

# Enovita konvolucijska nevronska mreža za detekcijo anomalij na teksturiranih površinah

Domen Rački<sup>1</sup>, Dejan Tomažević<sup>1,2</sup> in Danijel Skočaj<sup>3</sup>

<sup>1</sup>Sensum, sistemi z računalniškim vidom, d.o.o., Tehnološki park 21, 1000 Ljubljana, Slovenija

<sup>1</sup>Univerza v Ljubljani, Fakulteta za elektrotehniko, Tržaška 25, 1000 Ljubljana, Slovenija

<sup>1</sup>Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Večna pot 113, 1000 Ljubljana, Slovenija

E-pošta: domen.racki@sensum.eu

**Povzetek.** Pristopi, ki temeljijo na globokem učenju, se izkažejo za učinkovitejše od klasičnih pristopov na različnih področjih računalniškega vida. V prispevku uporabimo globoko učenje na domeni avtomatiziranega pregledovanja površin. Predlagamo enovito konvolucijsko nevronska mrežo za detekcijo anomalij na teksturiranih površinah, ki se sestoji iz segmentacijskega in klasifikacijskega dela. Raziščemo, ali je kompaktna arhitektura mreže z manjšim številom parametrov primerna za uporabo na domeni vizualnega pregledovanja. Predlagano arhitekturo evalviramo na podatkovni bazi, ki se sestoji iz različno teksturiranih površin s prisotnimi različnimi in šibko označenimi anomalijami. Predlagani pristop doseže sodobne rezultate tako v smislu segmentacije kot klasifikacije slik glede na prisotnost anomalij na teksturiranih površinah.

**Ključne besede:** globoko učenje, konvolucijske nevronske mreže, segmentacija, detekcija anomalij, klasifikacija

## A unified convolutional neural network for textured-surface anomaly detection

Deep-learning approaches have proven to outperform other non-deep approaches in various computer vision tasks. In this paper we apply deep learning to the domain of automated visual surface inspection. We design a unified convolutional neural-network-based framework for segmentation and detection of surface anomalies. We investigate whether a compact network architecture, with few parameters that need to be learned, is suitable for usage in the visual inspection domain. We evaluate the proposed compact network architecture on a dataset consisting of diverse textured surfaces with variously-shaped weakly-labeled anomalies. With the proposed approach we achieve state-of-the-art results in terms of anomaly segmentation as well as image classification.

**Keywords:** deep learning, convolutional neural networks, segmentation, anomaly detection, classification

## 1 UVOD

Pri nekaterih realnih problemih v računalniškem vidu se domenski strokovnjaki soočijo s težavnim snovanjem formalnih pravil z zadostno kompleksnostjo za modeliranje teh problemov. V takšnih primerih je zmožnost učenja in razpoznavanja vzorcev na avtomatizirani način nepogrešljiva prednost. To velja zlasti tedaj, ko je za določen problem že samo snovanje značilnic težka naloga. Navedeni primeri se pojavljajo pri avtomatiziranem pregledovanju površin izdelkov, kjer kompleksnost modeliranja narašča z naraščajočo kompleksnostjo objektov. V danem kontekstu se kot površinske anomalije

upoštevajo praske, razpoke, umazanija in udrtine, ki se nahajajo na površini danega objekta, medtem ko del kompleksnosti pomeni že sama pojavnost ali tekstura površine objekta. Avtomatizirano pregledovanje igra ključno vlogo pri zagotavljanju kakovosti, tj. odkrivanju anomalij na različnih površinah, kjer je med proizvodnim procesom treba pregledati veliko količino izdelkov z namenom detekcije izdelkov z vidnimi napakami. Glavna vprašanja pri razvoju takšnih sistemov za pregledovanje izhajajo iz dejstva, da je v nekaterih primerih pridobivanje ustreznih reprezentacij, tj. modelov, težavna naloga. Ti in drugi dejavniki dodatno otežujejo nalogo odkrivanja površinskih anomalij ter pripomorejo h kompleksnosti pojavitev do stopnje, da le-te ni mogoče dovolj natančno opisati s snovanjem značilnic. Globoko učenje rešuje to osrednjo težavo pri učenju reprezentacij s kompleksnimi koncepti iz preprostejših konceptov. Izraz *globoko učenje* se uporablja, saj učenje konceptov poteka v hierarhičnem smislu, le-to pa lahko predstavimo kot plasti v grafični predstavitvi modela. Pri sliki se holistično razumevanje le-te doseže na hierarhičen način skozi množico plasti naučenih značilnic. Tipičen primer modela globokega učenja je tako imenovana konvolucijska nevronska mreža (CNN), omenjeni pristop pa se izkaže uporaben v primerih, kjer je že samo snovanje primernih značilnic posebej težak problem.

Glavni prispevek članka je enovita konvolucijska nevronska mreža za segmentacijo in detekcijo anomalij na teksturiranih površinah. Za dani primer površine mreža kot izhod poda segmentacijo anomalije, kateri priredi oceno gotovosti. Oceno gotovosti lahko interpretiramo kot gotovost mreže, da se na danem primeru nahaja

anomalija. Mreža se reprezentacije anomalij lahko nauči iz majhnega nabora grobo označenih učnih primerov. Predlagano rešitev evalviramo na različnih naborih učnih slik podatkovne baze, sestavljene iz različno teksturiranih površin z znotrajrazredno variacijo tako tekstur na ozadju kot anomalij.

