in izvajanje podatkovne redukcije

Binarizacijo območja zgornjih okončin izvajamo na globinskem podatkovnem toku. Binarizacijo smo implementirali z algoritmom poplavljanja (angl. flood fill) [4]. Algoritem poplavljanja loči točke med seboj glede na kriterij pripadnosti. Če točka zadostuje kriteriju pripadnosti, jo označimo s pozitivno vrednostjo, v nasprotnem primeru z negativno vrednostjo.

Kriterij pripadnosti smo razdelili na dva dela. Prvi del kriterija pripadnosti se izvaja z računanjem razlike globin med točkami in njenimi sosedi. Drugi del je maksimalna dovoljena razlika med referenčno in trenutno točko.

Referenčna točka je osnovni podatek o zaznani razdalji človeškega obraza.

2.4 Postopek tanjšanja

Nad binarizirano sliko zgornjih okončin izvajamo morfološko operacijo tanjšanja. S postopkom tanjšanja si v naslednjih korakih pridobimo bolj učinkovito prepoznavanje pozicije trupa in zgornjih okončin. Tanjšanje izvajamo z Zhang-Suen algoritmom za tanjšanje [5].

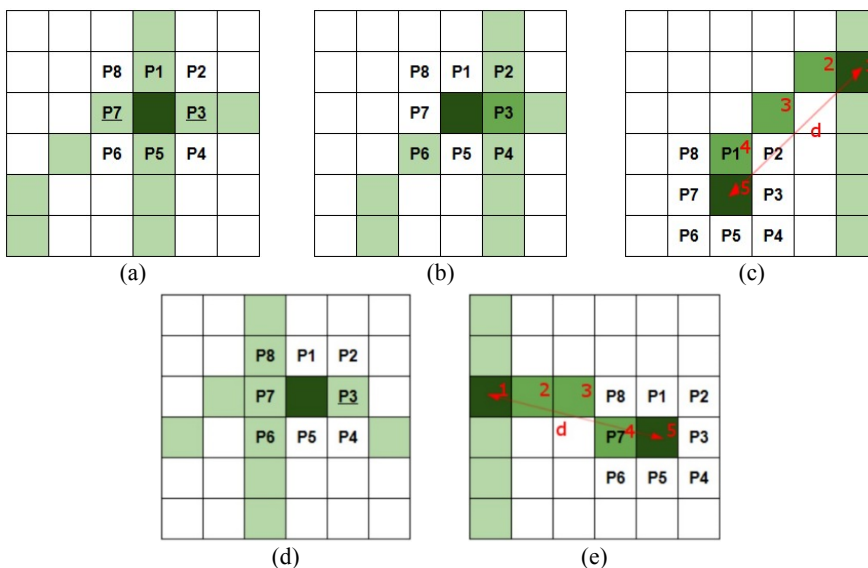
2.5 Detekcija zgornjih okončin

V sliki stanjšanih zgornjih okončin in trupa poiščemo najvišjo točko skeleta na sliki. Točko iščemo v območju glave iz prvega koraka. S tem pripomoremo k časovni optimizaciji algoritma.

Za nadaljevanje izvajanja označimo sosede vsake nerobne točke slike z oznakami P1 do P8 (oznake na sliki 2a okoli temnega kvadrata). Od najvišje do najnižje točke iterativno sledimo črti. V vsaki iteraciji naredimo 4 korake, ki so predstavljeni tudi v grafični obliki na sliki 3:

1. Preštejemo sosede s pozitivno vrednostjo na točkah od P1 do P8. Če je število sosedov 3 ali več, smo našli stičišče in nadaljujemo z izvajanjem preostalih korakov. V nasprotnem primeru preostale korake preskočimo in se prestavimo na novo točko (slika 3a).

2. Za detekcijo leve roke preverimo nahajanje pozitivne vrednosti na pozicijah od P6 do P8. To točko imenujemo potencialna točka leve roke. Za detekcijo desne roke preverimo nahajanje pozitivne vrednosti na pozicijah od P2 do P4. To točko imenujemo potencialna točka desne roke.
3. Želimo se prepričati, da sta potencialni točki leve ali desne roke stičišči trupa in rok. Preprečiti želimo lažno pozitivne primere, pri katerih je potencialna točka samo preskok črte po algoritmu tanjšanja. Prestavimo se na potencialno točko in pri potencialni točki leve roke preverimo vrednosti na pozicijah P1 in od P5 do P8. Pri potencialni točki desne roke preverimo vrednosti na pozicijah od P1 do P5. Vsaj ena točka mora biti pozitivna in potencialna točka roke postane koren (ramena) roke (sliki 3b in 3d).
4. Iterativno sledimo črti roke. Ko pridemo do konca, izračunamo razdaljo med stičiščem in končno točko. Če je razdalja med točkama od tri do pet kratnik velikosti glave, točko shranimo in zapišemo kot levo/desno zapestje (sliki 3c in 3e).



Slika 2: Postopek iskanja rok

Binarizirano sliko zgornjih okončin in rezultat tanjšanja z najdenimi okončinami vidimo na sliki 3.



Slika 3: (a) Binarizirana slika zgornjih okončin, (b) stanjšana slika z najdenimi rameni (modra barva) in zapestji (rdeča barva).

