

Konvolucione neuronske mreže u sistemima automatskog upravljanja – pregled stanja u oblasti istraživanja

NATALIJA B. PERIŠIĆ, Univerzitet u Beogradu,

Mašinski fakultet, Beograd

RADIŠA Ž. JOVANOVIĆ, Univerzitet u Beogradu,

Mašinski fakultet, Beograd

Pregledni rad

UDC: 004.8.032.26

681.51:004.4

DOI: 10.5937/tehnika2304433P

Konvolucione neuronske mreže su vrsta dubokih neuronskih mreža koje se koriste u zadacima klasifikacije, identifikacije, predikcije i detekcije objekata, a pogodne su za rad sa ulaznim podacima različitih dimenzija, kao što su signali, slike, video zapisi. O njihovom značaju svedoči činjenica da su u upotrebi više od bilo koje druge vrste dubokih mreža. Upravo zbog toga se konstantno radi na razvoju novih algoritama koji usavršavaju postojeće modele ili kreiranju novih modela koji ubrzavaju ili poboljšavaju proces učenja. Primenu ostvaruju u najrazličitijim oblastima nauke i industrije zbog mogućnosti postizanja visoke tačnosti i jednostavnosti implementacije. U ovom radu se predstavlja struktura konvolucionih mreža, a naročito se razmatraju novosti u sferi istraživanja konvolucionog sloja, gde se tumače različiti tipovi konvolucija. Takođe, posebno se obraća pažnja na upotrebu ovih mreža u sistemima automatskog upravljanja poslednjih godina, kao rezultata pojave Industrije 4.0. Prilikom analiziranja naučnih radova, primena konvolucionih mreža je razgraničena prema dimenzionalnosti ulaznih podataka, odnosno prema dimenzionalnosti mreža i zadacima koje je moguće rešiti pomoću njih.

Ključne reči: *Konvolucione neuronske mreže, konvolucija, konvolucioni sloj, automatsko upravljanje*

1. UVOD

Četvrta industrijska revolucija podrazumeva razvijanje i primenu novih tehnologija u gotovo svim sektorima privrednih delatnosti sa ciljem da se industrijski procesi optimizuju, unaprede, povećaju bezbednost, kvalitet i produktivnost. Osnovne tehnologije koje su zastupljene u industriji 4.0 su 3D štampa (eng. 3D printing), veštačka inteligencija (eng. artificial intelligence), velike baze podataka (eng. big data), internet inteligentnih uređaja (eng. internet of things) i računarstvo u oblaku (eng. cloud computing) [1]. Pod pojmom veštačke inteligencije smatra se skup svih metoda koje omogućavaju da mašina oponaša ljudsku inteligenciju. Mašinsko učenje, kao podskup veštačke inteligencije, podrazumeva da sistem samostalno uči na osnovu sopstvenog iskustva. To znači da sistem sam utvrđuje obrazac za generisanje odgovarajućeg izlaza na osnovu ulaznih podataka, koristeći pri tome statisti-

čki algoritam učenja koji se automatski modifikuje i poboljšava, bez ljudske pomoći [2]. Metoda mašinskog učenja, koja se zasniva na proučavanju i primeni veštačkih neuronskih mreža koje simuliraju analitičke procese u ljudskom mozgu se naziva duboko učenje [3]. Razvoj veštačkih neuronskih mreža počeo je četrdesetih godina prošlog veka, kada su McCulloch i Pitts predstavili prvi model veštačkog neurona [4] i on još uvek traje, te danas postoji mnoštvo različitih modela i arhitektura mreža, specijalizovanih za izvršenje određenih tipova zadataka.

Konvolucione neuronske mreže (eng. Convolutional Neural Networks - CNN) predstavljaju vrstu dubokih neuronskih mreža koje su od svog nastanka do danas našle primenu u najrazličitijim oblastima nauke i industrije, te su neprestano u sferi interesovanja istraživača koji unapređuju i optimizuju modele, razvijaju algoritme i pronalaze im nove primene. Preteča modernijih CNN je Neokognitron [5], veštačka neuronska mreža koja je dizajnirana za prepoznavanje obrazaca. Ideja njenog funkcionisanja je bazirana na istraživanju vizuelnog korteksa mozga mačaka koje su sproveli naučnici Hubel i Wiesel 1962. godine [6], a u kome je dokazano da se određeni neuroni u mozgu pobuđuju isključivo kad su svetlosne trake postavljene

Adresa autora: Natalija Perišić, Univerzitet u Beogradu, Mašinski fakultet, Beograd, Kraljice Marije 16

e-mail: nperisic@mas.bg.ac.rs

Rad primljen: 15.06.2023.