## 2 SORODNO DELO

Klasični pristopi za detekcijo napak na površinah objektov, ki ne temeljijo na globokem učenju, sledijo bolj ali manj enaki paradigmi, tj. klasifikator, kot je SVM, naučen na značilnicah, pridobljenih iz predprocesiranih slik. Korak predprocesiranja zagotavlja, da je problem dobro pogojen za ključni proces snovanja primernih značilnic. Nadaljnja pomanjkljivost teh pristopov je, da posebej zasnovane značilnice ni vedno mogoče uporabiti za težave drugega tipa v podobnih domenah. Globoko učenje se razlikuje od omenjenih pristopov z učenjem primernih značilnic namesto ročnega snovanja primernih in običajno suboptimalnih značilnic. In ker so značilnice naučene, se globoke arhitekture lažje posplošijo na druge vrste težav. Omenjena pristopa tako lahko razdelimo v dve skupini, tj. *snovanje značilnic* ter *učenje značilnic*. Pristope snovanja značilnic lahko v grobem razdelimo na: (i) pristope s filtriranjem, kjer se običajno uporabljajo matematične transformacije in filtri [1]; (ii) pristope s strukturiranjem, ki večinoma vključujejo morfološke operacije in detekcijo robov [2]; (iii) pristope z modeliranjem, ki uporabljajo model, kot je skriti markovski model (SMM) [3]; (iv) statistične pristope, ki spadajo med pogosteje uporabljene pristope v problemih klasifikacije z uporabo preprostega upravljanja [4].

Med pristope učenja značilnic pa uvrščamo tako imenovane konvolucijske nevronske mreže (CNN). Motivacija za uporabno CNN izhaja iz prej omenjene težave, kjer se za nekatere realne probleme v računalniškem vidu domenski strokovnjaki soočijo s težavnim snovanjem formalnih pravil, temelječih na geometrijskih in oblikovnih značilnicah za detekcijo anomalij. Pri problemu detekcije anomalij na površinah v [5] avtorji predstavijo mrežo, naučeno z nadzorovanim učenjem na neobdelanih primerih slik jeklenih površin s prisotnimi defekti, s katero izboljšajo točnost detekcije v primerjavi z uporabo značilnic, pridobljenih s klasičnimi metodami za detekcijo napak na teksturiranih materialih. Te vključujejo večplastni perceptron (MLP) in klasifikator SVM z jedrom RBF naučen na značilnicah, pridobljenih s HOG, PHOG, rotacijsko invariantnimi merami lokalne variance in lokalnimi binarnimi vzorci (LBP, LBP-Fourier). Avtorji v [6] predstavijo pregled različnih hevristik CNN za industrijsko pregledovanje. V članku avtorji raziskujejo vpliv različnih hiperparametrov glede na točnost detekcije. Na splošno se avtorji poslužijo arhitektur, objavljenih v [7], in te evalvirajo na podatkovni bazi, sestavljeni iz umetno generiranih slik z različno teksturiranimi površinami in dodanimi

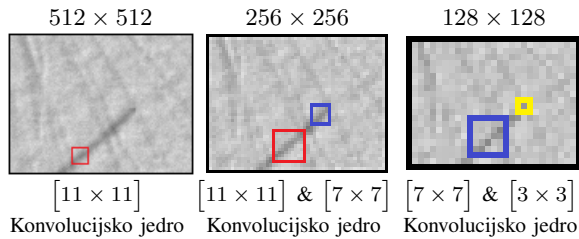
defekti v obliki pik in črt. Za učenje značilnic avtorji z uporabo drsečega okna generirajo veliko učnih primerov z namenom minimiziranja števila napačno klasificiranih slikovnih elementov, upoštevajoč ročno anotirane maske z grobo označenimi slikovnimi elementi, ki prikazujejo anomalijo. Druga dela na pretežno ploskih jeklenih površinah, kot so detekcija napak iz fotometričnih stereo-slik železniških tračnic [8] in ekstrahiranje ustreznih značilnic za detekcijo anomalij [9] kažejo, da je uporaba globokega učenja za ekstrahiranje značilnic v ne-trivialnih domenah mogoča in uporabna.

V literaturi na področju odkrivanja anomalij se tako diskriminativne mreže uporabljajo večinoma kot klasifikator na neobdelanih ali obdelanih slikah, ki ponazarjajo ploskve površine z namenom določiti, ali je na dani sliki prisotna anomalija ali ne. Izhod mreže kot take je torej zgolj klasifikacija celotne slike. Kljub številnim raziskavam na temo segmentacije z uporabo mrež, kot so [10], [11], je uporaba slednjih za namen segmentiranja anomalij redko obravnavana v literaturi detekcij nepravilnosti na površinah.

## 3 ZASNOVA ARHITEKTURE

Pristopi, ki temeljijo na globokem učenju, se izkažejo za uspešnejše od neglobokih pristopov pri reševanju različnih nalog, ki segajo od detekcije [5] do segmentacije [10]. Običajna praksa pri uporabi globokega učenja je prirejanje obstoječih arhitektur, kot na primer VGG [12], za specifično nalogo na način, ki uporabi vnaprej naučene uteži nižjih plasti na velikih učnih bazah, ter douči uteži na višjih plasteh na novi problemski domeni. Na splošno takšen pristop ni najbolj primeren za uporabo na specifičnih problemskih domenah, kot je detekcija napak na teksturiranih površinah. In to zato, ker velike arhitekture implicitno kažejo potrebo po velikih učnih množicah, ki pa jih v določenih industrijskih okoljih ni na voljo, saj je lahko že sam proces pridobivanja teh drago oz. nepraktično opravilo.