2.6 Detekcija spodnjih okončin

Detekcija kolkov in nog se izvaja v dveh korakih:

1. Iskanje rok \hat{q} zaključeno na višini kolkov. S tem podatkom pridobimo informacijo o višini kolkov. Za iskanje pozicije levega kolka na binarni sliki rekurzivno izvajamo premike v levo smer, dokler je izpolnjen pogoj pozitivne vrednosti trenutne točke. Pozicijo desnega kolka dobimo na enak način, s premikom v desno smer. Obe točki shranimo kot levi/desni kolk. Točki se nahajata na prehodu med pozitivno in negativno vrednostjo v binarni sliki.
2. Izhodišče za iskanje stopal sta najdeni točki kolkov. Iz izhodiščne točke se iterativno premikamo po y koordinati navzdol in iščemo preskok med pozitivno in negativno vrednostjo. Iščemo v območju P4 do P6. Postopek ponavljamo do iteracije, ko ne najdemo več preskoka. Renesančna predstavitev razmerij človeškega telesa definira dolžino nog kot štiri kratnik velikosti glave. Preverimo ali je dolžina med kolki in stopali med 3,5 in 4,5-kratnikom. Če se nahaja v teh okvirjih, shranimo točko kot levo/desno stopalo.

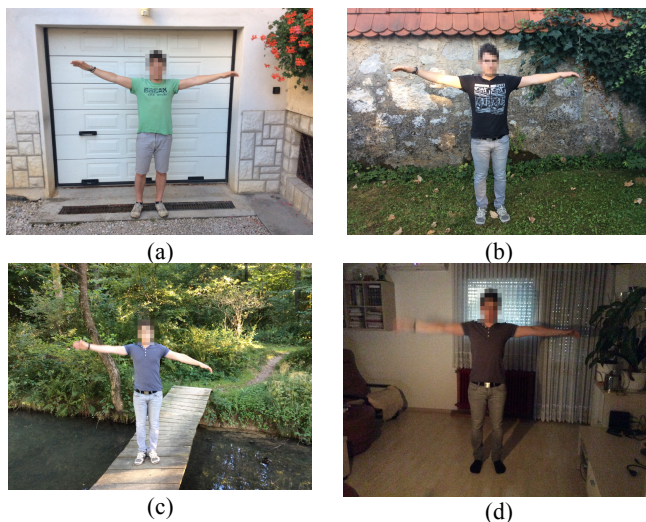
3. REZULTATI

Analizo smo naredili na podlagi štirih scenarijev. S scenariji smo testirali vse aspekte, ki so pomembni pri detekciji človeškega telesa. Scenariji se delijo na podscenarije. Vsak scenarij (ali podscenarij) smo testirali na 9 pozah človeškega telesa.

Testirani scenariji so bili:

- Urbano okolje: Lokacija scenarija je v urbanem okolju pred garažnimi vrati. Za preverjanje kvalitete procesiranja smo z različnimi oblačili izvedli dva podscenarija. Ozadje je bilo v dosegu globinskega senzorja (slika 4a).

- Mestni park: Lokacija scenarija je bila v mestnem parku ob obzidju. Ozadje je bilo zelo podobno barvi kože in v dosegu globinskega senzorja (slika 4b).
- Naravno okolje: Lokacija je bila v naravi ob jezeru. Na lesenem mostu smo izvedli dva podscenarija z različnimi oblačili. Scenarij je preverjal delovanje algoritma, kjer je ozadje dlje od dosega globinskega senzorja (slika 4c).
- Notranji prostor: Lokacija scenarija je bila notranji prostor. Pri tem scenariju z umetno svetlobo smo preverjali učinkovitost detekcije ob različnih svetlobnih pogojih. Ozadje je bilo v dosegu globinskega senzorja (slika 4d).



Slika 4: (a) Urbano okolje, (b) mestni park, (c) naravno okolje, (d) notranji prostor.

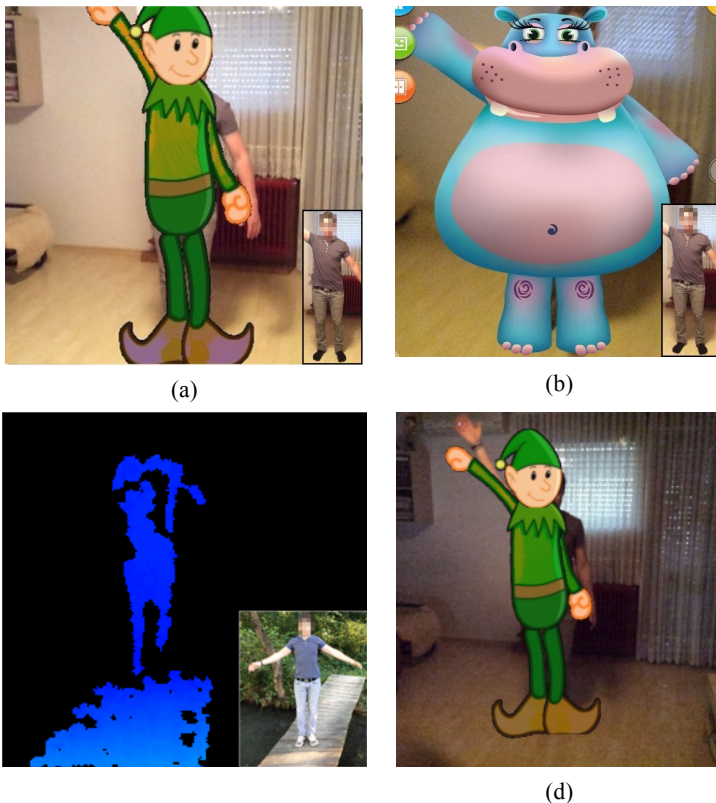
Med seboj smo primerjali metodo z RGB podatkovnim tokom (Cartoob [6]), našo metodo z globinskim tokom in našo metodo, ki združuje oba podatkovna toka (tabela 1).

Tabela 1: Primerjava uspešnosti detekcije človeškega telesa v različnih scenarijih in z različnimi podatkovnimi toki.

scenarij	podscenarij/algoritem [pravilno detektirane klasifikacije (čas izvajanja)]	Cartoob (RGB p.t.)	globinski p.t.	RGB & globinski p.t.
urbano okolje	temnejša oblačila	66% (41ms)	44% (880ms)	88% (110ms)
	svetlejša oblačila	66% (41ms)	44% (850ms)	88% (100ms)
mestni park	temnejša oblačila	66% (41ms)	22% (950ms)	88% (130ms)
Naravno okolje	temnejša oblačila	77% (41ms)	0% (3000ms)*	0% (35ms)*
	svetlejša oblačila	77% (41ms)	0% (3800ms)*	0% (40ms)*
notranji prostor	difuzna umetna svetloba	66% (41ms)	44% (3800ms)	88% (35ms)
	delno zatemnjen prostor	22% (41ms)	33% (3800ms)	88% (40ms)
	zatemnjen prostor	0% (41ms)	33% (3800ms)	33% (40ms)

*Anomalije v podatkih zaradi strojne opreme.

Slika 5a prikazuje prikrivanje z uporabo obeh podatkovnih tokov, slika 5b prikazuje prekrivanje z uporabo samo RGB podatkovnega toka v Cartoobu. Pojav anomalij v podatkih z globinskega podatkovnega toka pri močnejši zunanji svetlobi vidimo na sliki 5c (problem Structure senzorja). Scenarij zaprtega prostora s srednje dobro osvetlitvijo in z uporabo algoritma, ki združuje oba podatkovna toka vidimo na sliki 5d. Pri tej osvetlitvi Cartoob ni deloval.



Slika 5: (a) Primer rezultata algoritma z obema podatkovnima tokoma (desno spodaj je podana vhodna slika), (b) Cartooba, (c) anomalije na globinskem podatkovnem toku Structure senzorja (črna barva predstavlja nedefinirane vrednosti), (d) scenarij zaprtega prostora s srednje dobro osvetlitvijo in rezultat algoritma z obema podatkovnima tokoma.

Analiza vseh scenarijev je pokazala, da naš algoritem bolje detektira človeško telo v različnih okoljih in z več pozami kot Cartoob. Naš algoritem je manj občutljiv na svetlobo kot algoritem v Cartoobu. Združevanje obeh podatkovnih tokov je dobra izbira.

Cartoob izvaja prekrivanje na osnovi predefinirane množice 27. poz. Naš algoritem izvaja prekrivanje na osnovi posameznih delov človeškega telesa. Število poz ni omejeno. Detektiran odklon udov od telesa se v enakem razmerju prenese na izris tekstur

delov človeškega telesa. Zaradi tega je natančnost prileganja veliko boljša kot pri uporabi Cartoob. S tem je tudi uporabniška izkušnja bolj pristna.

Slabost našega algoritma so slabši povprečni odzivni časi. Cartoob izvaja eno iteracijo detekcije človeškega telesa in animacijo konstantno z 41 ms. Naš algoritem potrebuje povprečno 77,5 ms.

Naš algoritem ima z uporabo Structure senzorja slabe rezultate v okoljih, kjer ozadje sega izven maksimalnega dosega globinskega senzorja. Če je namreč razdalja večja od maksimalnega dosega naprave, se na sliki pojavljajo nedefinirana območja, ki segajo tudi v območja, v katerih bi naprava morala zaznati globino. Do teh anomalij prihaja, če je nedefinirano območje večje od 1/3 velikosti celotne slike. To je ključna slabost Structure senzorja.

Do anomalij v podatkih prihaja tudi, kjer nimamo nedefiniranih območij, vendar imajo posamezni slikovni elementi nedefinirane vrednosti. Te anomalije niso kritične in jih lahko minimiziramo z izvajanjem interpolacije v okolici.

V našem algoritmu smo zaznali nekoliko večje trepetanje izrisane lika kot v Cartoob algoritmu. Trepetanje smo v večji meri odpravili z računanjem razlike med novo in prejšnjo sliko. Teksturo smo izrisali po novih podatkih, ko smo zaznali spremembo večjo od 5. enot (slikovnih elementov ali stopinj v rotaciji).

Ključne primerjave med Cartoobom in našim algoritmom so povzete v tabeli 2.

Tabela 2: Primerjava Cartoob in našega algoritma

	Cartoob	Naš algoritem
Podatkovni tok	RGB	RGB in globinski
Skupni rezultat – točnost (brez scenarija 3 [*])	36%	56%
Prekrivanje	Predefinirana množica (27 poz)	Na osnovi posameznih delov telesa
Odzivni čas (brez scenarija 3 [*])	Konstantno: 41 ms	Povprečje: 77,5 ms
Trepetanje izrisane lika	Malo trepetanja	Nekoliko večje trepetanje posameznega dela telesa
Slabši svetlobni pogoji	Difuzna umetna svetloba že povzroči nedelovanje	Deluje v temnejšem prostoru
Delovanje na prostem	Dobra detekcija	Slaba detekcija (problem Structure senzorja [*])

^{*}Anomalije v podatkih zaradi strojne opreme.

4. ZAKLJUČEK

Za klasifikacijo nismo uporabljali baze vnaprej posnetih slik človeškega telesa. S tem smo omogočili boljše prekrivanje brez omejitev, ki jih prinaša baza.

Naredili smo kvalitativno in kvantitativno analizo implementiranega algoritma ter ga primerjali z algoritmom, ki je vključen v aplikacijo Cartoob in algoritmom, ki uporablja samo globinski podatkovni tok.

Naš algoritem se je izkazal bolje od obeh preostalih algoritmov. Prekrivanje izvaja na osnovi posameznih delov telesa (ne samo 27 poz, kot to dela Cartoob). Uporabniška izkušnja in kvaliteta prekrivanja je s tem veliko boljša. Zaradi izvajanja na dveh podatkovnih tokovih so bili časi procesiranja nekoliko slabši, kot pri Cartoobu. Pri slabših svetlobnih pogojih se je naš algoritem izkazal bolje, ker je pri detekciji uporabljal tudi globinski podatkovni tok. Na zunanji sončni svetlobi Structure senzor ne deluje. Posledično naš algoritem ne najde človeškega telesa. Možne izboljšave so v smeri optimizacije, trepetanja, časa procesiranja ipd.