Tako lahko določimo kriterije, potrebne za mrežo, namenjeno za uporabo v sistemih za avtomatizirano pregledovanje: (i) kompaktnost, tj. zmožnost učenja potencialne detekcije anomalij iz majhnega števila defektnih učnih primerov; (ii) robustnost, tj. v podobni problemski domeni so potrebni zgolj majhni popravki hiperparametrov; (iii) interpretabilnost, tj. zmožnost vizualne lokalizacije anomalije in razlage klasifikacije domenskemu strokovnjaku. Glede na navedene kriterije predlagamo enovito konvolucijsko nevronske mrežo za detekcijo anomalij na teksturiranih površinah, ki se sestoji iz segmentacijskega in klasifikacijskega dela. Naloga klasifikacijskega dela je, da danemu primeru priredi oceno gotovosti, ki jo lahko interpretiramo kot gotovost mreže, da se na danem primeru nahaja anomalija. Naloga segmentacijskega dela je prikaz vizualne lokalizacije anomalije in razlage klasifikacije domenskemu strokovnjaku. Predlagano arhitekturo mreže prikazuje slika 2.

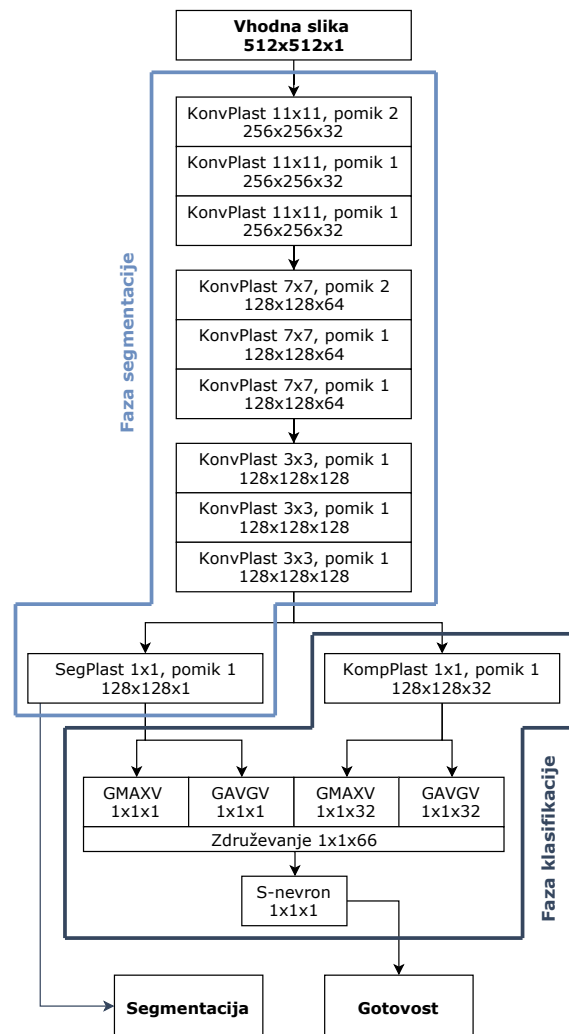


Slika 1: Dimenzije konvolucijskih jeter v primerjavi s skalirano anomalijo površine 4, prikazane na sliki 3. Prva slika prikazuje originalno velikost vhoda v izmeri  $512 \times 512$  slikovnih elementov, medtem ko vsaka nadaljnja slika prikazuje podvzorčeno predhodno sliko. Rdeči kvadrat je filter v velikosti  $11 \times 11$ , modri kvadrat je filter velikosti  $7 \times 7$  in rumeni kvadrat filter v velikosti  $3 \times 3$  slikovnih elementov.

Segmentacijski del mreže je sestavljen iz treh konvolucijskih blokov, vsak blok pa je sestavljen iz treh konvolucijskih plasti. Na splošno se število filtrov podvoji v vsakem konvolucijskem bloku, medtem ko se velikost filtrov zmanjša. Pri izbiri velikosti filtrov lahko naredimo veljavne predpostavke o problemski domeni ter jih prenesemo v zasnovno arhitekturo. Velikost filtra mora biti takšna, da pokrije del anomalije, ki jo želimo zaznati, kot je prikazano na sliki 1. Tu izbiramo velikosti filtrov tako, da se razmerje med soslednimi velikostmi ohranja skozi faze podvzorčenja. Tako zagotovimo, da filter vseskozi pokriva dovolj velik del anomalije. Dejanske reprezentacije v mreži se seveda razlikujejo od prikazanih, ki tu služijo zgolj za ponazoritev in obrazložitev izbire velikosti filtrov. Na splošno podvajanje števila filtrov in redukcija velikosti pri vsakem koraku podvzorčenja v vsakem konvolucijskem bloku ohranja kompaktnost arhitekture v smislu števila parametrov.

Klasifikacijski del mreže se zanaša na segmentacijski del. Ocena gotovosti za dani primer se doseže s pomočjo maksimalne in povprečne globalne združitve, dobljene iz segmentacijske (*SegPlast*) in kompresijske plasti (*KompPlast*), kot je prikazano na sliki 2. Segmentacijska plast je povzeta po popolnoma konvolucijskih mrežah [10] ter zagotavlja segmentacijski izhod iz prejšnje plasti. Kompresijski sloj na drugi strani služi za kompresijo volumna prejšnje plasti, saj se s tem zmanjša število parametrov, iz katerih se oceni gotovost in robustificira klasifikacijski rezultat. Uporaba maksimalnega in povprečnega globalnega združevanja se izkaže za zanesljivo v primerih, ko bi se sicer večje območje brez napak segmentiralo kot območje z napako. Uporaba zgolj maksimalnega združevanja bi v tem primeru neuspešno razlikovala med velikim in veliko manjšim potencialno defektnim segmentiranim območjem. To je glavni razlog, zakaj dodatno opravljamo povprečno združevanje, saj tako lahko ugotovimo razliko v prej omenjenem primeru.