Dokazali smo, da globinski senzorji v mobilnih napravah prinašajo nov nivo uporabniške izkušnje [7].

LITERATURA

- [1] (2017) Lenovo Pnhab 2 Pro. Dostopno na: <http://shop.lenovo.com/us/en/tango/>.
- [2] (2017) Occipital, Structure sensor. Dostopno na: <http://structure.io/>.
- [3] P. Viola, M. J. Jones, Robust real-time face detection, *International journal of computer vision*, 57 (2), str. 137-154, 2004.
- [4] P. S. Heckbert, A seed fill algorithm, *Glassner*, 1 (3), str. 275-277, 1990.
- [5] T. Zhang, C. Y. Suen, A fast parallel algorithm for thinning digital patterns, *Communications of the ACM*, 27 (3), str. 236-239, 1984.
- [6] (2017) Cartoob spletna stran. Dostopno na: <http://cartoobapp.com/>.
- [7] N. Škerjanc, Obogatena resničnost gibanja uporabnika v realnem času na mobilni napravi s pomočjo podatkov globinskega senzorja, magistrska naloga, Fakulteta za računalništvo in informatiko Univerze v Ljubljani, Slovenija, 2016.

ZAKRIVANJE OSEB NA POSNETKIH GRADBIŠČA

Gašper Sedej, Martin Šavc

Laboratorij za sistemsko programsko opremo
Fakulteta za elektrotehniko, računalništvo in informatiko
E-pošta: gasper.sedej@um.si, martin.savc@um.si
URL: <http://lspo.feri.um.si>

POVZETEK: *Kamere s katerimi nadziramo gradbišča zajemajo tudi zaposlene na gradbišču in s tem posegajo v pravico do zasebnosti zaposlenega. Pri obdelavi in posredovanju takšnih slik v druge informacijske sisteme moramo zato zaposlene na slikah zakriti. Z uporabo globokega učenja in regijskih konvolucijskih nevronskih mrež smo iskali osebe na slikah gradbišč. Osnovna metoda zaznavanja je zaznala le 25 % vseh oseb, medtem ko je predlagana metoda detektirala 88 % vseh oseb na gradbišču. Z uporabo metode »inpaint« smo zaznane osebe na slikah še na koncu zakrili.*

1. UVOD

Gradbeniki v zadnjih desetletjih uporabljajo sistem BIM (*angl. Building Information Modelling*) za modeliranje gradnje. Primerjava modela z dejanskim stanjem na gradbišču je mogoča s pomočjo slik visoke ločljivosti. Takšne slike lahko poleg gradbenih objektov vsebujejo tudi veliko oseb, tj. zaposlenih na gradbišču. Snemanje zaposlenih na delovnem mestu spada pod zajem osebnih podatkov. Da bi lahko slike posredovali zunanjemu podjetju za obdelavo ali objavo v javnosti, je potrebno osebne podatke zaposlenih na slikah skriti.

Snemanje in video nadzor delovnega mesta ureja in omejuje Zakon o varstvu osebnih podatkov (pravica zasebnosti in delne samostojnosti, 77. člen ZVOP-1 [1]). Ta zakon vpliva tudi na druge aplikacije zajemanja slik in video posnetkov v javnosti. Problem običajno rešujemo z zakrivanjem oseb v zajetih fotografijah [2].

Naš cilj je uporaba računalniškega vida za avtomatizirano zaznavo oseb v slikah gradbišča. Zaznane regije pa želimo nato zakriti z minimalnim estetskim vplivom na sliko.

Za zaznavo osebe oziroma obraza obstaja več postopkov. Med bolj znanimi so Viola-Jones oziroma kaskadni klasifikator, metoda na osnovi lastnih obrazov (*angl. eigenfaces*), ki uporablja analizo poglavitnih komponent in detektor kože z verjetnostjo barvnega prostora.

Za prepoznavo in detekcijo predmetov v sliki so globoke nevronske mreže trenutno eden izmed uspešnejših pristopov. Dobri modeli prepoznajo že 1000 razredov

(kategorij) z natančnostjo 66 % mAP (*angl. Median Average Precision*) [3] oziroma 20 razredov z 82 % mAP [4]. Pogosto so v te razrede vključene tudi osebe.

Ko osebo na sliki lokaliziramo, jo je potrebno zakriti. Na voljo je več načinov, ki imajo različen vizualni poseg v sliko. V tem prispevku bomo predstavili zapolnitev zaznane regije s črno barvo, povprečno barvo, zameglitev regije, pikselizacijo in metodo »inpaint«.

V tej raziskavi predstavljamo uporabo obstoječega detektorja, ki temelji na nevronskih mrežah. Le-tega smo prilagodili za namen zakrivanja oseb na gradbišču. Rešitev smo zasnovali tako, da smo zmanjšali število nedetektiranih oseb, kar je za končno aplikacijo izredno pomembno. Predlagano metodo smo validirali na posnetkih dveh različnih gradbišč.

2. OPIS METODE

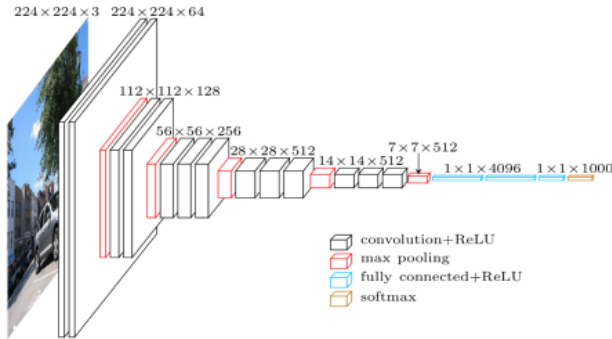
Za zaznavanje oseb smo modificirali obstoječi detektor, ki temelji na uporabi nevronskih mrež. Zaznane osebe nato zamaskiramo z eno izmed metod zakrivanja. Najprej si bomo pogledali delovanje uporabljene nevronske mreže, v nadaljevanju pa bomo še primerjali različne metode zakrivanja.

2.1 Zaznava oseb s pomočjo nevronskih mrež

Za ogrodje konvolucijske nevronske mreže (CNN) smo uporabili ogrodje Caffè [5]. To ogrodje omogoča fleksibilno uporabo in modifikacijo posameznih komponent. Na voljo so različni naučeni modeli, možno pa je naučiti tudi lasten model. Knjižnica podpira različne jezike (tj. programski jezik C++, Python in Matlab). Ogrodje Caffè podpira tudi lasten standard za modele imenovan »caffemodel«.

Uporabili smo že naučen model VGG16 [6], ki spada med najboljše modele za razpoznavo objektov na slikah. Zgradbo modela prikazuje slika 1. Model sestoji iz 13 konvolucijskih slojev ter treh polno povezanih slojev. Konvolucijskim slojem sledi »Max pooling« sloj, končni rezultat pa poda funkcija »Softmax«. S pomočjo tega modela lahko prepoznamo 21 različnih razredov, in sicer 20 objektov ter ozadje [7]. Med dvajsetimi razredi najdemo razrede, kot so avtomobil, letalo, ptica, mačka, pes in tudi razred oseba, ki je za nas ključnega pomena.

Za zaznavo objekta na sliki uporabimo regijsko konvolucijsko nevronske mrežo (R-CNN) [8]. Ta mreža je zmožna ob prepoznavi objekta podati še njegovo lokacijo in velikost na sliki. Sliko najprej razdeli z algoritmom »Selective search« na regije kandidatke. Te regije nato posredujemo na vhod CNN. Mreža kot rezultat vrne dve matriki, in sicer »Scores« in »Boxes«. Matrika »Scores« vsebuje ocene verjetnosti in je dimenzije $M \times N$, kjer M predstavlja število kandidatov in N število razredov (v našem primeru 21). Matrika »Boxes« pa je dimenzije $M \times 4N$. V tej matriki je za vsako oceno opisana pozicija in velikost pravokotnika na sliki.



Slika 1: Diagram mreže VGG16. Vir: [9]

2.2 Zakrivanje zaznanih oseb

Rezultat zaznavanja oseb so okvirji (*angl. bounding box*), kjer se nahajajo osebe na sliki. Te regije je potrebno v sliki prekriti, tako da odstranimo vse informacije o osebi. To operacijo običajno imenujemo zakrivanje, zanjo pa obstaja več uveljavljenih pristopov. V grobem jih lahko delimo na postopke, ki ne dodajajo umetne vsebine (npr. črni kvadrat, povprečje barve, zameglitev, »pixelization«) in na postopke, ki regijo zapolnijo z umetno vsebino.

Za zakrivanje lahko uporabimo nekaj preprostih postopkov. Najenostavnejši postopek je zapolnitev zaznane regije s črno barvo. To ima velik negativni vizualni vpliv na sliko. Preostali postopki poskušajo ta vpliv zmanjšati. S povprečjem barve regijo nadomestimo s povprečno vrednostjo po vseh barvnih kanalih. Pravokotnik regije je še vedno viden, vendar je v slikah gradbišča manj vpaddljiv kot črn pravokotnik. Zameglitev lahko naredimo z uporabo Gaussovega filtra z velikim jedrom. Skrito območje je manj očitno, vendar še vedno opazno zaradi odsotnosti ostrih robov. Hkrati območje ni popolnoma zakrito, saj je mogoče prepoznati obris osebe. Postopek »Pixelization« oziroma »Mosaic« je tehnika, ki območju umetno zmanjša ločljivost in nato vsebino interpolira s tehniko najbližji sosed. Tudi ta način močno posega v sliko. Vse te metode vnašajo popravke, ki so za opazovalca očitni. Demonstracija opisanih postopkov je prikazana na sliki 2.

Če želimo manj vsiljiv poseg v sliko, moramo uporabiti kompleksnejšo metodo. Tipično se opremo na podatke iz okolice regije. Takemu postopku se v angleščini reče »inpainting« (orodja podjetja Adobe pa jih imenujejo »Content Aware Fill«). Postopek je namenjen rekonstrukciji oziroma restavriranju slik (npr. razpoke ali odrgnine razvitih fotografij) in je tradicionalno ročni poseg v sliko. Računalniški algoritem posnema ta postopek tako, da iz okolice vstavi najprimernejše teksture v izbrano regijo. Primerjava različnih postopkov »Inpaint« je prikazana na sliki 3. Postopki so natančneje opisani v poglavju 3.2.