Za dano vhodno sliko teksturirane površine v velikosti  $512 \times 512$  slikovnih elementov predlagana mreža kot izhod poda segmentacijo v velikosti  $128 \times 128$  slikovnih



Slika 2: Predlagana CNN-arhitektura, ki za dani primer kot izhod poda segmentacijo napake in oceno gotovosti, da gre za napako. Kratice KonvPlast, SegPlast in KompPlast označujejo konvolucijsko, segmentacijsko in kompresijsko plast. Kratici GMAXV ter GAVGV označujeta globalno maksimalno in globalno povprečno vzorčenje, medtem ko S-nevron stoji za ocenjevalni nevron.

elementov ter oceno gotovosti, da se na danem primeru nahaja anomalija [13]. Za vsako plastjo v mreži se uporabi aktivacijska funkcija ReLU, ki ji sledi paketna normalizacija. Izjemi sta SegPlast in S-nevron, kjer uporabimo linearno oz. sigmoidno aktivacijsko funkcijo. Vse začetne uteži filtrov se inicializirajo z normalno distribucijo, porazdeljeno okoli ničle, kot je predlagano v [14]. Učenje mreže poteka v dveh fazah. V prvi fazi, tj. fazi segmentacije, naučimo korak segmentacije anomalije skozi 25 epoh. V drugi fazi, tj. fazi klasifikacije, naučimo korak ocene gotovosti anomalije skozi 10 epoh. Ločeno učenje v dveh fazah, prikazano na sliki 2, je ključno. Pri učenju faze segmentacije se tako zamrznejo vse uteži faze klasifikacijske, podobno se pri učenju faze

klasifikacijske zamrznejo vse uteži faze segmentacije. S takšnim načinom učenja zagotovimo, da se faza klasifikacije uči iz smiselnih segmentiranih reprezentacij anomalij. V obeh primerih se mreža uči napovedovanja regresijske vrednosti v intervalu  $[-1, 1]$ , ki jo priredi vsakemu slikovnemu elementu v fazi segmentacije, oz. vrednosti v intervalu  $[0, 1]$ , ki jo priredi posameznemu primeru v fazi klasifikacije. V prvi fazi učenja tako minimiziramo funkcijo srednje kvadratne napake, tj.

$$\mathcal{L}_S = \frac{1}{np} \sum_{i=1}^n \sum_{j=1}^p \|x_i^{(j)} - \hat{x}_i^{(j)}\|^2 \quad (1)$$

kjer  $n$  označuje število primerov,  $p$  številno slikovnih elementov,  $x_i$  anotirano vrednost slikovnih elementov na sliki in  $\hat{x}_i$  napovedano vrednost slikovnih elementov. V fazi klasifikacije minimiziramo binarno navzkrižno entropijo, tj.

$$\mathcal{L}_C = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (2)$$

kjer  $y_i$  označuje anotirano vrednost danega primera in  $\hat{y}_i$  regresijsko napoved za dani primer. Pri obeh kriterijskih funkcijah se uporablja optimizator Adadelta [15] z vrednostjo parametrov kot predlaganih v članku.

#### 4 EKSPERIMENTALNI REZULTATI

Predlagano arhitekturo smo evalvirali na javno dostopni bazi za Industrijsko Optično Pregledovanje DAGM\*, ki se sestoji iz umetno generiranih teksturiranih površin. Kot je razvidno s slike 3, je baza sestavljena iz desetih različnih razredov površin z različnimi anomalijami, ki ponazarjajo razpoke, udrtine in nečistoče. Dani primer označimo kot *pozitiven*, če se na njem nahaja anomalija, sicer dani primer označimo kot *negativen*. Celotna baza se sestoji iz 8050 učnih primerov, od katerih jih 1046 vsebuje anomalijo, in 8050 testnih primerov, od katerih jih 1056 vsebuje anomalijo. Če dana površina v velikosti  $512 \times 512$  slikovnih elementov vsebuje anomalijo, ta vsebuje točno eno šibko označeno anomalijo na teksturiranem ozadju. Šibke labele so podane v obliki elips, ki grobo pokrivajo območje z anomalijo na danem primeru, toda deloma tudi vključujejo območje brez anomalije, kot je razvidno s slike 3. V določenih primerih se lahko izkaže, da premajhno število pozitivnih učnih primerov ni dovolj reprezentativno za dano problematiko, v takšnih primerih uporabimo dodatno bogatene pozitivne učne primere. Natančni postopek bogatenja opišemo v razdelku 4.2.

Čeprav označba v celoti pokriva anomalijo na teksturirani površini, je v nekaterih primerih vanjo zajet znaten delež ozadja. Posledično je del slikovnih elementov napačno označen kot slikovni elementi, ki pripadajo anomaliji, kar pa lahko vpliva na učni proces. Tovrstno

problematiko je sicer pričakovati v številnih realnih problemskih domenah, saj je natančno označevanje težaven in drag postopek. Tudi sicer je dostop do industrijskih baz, ki prikazujejo realne problemske domene otežen, glavni razlogi za to pa ležijo v sporazumih o nerazkritju, ki so namenjeni preprečitvi razkritja določenih procesov, kot so procesi zajemanja slik ali drugi ključni procesi, ki zagotavljajo konkurenčno prednost podjetja.

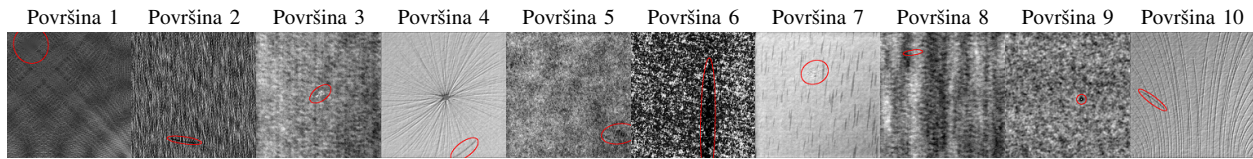
##### 4.1 Vpliv različnih konfiguracij

Segmentacijski del mreže se sestoji iz treh blokov, pri čemer se vsak blok sestoji iz treh konvolucijskih plasti, število filtrov pa se podvoji v vsakem naslednjem bloku. Ker sama velikost filtrov v plasti lahko vpliva na segmentacijsko in klasifikacijsko točnost, naprej proučimo pet različnih konfiguracij mreže: (i) *v1173*, kjer se velikost filtrov v vsakem naslednjem bloku zmanjša, tj.  $[11, 11, 11] \rightarrow [7, 7, 7] \rightarrow [3, 3, 3]$ . (ii) *v3711* pomeni obratno konfiguracijo (i), torej se velikost filtrov poveča v vsakem nadaljnjem bloku, tj.  $[3, 3, 3] \rightarrow [7, 7, 7] \rightarrow [11, 11, 11]$ ; (iii) *c1173* je podoben princip kot (i), s to razliko, da so tu filtri konfigurirani v piramidni shemi v enem bloku, torej se velikost filtrov v vsakem bloku spreminja od večjega proti manjšemu, tj.  $[11, 7, 3] \rightarrow [11, 7, 3] \rightarrow [11, 7, 3]$ ; (iv) *c3711* pomeni nasprotno konfiguracijo (iii), tj.  $[3, 7, 11] \rightarrow [3, 7, 11] \rightarrow [3, 7, 11]$ ; (v) *v333* sledi principu zasnovanih arhitektur VGG [12] s konstantno velikostjo filtrov, tj.  $[3, 3, 3] \rightarrow [3, 3, 3] \rightarrow [3, 3, 3]$ .

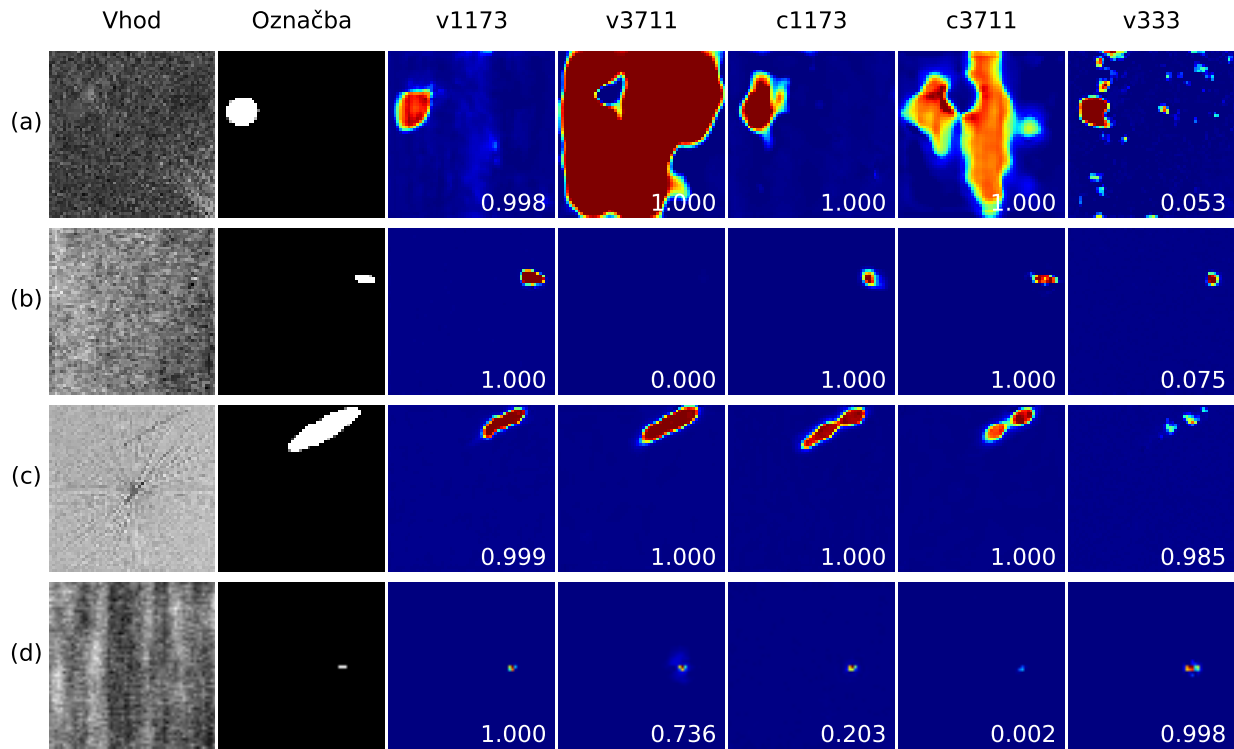
Točnost predlagane mreže evalviramo v smislu deleža pozitivnih primerov, ki so pravilno identificirani kot pozitivni (TPR), ter delež negativnih primerov, ki so pravilno identificirani kot negativni primeri (TNR). Točnost pogledamo tudi v smislu absolutnega števila napačno identificiranih testnih primerov. Prikazani rezultati so dobljeni tako, da na učni množici vzamemo TPR in TNR pri najvišji klasifikacijski točnosti ter s tem pragom klasificiramo primere v testni množici.