Slika 2: Tehnike zakrivanja (od leve proti desni): originalna slika, zapolnitev s črno barvo, zameglitev, poprečje barv in "Pixelization"



Slika 3: Primerjava različnih postopkov »Inpaint« (od leve proti desni): originalna slika, postopek »Inpaint«: Scikit-image, postopek »Inpaint« iz knjižnice OpenCv in postopek »Inpaint« iz orodja Gimp - G'MIC

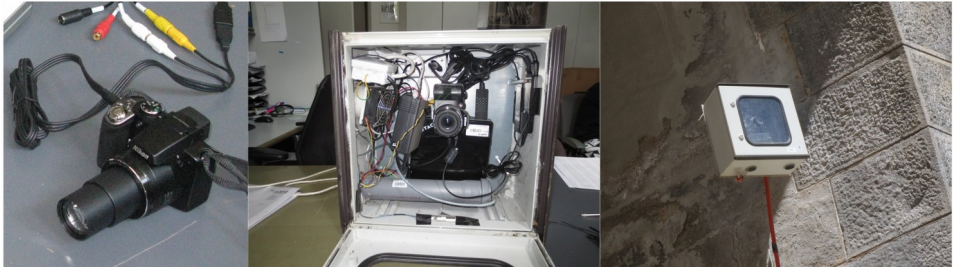
3. EKSPERIMENT

Naš predlagan postopek smo validirali z naslednjim eksperimentom. Eksperiment smo izvedli na slikah gradbišč. Na voljo smo imeli slike iz dveh gradbišč: a) gradnja fakultete v mestu in b) gradnja razglednega stolpa na vrhu hriba. Pri gradnji fakultete imamo slike zajete iz dveh različnih lokacij, medtem ko pri gradnji stolpa le slike iz enega zornega kota.

Za zajem slik smo uporabili strojno-programski sistem, ki smo ga izdelali v Laboratoriju za sistemsko programsko opremo. Sistem je opisan v diplomskem delu [10] in na spletni strani projekta 4D-ACT [11]. Na sliki 4 je prikazan uporabljen fotoaparati Nikon Coolpix P100, ki je vgrajen v ohišje skupaj z modulom za nadzor fotoaparata in modulom za komunikacijo z oddaljenim sistemom. Na sliki lahko vidimo tudi montirani sistem na eni izmed lokacij gradbišča fakultete.

Za eksperiment smo izmed vseh zajetih slik izbrali 40 slik, na katerih se nahajajo osebe. Število oseb na posamezni sliki je variiralo med ena in osem. Ena izmed testnih slik ni vsebovala nobene osebe. Poleg oseb scena vsebuje tudi druge objekte kot so gradbeni stroji, drogovi, avtomobili ipd. Osebe so včasih deloma prekrivane. Tudi poza oseb je na nekaterih slikah neobičajna, na primer, čepenje pri polaganju talne napeljav. Slike so posnete pri različnih vremenskih in svetlobnih pogojih. Vključujejo tako sončno kot deževno vreme, osebe se nahajajo tako na svetlobi kot v senci ter deloma v senci in deloma na sončni svetlobi.

Kot osnovno programsko opremo smo uporabili delujoči vzorčni program v projektu »Faster R-CNN (Python implementation)« [12]. Demo.py je python program, ki s pomočjo nevronske mreže »caffe« in naučenega modela VGG16 zaznava objekte na sliki.



Slika 4: Strojno-programski sistem za zajem slik na gradbišču (od leve proti desni): fotoaparata Nikon Coolpix, vgrajen v sistem, sistem in montiran sistem na gradbišču

3.1 Detekcija oseb na gradbišču

Obstoječa rešitev je prilagojena slikam z nižjo ločljivostjo (500x400 pikselov, kar je približno 0,2 MP). V primeru vhoda slike z visoko ločljivostjo (2600x2000 pikselov, ~5 MP) zazna samo velike objekte, medtem ko so manjši objekti izpuščeni, čeprav so v sliki še vedno jasno vidni.

Prva modifikacija, ki smo jo izvedli, je bila specializacija za prepoznavo oseb. Izmed vseh rezultatov, ki jih dobimo po detekciji, preverjamo samo verjetnost prisotnosti osebe. Naša želja je detekcija vseh oziroma čim več oseb na sliki. Zato smo prag, uporabljen pri prepoznavi, znižali iz privzete 0,8 na 0,3.

S pomočjo druge modifikacije obstoječe rešitve smo odpravili pomanjkljivost pri detekciji na slikah z visoko ločljivostjo. To smo naredili tako, da smo vhodno sliko razrezali na manjša prekrivajoča se podokna izbranih dimenzij. Izvedli smo več eksperimentov z različnimi velikostmi podokna: 1100x1100 pikselov, 800x800 pikselov, 400x400 pikselov in 200x200 pikselov. Eksperimenti so pokazali, da je najbolj ustrezna ločljivost 400x400 pikselov, s prekrivanjem 30 % okna. Pri oknih z višjo ločljivostjo detektor pogosto zgreši manjše objekte (osebe), nižja ločljivost pa pomeni več oken in posledično več računskega časa. Hkrati bi slabo zaznali tudi srednje velike objekte, saj bi bilo okno samo del te osebe. Velike objekte zaznamo že, ko damo na vhod kar sliko v originalni ločljivosti.

3.2 Zakrivanje oseb na gradbišču

V tej raziskavi smo preverili različne implementacije postopka »Inpaint«. Obstaja več aplikacij, ki podpirajo to funkcionalnost. Adobe Photoshop, na primer, ponuja »Content-Aware Fill« od različice CS5 [13] naprej. Namenska aplikacija »The Inpaint« ponuja enostavno uporabo te funkcionalnosti tudi za neuke uporabnike [14].

Odprto-kodno orodje Gimp omogoča postopek »Inpaint« preko dveh različnih vstavkov (*angl. plugin*). Prvi je Resynthesizer [15], ki funkcijo imenuje »Heal Section«, drugi pa »G'MIC« [16], in ponuja več načinov popravkov »Inpaint«. Vstavek »G'MIC« je na voljo tudi kot samostojna aplikacija.

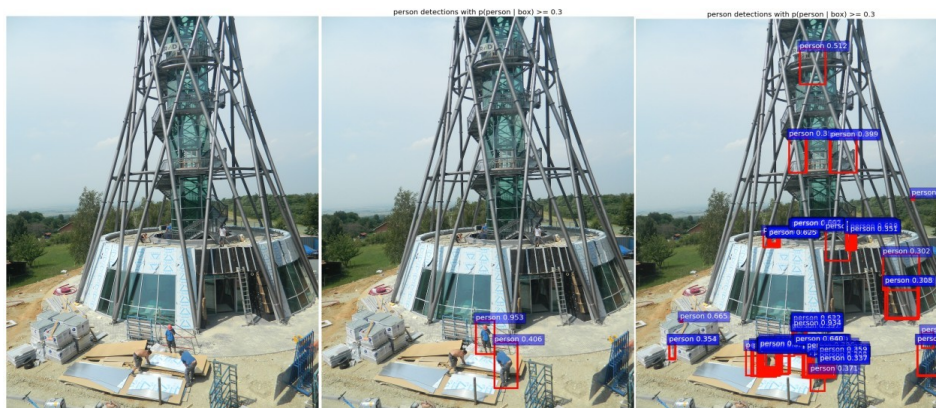
Integracija je enostavnejša z metodami, ki so implementirane v knjižnicah. Preverili smo dve prosto dostopni knjižnici: knjižnico Scikit-image in OpenCv. Knjižnica Scikit-

image je knjižnica napisana v jeziku Python in je namenjena obdelavi slik. Funkcionalnost »Inpaint« je implementirana po algoritmu »Biharmonic equation« [17]. Rezultat je podoben nizkemu filtriranju tega dela slike. Knjižnica OpenCV je zelo pogosto uporabljena knjižnica za obdelavo slik. Ponuja dva načina »Inpaint«: metodo Navier-Stokes in metodo Telea [18]. Obe metodi dajeta podobne rezultate. Po naši subjektivni oceni so rezultati boljši od rezultatov, ki jih dobimo s pomočjo knjižnice Scikit-image. Primerjava različnih metod »Inpaint« je prikazana na sliki 3. V tej raziskavi smo se odločili za uporabo knjižnice OpenCV in metode Telea.

4. REZULTATI

Rezultate detekcije smo ocenjevali ročno. Ob prvem pregledu smo prešteli osebe na izvornih slikah. Ob drugem pregledu smo prešteli, koliko oseb je našla nevronska mreža brez uporabe podoken. Naslednji pregledi so vključevali rezultate detekcij metod s podokni. Nevronska mreža je med obdelovanjem našla še nekatere nove osebe. Teh v prvem koraku sami nismo našli in to predvsem zaradi kompleksne scene, slabih svetlobnih pogojev in močnih kontrastov (senca in obsijan prostor). Novo najdene osebe smo dodali v skupno število oseb na sliki. Skupno število oseb na 40 slikah je tako bilo 164.

Zaradi več-nivojskega delovanja je prišlo do večkratne detekcije istega objekta pri različnih velikostih. Prekrivajoče in sosednje regije smo zato smiselno združili. Primer tega je viden na sliki 5. V kolikor je regija prekrivala večinski del osebe, smo označili kot pravilno zaznavo TP (*angl. True Positive*). Če je mreža zaznala osebo, kjer oseba ni bila prisotna, smo to označili za napačno zaznavo FP (*angl. False Positive*). Manjkajoče osebe FN so razlika med vsemi in zaznanimi osebami (*angl. False Negative*). Iz teh podatkov smo lahko izračunali mero natančnosti (*angl. Precision*) in priklica (*angl. Recall*).



Slika 5: Detekcija oseb na sliki gradbišča (od leve proti desni): vhodna slika s šestimi osebami, rezultat metode brez podoken in rezultat izboljšane metode s podokni.

Tabela 1: Priklic in natančnost zaznavanja oseb z različnimi velikostmi podokna

	brez podokna	podokno 1100x1100 pikslov	podokno 800x800 pikslov	podokno 400x400 pikslov
pravilno zaznani TP	42	119	127	143
manjkajoče osebe FN	122	45	37	21
napačno zaznani FP	4	25	67	266
priklic	25,6 %	72,6 %	77,3 %	88,2 %
natančnost	91,3 %	82,6 %	62,6 %	35 %

Podrobni rezultati so prikazani v tabeli 1. V vrsticah so navedeni rezultati za pravilno zaznane osebe (TP), osebe, ki jih nismo detektirali (FN), napačno zaznane osebe (FP), ter metriki priklic in natančnost. V stolpcih so navedeni rezultati različnih pristopov k detekciji oseb. Prvi pristop je osnovna oziroma originalna metoda, ki ji sledijo metode z različnimi velikostmi podokna.

5. DISKUSIJA IN SKLEP

Z uporabo nevronske mreže in osnovne metode smo zaznali samo četrtnino vseh oseb na slikah gradbišč. Metodo smo izboljšali tako, da smo uporabili podokna in bistveno izboljšali metriko priklic. Uspeli smo zaznati in zakriti 88,2 % oseb v zelo kompleksnih scenah in slabih svetlobnih ter vremenskih pogojih.

Žasmo hkrati zaznali veliko regij, kjer osebe niso bile prisotne. To je vplivalo na ocenjeno natančnost sistema. Le ta je padla iz 91,3 % pri metodi brez podokna na 35 % pri metodi z najmanjšim podoknom (400 x 400 pikslov). Zakrivanje regij ima precejšen vpliv na videz slike in vsebnost uporabnih informacij. Izboljšanje natančnosti je pomemben korak v nadaljnjem delu. Želeli bi še dodatno izboljšati zaznavo oseb na gradbiščih. To bi lahko storili z dodatnim uglaševanjem nevronske mreže (*angl. fine-tuning*) s slikami oseb na gradbiščih.