Slika 4 prikazuje izhode mreže, tj. segmentacijo anomalije in oceno gotovosti anomalije za različne površine. Iz slednjih je bolj razviden vpliv različnih konfiguracij. Pri konfiguraciji *v333* opazimo artefakte pri segmentiranju ozadja za primer (a), medtem ko je v primeru (c) razvidna slabo segmentirana anomalija. Druge konfiguracije z uporabo večjih filtrov bolje upoštevajo kontekst ozadja, izbrano zaporedje velikost pa vpliva na točnost segmentacije. Konfiguraciji *v1173* in *c1173* upoštevata lokalni kontekst po piramidni shemi, tj. velikost filtra se skozi arhitekturo manjša. Ker se velikost filtra manjša, s podvzorčenjem slike opazimo natančnejše segmentacije. Nasprotno delujeta konfiguraciji *v3711* in *c3711*, kjer se velikost filtra večja skozi arhitekturo, torej majhni filtri iščejo napake na velikih slikah, večji filtri pa na podvzorčenih reprezentacijah v mreži. V nekaterih primerih tu opazimo precej slabo oz. celo napačno segmentirane anomalije, kot je razvidno iz prvega in tretjega primera. Slika 5 prikazuje grafični povzetek

\*<https://hci.iwr.uni-heidelberg.de/node/3616>



Slika 3: Primeri različno teksturiranih površin podatkovne baze. Vsaka površina vsebuje znotrajrazredne variacije v teksturi ozadja in anomaliji. Rdeče elipse ponazarjajo grobe labele, ki označujejo defektno območje.



Slika 4: Segmentacijski izhodi mreže za različne primere. Števila v spodnjem desnem kotu prikazujejo pripisano gotovost mreže, da se na individualnem primeru nahaja anomalija.

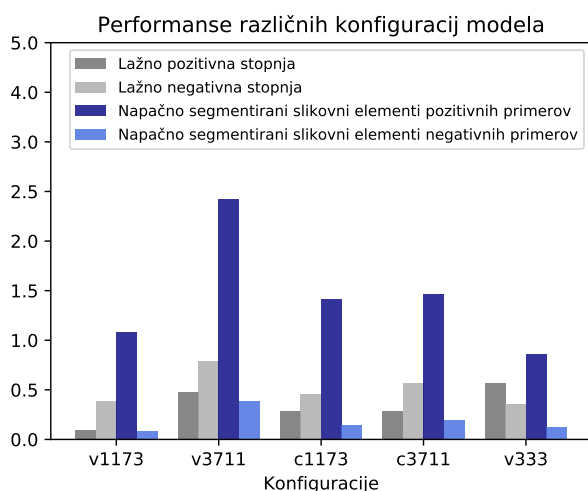
razmerja med segmentacijo šumnega ozadja in segmentacijo anomalije ter točnost mreže pri različnih konfiguracijah konvolucijskih blokov. Razmerje se izračuna iz binariziranih testnih slik z izračunom razmerja med napačno segmentiranimi slikovnimi elementi v ozadju ter vsemi slikovnimi elementi, ki pripadajo ozadju. Slike se binarizirajo pri pragu 0.15, razmerje pa se izračuna kot ločeno povprečje za pozitivne in negativne primere za vsako površino. Kot najboljša se izkaže konfiguracija *v1173*, saj z njo dosežemo najvišjo klasifikacijsko točnost in najnižje razmerje napačno klasificiranih slikovnih elementov v ozadju za negativne primere. Rezultati kažejo, da je takšna konfiguracija filtrov robustna na različne texture v ozadju, segmentirana območja pa po večini pripadajo defektnim regijam. Zato pri nadaljnjih poizkusih uporabljamo to konfiguracijo.

#### 4.2 Velikost učne množice

Pridobivanje primerov z defekti je draga, lahko pa tudi precej zahtevna operacija v realni domeni delovanja, medtem ko so običajno primeri brez defektov na voljo

v veliko večjih količinah. Mreža, zasnovana za kontrolne sisteme, bi se tako morala biti sposobna naučiti razločevanja defektov iz majhnega nabora defektnih primerov. V ta namen predlagano mrežo evalviramo na različnih velikostih učnih množic, opisanih v tabeli 1. Poizkusi se sestojijo iz učenja mreže na zgolj pozitivnih učnih primerih, označenih s *Pos*, ter pozitivnih in negativnih učnih primerih, označenih s *PosNeg*. Ker pa se v določenih primerih lahko izkaže, da premajhno število pozitivnih učnih primerov ni dovolj reprezentativno za dano problematiko, predlagano mrežo naučimo tudi na bogatenih pozitivnih učnih primerih, označenih s *Pos-aug* in *PosNeg-aug*. Pri slednjih se vsak dani pozitivni učni primer rotira za  $180^\circ$  ter zrcali prek horizontalne in vertikalne osi. Za vsak pozitivni učni primer tako dobimo dodatne tri bogatene primere, s čimer trikrat povečamo število pozitivnih primerov v učnih množicah ter pustimo številno negativnih primerov nespremenjeno.

Tabela 2 prikazuje točnost mreže pri učenju z različno velikimi množicami učnih primerov. Kot vidimo pri



Slika 5: Grafični povzetek razmerja med segmentacijo šumnega ozadja in segmentacijo anomalije ter točnost mreže pri različnih konfiguracijah konvolucijskih blokov

Kratica	Uporabljeni učni primeri
<i>Pos</i>	pozitivni učni primeri
<i>PosNeg</i>	pozitivni in negativni učni primeri
<i>Pos-aug</i>	<i>Pos</i> z bogatenjem podatkov
<i>PosNeg-aug</i>	<i>PosNeg</i> z bogatenjem podatkov
Pozitivni primeri	primeri z anomalijami
Negativni primeri	primeri brez anomalij

Tabela 1: Oznake različnih velikosti učnih množic

učanju na učni množici *PosNeg-aug*, mreža zgreši zgolj en pozitiven testni primer, pri čemer ohrani visoko točnost pri detekciji negativnih primerov, tj. primerov brez anomalij. Na splošno se opazita dva trenda pri večanju učne množice. Z bogatenjem pozitivnih primerov se precej poveča točnost detekcije pozitivnih primerov. Z dodajanjem negativnih primerov pa se poveča točnost detekcije tako pozitivnih kot negativnih primerov ter posledično se zmanjšajo artefakti na ozadju pri segmentiranju teksture ozadja dane površine.

Tabela 3 prikazuje primerjavo predlagane arhitekture mreže z obstoječo sodobno rešitvijo, ki temelji na globokem učenju [6], ter s klasičnimi neglobokimi pristopi [16], [17], [18]. S predlagano mrežo izboljšamo detekcijsko točnost pozitivnih, tj. primerov z anomalijami na vseh površinah, pri čemer ohranja visoko točnost detekcije negativnih primerov, tj. primerov brez anomalij. Dosedanji pristopi poleg tega za učenje zahtevajo precej več učnih primerov ter kot del izhoda ne podajo segmentacije anomalije, kar oteži zmožnost morebitne vizualne verifikacije anomalije domenskemu strokovnjaku. Čeprav so slikovni elementi, ki ponazarjajo anomalijo, šibko označeni, saj v določeni meri vključujejo slikovne elemente, ki pomenijo ozadje, se predlagani pristop izkaže kot robusten na tovrstno označevanje in uspe doseči učinkovito segmentacijo anomalij.

## 5 SKLEP

V prispevku predlagamo enovito konvolucijsko nevronska mrežo za detekcijo anomalij na teksturiranih površinah. Nevronska mreža se sestoji iz segmentacijskega in klasifikacijskega dela, katerih učenje poteka v dveh ločenih fazah. Primernost kompaktne arhitekture z manjšim številom parametrov za domeno pregledovanja evalviramo na podatkovni bazi, ki se sestoji iz različno teksturiranih površin s prisotnimi različnimi, šibko označenimi anomalijami. S konfiguracijo *v1173* se izboljšata detekcija in segmentacijska točnost primerov z anomalijami na vseh površinah, pri čemer se ohrani visoka točnost detekcije primerov brez anomalij. Predlagana arhitektura mreže ohranja kompaktnost v smislu števila parametrov in učenja anomalij iz majhnega števila učnih primerov; je robustna v podobni problemski domeni, tj. pri variaciji teksture v ozadju in variaciji anomalij, ter daje vizualno lokalizacijo, tj. segmentacijo anomalije za morebitno razlago klasifikacije domenskemu strokovnjaku.

## ZAHVALA

Raziskavo so omogočili Ministrstvo za gospodarski razvoj in tehnologijo (MGRT) Republike Slovenije, Evropski sklad za regionalni razvoj (ERDF) Evropske unije v okviru programa 631-63/2017/1, ter Sensum, sistemi z Računalniškim Vidom.

## LITERATURA

- [1] W. Polzleitner, "Defect detection on wooden surface using gabor filters with evolutionary algorithm design," in *Neural Networks, 2001. Proceedings. IJCNN'01. International Joint Conference on*, vol. 1. IEEE, 2001, pp. 750–755.
- [2] K.-L. Mak, P. Peng, and K. Yiu, "Fabric defect detection using morphological filters," *Image and Vision Computing*, vol. 27, no. 10, pp. 1585–1592, 2009.
- [3] A. Mukherjee, S. Chaudhuri, P. K. Dutta, S. Sen, and A. Patra, "An object-based coding scheme for frontal surface of defective fluted ingot," *ISA transactions*, vol. 45, no. 1, pp. 1–8, 2006.
- [4] H.-G. Bu, X.-B. Huang, J. Wang, and X. Chen, "Detection of fabric defects by auto-regressive spectral analysis and support vector data description," *Textile Research Journal*, vol. 80, no. 7, pp. 579–589, 2010.
- [5] J. Masci, U. Meier, D. Ciresan, J. Schmidhuber, and G. Fricout, "Steel defect classification with max-pooling convolutional neural networks," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2012, pp. 1–6.
- [6] D. Weimer, B. Scholz-Reiter, and M. Shpitalni, "Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection," *CIRP Annals-Manufacturing Technology*, 2016.
- [7] K. Simonyan and A. Zisserman, "Very deep convolutional neural networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556v6*, 2015.
- [8] D. Soukup and R. Huber-Mörk, "Convolutional neural networks for steel surface defect detection from photometric stereo images," in *International Symposium on Visual Computing*. Springer, 2014, pp. 668–677.
- [9] S. Faghieh-Roohi, S. Hajizadeh, A. Núñez, R. Babuska, and B. De Schutter, "Deep convolutional neural networks for detection of rail surface defects," in *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE, 2016, pp. 2584–2589.