Za zakrivanje trenutno uporabljamo samo informacije iz slike, v kateri je bila oseba zaznana. Za boljše rezultate potrebujemo ustrezne podatke. Te bi lahko pridobili tudi iz časovno sosednjih slik, v katerih oseba ni bila prisotna. Estetski rezultat zakrivanja bi lahko izboljšali z uporabo kvalitetnejših metod. Primer take metode je orodje G³MIC.

Uporabljeni model mreže VGG16 je izdan pod licenco, ki ne dovoljuje komercialne uporabe. Želeli bi preveriti druge modele, ki dovoljujejo komercialno uporabo, kot so GoogLeNet [19] in AlexNet [20]. Na ta način bi lahko takšen postopek vgradili znotraj našega sistema za spremljanje gradnje in naročniku ponudili rešitev, ki ni sporna glede zakonodaje in varstva osebnih podatkov.

LITERATURA

1. <https://zakonodaja.com/zakon/zvop-1/77-clen-delovni-prostori>, zadnji dostop: februar 2017
2. <https://www.ip-rs.si/vop/google-street-view-1943/>, zadnji dostop: februar 2017
3. <http://image-net.org/challenges/LSVRC/2016/results>, zadnji dostop: februar 2017
4. <http://host.robots.ox.ac.uk/pascal/VOC/voc2012/results/index.html>, zadnji dostop: februar 2017
5. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Guadarrama, S. (2014) *Caffe: Convolutional Architecture for Fast Feature Embedding* arXiv:1408.5093
6. Simonyan, K., Zisserman A. (2014) *Very Deep Convolutional Networks for Large-Scale Image Recognition* arXiv:1409.1556
7. <http://host.robots.ox.ac.uk/pascal/VOC/>, zadnji dostop: februar 2017
8. Girshick, R., Donahue, J., Darrell, T., Malik, J. (2014) *Computer Vision and Pattern Recognition: Rich feature hierarchies for accurate object detection and semantic segmentation*
9. <https://blog.heuritech.com/2016/02/29/a-brief-report-of-the-heuritech-deep-learning-meetup-5/>, zadnji dostop: februar 2017
10. Sedej, G. (2011), *Uporaba digitalnega fotoaparata pri multimodalnem nadzoru okolja*, diplomsko delo, Mesto izdaje: Univerza v Mariboru, Fakulteta za elektrotehniko računalništvo in informatiko
11. <https://sites.google.com/site/fourdact/about>, zadnji dostop: februar 2017
12. Ren, S., He, K., Girschak, R., Sun, J. (2015) *Advances in Neural Information Processing Systems: Faster R-CNN Towards Real-Time Object Detection with Region Proposal Networks*
13. <https://helpx.adobe.com/photoshop/how-to/fix-photo-content-aware.html>, zadnji dostop: februar 2017
14. <https://www.theinpaint.com/>, zadnji dostop: februar 2017
15. <http://www.logarithmic.net/pfh/resynthesizer>, zadnji dostop: februar 2017

16. <http://gmic.eu/gimp.shtml>, zadnji dostop: februar 2017
17. http://scikit-image.org/docs/dev/auto_examples/filters/plot_inpaint.html, zadnji dostop: februar 2017
18. <http://docs.opencv.org/2.4/modules/photo/doc/inpainting.html>, zadnji dostop: februar 2017
19. https://github.com/BVLC/caffe/tree/master/models/bvlc_googlenet, zadnji dostop: februar 2017
20. https://github.com/BVLC/caffe/tree/master/models/bvlc_alexnet, zadnji dostop: februar 2017



SLOVENSKI PODJETNIŠKI SKLAD

Slovenski podjetniški sklad (SPS) je javna finančna institucija Republike Slovenije, ustanovljena z namenom dostopa mikro, malim in srednje velikim podjetjem (MSP) do ugodnih finančnih virov na trgu.



KOGA SPS FINANCIRA?

- **MSPje z velikim povpraševanjem**, pri katerih obstaja tržna vrzel pri pridobivanju finančnih sredstev
- **posebne skupine MSP-jev in ostalih podjetij**, ki potrebujejo posebne finančne spodbude za prodor na trg ali rast na določenem ranljivem trgu

KAJ SPS PONUJA?

- **Garancije za zavarovanje bančnih kreditov** SPS nudi skupaj z znižano obrestno mero podjetjem, ki sama nimajo zadostnih jamstev za zavarovanje bančnega kredita. Garancije podjetjem omogočijo lažji, cenejši in hitrejši vir financiranja za razvojne investicije in obratna sredstva.
- **Mikrokrediti** v obliki neposrednih kreditov so na voljo posebnim ciljnim skupinam za likvidno in tekoče poslovanje.
- **Zagonske spodbude** (nepovratna sredstva) so na voljo novonastalim podjetjem:
 - za zagon inovativnih mladih podjetij
 - za zagon podjetij na problemskih območjih
 - za zagon podjetij, ki se ukvarjajo z rabo lesa
- **Semenski kapital** v obliki konvertibilnih posojil in kapitalskih vložkov lahko prejmejo mlada tehnološko inovativna podjetja, ki vstopajo na trg ali želijo na trgu hitro rasti
- **Tvegan kapital** je namenjen visokotehnološkim podjetjem s potencialom hitre globalne rasti

POSLOVNI SERVIS:

1.

HITRE, PREGLEDNE IN ENOSTAVNE INFORMACIJE

- spletna stran: www.podjetniskisklad.si
- info e-mail: info@podjetniskisklad.si
- svetovalni telefon: 02 234 12 60
- e-novice

2.

ENOSTAVNI PRIPOMOČKI ZA PRIPRAVO VLOGE

- elektronski obrazec za prijavo na razpis
- elektronski pripomočki za pripravo poslovnega in amortizacijskega načrta

3.

TRANSPARENTNOST

- vse informacije so javno dostopne
- javni razpisi so objavljeni v Uradnem listu in na spletni strani SPSa