Surface	Pos		PosNeg		Pos-aug		PosNeg-aug	
	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR
1	100 [0]	97.6 [12]	100 [0]	96.4 [18]	98.6 [1]	97.8 [11]	100 [0]	98.8 [6]
2	97.6 [2]	99.6 [2]	98.8 [1]	99.6 [2]	100 [0]	99.8 [1]	100 [0]	99.8 [1]
3	100 [0]	97.8 [11]	100 [0]	97.1 [14]	100 [0]	96.7 [16]	100 [0]	96.3 [18]
4	0.0 [68]	100 [0]	77.9 [15]	95.7 [22]	97.1 [2]	88.4 [59]	98.5 [1]	99.8 [1]
5	100 [0]	99.8 [1]	100 [0]	99.6 [2]	100 [0]	100 [0]	100 [0]	100 [0]
6	100 [0]	99.8 [1]	100 [0]	100 [0]	100 [0]	100 [0]	100 [0]	100 [0]
7	100 [0]	99.7 [3]	100 [0]	98.9 [11]	100 [0]	98.4 [16]	100 [0]	100 [0]
8	100 [0]	99.1 [9]	100 [0]	99.9 [1]	99.3 [7]	99.9 [1]	100 [0]	100 [0]
9	100 [0]	98 [20]	100 [0]	100 [0]	100 [0]	99.9 [1]	100 [0]	99.9 [1]
10	42.7 [86]	100 [0]	100 [0]	99.7 [3]	100 [0]	99.6 [4]	100 [0]	100 [0]
Σ	[156]	[59]	[16]	[73]	[4]	[109]	[1]	[27]

Tabela 2: Klasifikacijska točnost predlagane mreže na različnih velikostih učnih množic, prikazanih v tabeli 1. Števila v oglatih oklepajih označujejo absolutno število napačno klasificiranih primerov za vsako površino, medtem ko števila v oglatih oklepajih na dnu tabele označujejo skupno število napačno klasificiranih primerov.

Surface	Ours		Weimer et. al. [6]		Statistical feat. [16]		SIFT and ANN [17]		Weibull feat. [18]	
	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR
1	100	98.8	100	100	99.7	99.4	100	98.9	98.0	87.0
2	100	99.8	97.3	100	80.0	94.3	91.3	95.7	-	-
3	100	96.3	100	95.5	100	99.5	100	98.5	100	99.8
4	98.5	99.8	98.7	100	96.1	92.5	-	-	-	-
5	100	100	100	98.8	96.1	96.9	100	98.2	100	97.2
6	100	100	99.5	100	96.1	100	100	99.8	100	94.9
7	100	100	-	-	-	-	-	-	-	-
8	100	100	-	-	-	-	-	-	-	-
9	100	99.9	-	-	-	-	-	-	-	-
10	100	100	-	-	-	-	-	-	-	-

Tabela 3: Klasifikacijska točnost predlagane mreže v primerjavi z drugimi globokimi in klasičnimi pristopi

- [10] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.
- [11] G. Papandreou, L.-C. Chen, K. P. Murphy, and A. L. Yuille, "Weakly- and semi-supervised learning of a deep convolutional network for semantic image segmentation," in *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.
- [12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [13] D. Rački, D. Tomažević, and D. Skočaj, "A compact convolutional neural network for textured surface anomaly detection," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, March 2018, pp. 1331–1339.
- [14] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1026–1034.
- [15] M. D. Zeiler, "Adadelta: an adaptive learning rate method," *arXiv preprint arXiv:1212.5701*, 2012.
- [16] B. Scholz-Reiter, D. Weimer, and H. Thamer, "Automated surface inspection of cold-formed micro-parts," *CIRP Annals*, vol. 61, no. 1, pp. 531–534, 2012.
- [17] D. Weimer, H. Thamer, and B. Scholz-Reiter, "Learning defect classifiers for textured surfaces using neural networks and statistical feature representations," *Procedia CIRP*, vol. 7, no. Supplement C, pp. 347–352, 2013, forty Sixth CIRP Conference on Manufacturing Systems 2013.
- [18] F. Timm and E. Barth, "Non-parametric texture defect detection using weibull features," *IS&T SPIE Electronic Imaging. International Society for Optics and Photonics*, 2011.

**Domen Rački** je leta 2015 magistriral na dvojnem študijskem programu Fakultete za računalništvo in informatiko Univerze v Ljubljani in Tehniške univerze v Gradcu. Deluje kot član raziskovalne skupine v Sensum, sistemi z računalniškim vidom ter kot podiplomski študent na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Njegovi glavni raziskovalni interesi segajo na področje praktične uporabe globokega učenja in računalniškega vida za analizo slik.

**Dejan Tomažević** je leta 2005 doktoriral s področja elektrotehnike na Fakulteti za elektrotehniko Univerze v Ljubljani, leta 2008 pa doktoriral na Fakulteti za medicino Univerze v Utrechtu na Nizozemskem. Deluje kot aktivni član raziskovalne skupine v Sensum, sistemi z računalniškim vidom, in kot docent na Fakulteti za elektrotehniko Univerze v Ljubljani. Njegovi glavni raziskovalni interesi segajo na področje vizualne kontrole kakovosti, računalniškega in strojnega vida ter analizo biomedicinskih slik.

**Danijel Skočaj** je izredni profesor na Fakulteti za računalništvo in informatiko Univerze v Ljubljani, kjer je leta 2003 tudi doktoriral. Je predstojnik Laboratorija za umetne vizualne spoznavne sisteme. Njegovi glavni raziskovalni interesi segajo na področja računalniškega vida, razpoznavanja vzorcev, strojnega učenja in spoznavne robotike.