Rad prihvaćen: 24.07.2023.

u odgovarajućem pravcu, dok se sa promenom pravca pobuđuju drugi neuroni. To znači da se aktivacijom određenih neurona dobija jedna informacija o objektu posmatranja, koja se prenosi dalje kako bi se stvorila celokupna slika. Algoritam obučavanja sa povratnim prostiranjem greške (eng. backpropagation algorithm) za obučavanje CNN je prvi put upotrebljen 1989. godine sa ciljem da se mreža osposobi za raspoznavanje rukom pisanih cifara [7]. Ova mreža bila je prototip za razvijanje arhitekture LeNet-5 modela. [8] Kasnije, nakon velikog iskoraka koji je napravljen stvaranjem AlexNet [9] arhitekture, razvijeni su modeli poput ResNet [10], VGGNet [11], GoogleLeNet [12] i mnogih drugih. Konvolucione neuronske mreže, kao efikasno sredstvo kompjuterske vizije, koriste se za unapređivanje procesa u poljoprivredi [13], medicini [14], stomatologiji [15], u ekonomiji [16], itd.

U ovom radu se razmatraju osnovni koncepti CNN i novosti u sferi istraživanja, sa posebnim osvrtom na primenu koje je ovaj tip mreža našao u oblastima automatskog upravljanja.

2. OSNOVNA STRUKTURA KONVOLUCIONIH NEURONSKIH MREŽA

U građu svake CNN ulaze tri različita sloja – konvolucioni sloj (eng. convolutional layer), sloj sažimanja (eng. pooling layer) i potpuno povezani sloj (eng. fully connected layer). Konvolucioni i sloj sažimanja su slojevi čija je uloga izdvajanje karakteristika ulaznih podataka, dok je svrha potpuno povezanog sloja donošenje krajnjeg zaključka o tim podacima, odnosno generisanje izlaznih podataka mreže.

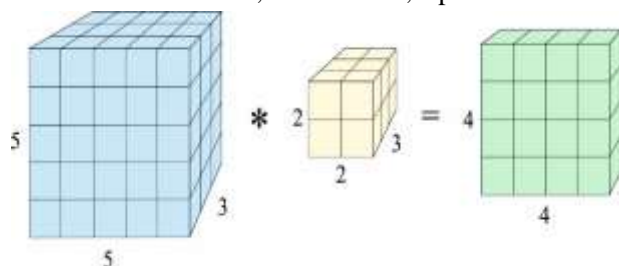
2.1. Konvolucioni sloj

U ovom sloju se odvija operacija konvolucije na ulaznom podatku, kako bi se izdvojile određene karakteristike ulaza. Kao njen rezultat, dobijaju se aktivacione mape (eng. activation map) koje se drugačije nazivaju i mape obeležja (eng. feature map), koje se prenose na ulaz narednog sloja, koji je najčešće sloj sažimanja ili još jedan konvolucioni sloj. U samom procesu konvolucije učestvuju dva elementa, a to su ulazni podatak i filter koji sadrži težinske koeficijente koji se modifikuju u procesu obučavanja. Prvi konvolucioni sloj izdvaja karakteristike nižeg reda, kao što su npr. linije, uglovi ili ivice, dok dublji slojevi izdvajaju složenije, odnosno karakteristike višeg reda [17]. Kako bi se obezbedilo što efikasnije izdvajanje bitnih elemenata ulaznih podataka, moguće je podesiti i hiper-parametre u ovom sloju, kao što su dodavanje nula (eng. zero padding), definisanje pomaka filtra (eng. stride), dimenzije filtra, kao i njegova dubina. U nastavku rada, biće predstavljeni neki od najopštijih tipova konvolucija koji se koriste u modelima CNN. Osim navedenih, postoje i druge modifikacije klasične

konvolucije i neprekidno se radi na istraživanjima konvolucionog sloja, tako da se modeli specijalizuju za rešavanje određenih zadataka, filteri osposobe za izdvajanje važnih karakteristika ulaznih podataka, a uočeni potencijalni nedostaci i mane postojećih modela prevaziđu.

Klasična konvolucija

Poznato je da su kod klasične konvolucije filteri matrice manjih dimenzija u odnosu na dimenziju matrice ulaznog podatka, koje prelaze preko njega sa određenim korakom, pri čemu se vrši matematička operacija, čiji rezultat formira aktivacionu mapu. Najjednostavnije rečeno, rezultat konvolucije je zbir proizvoda preklapljenih elemenata filtra i ulaza. Potrebno je istaći da ukoliko ulazni podatak poseduje osim širine i visine, treću dimenziju - dubinu, tj. broj kanala, klasična konvolucija podrazumeva korišćenje filtra iste dubine. Aktivaciona mapa tada ima samo dve dimenzije, a ukoliko je potrebno povećati njenu dubinu, povećava se broj filtara. Na slici 1 je predstavljen primer jedne klasične konvolucije, gde su dimenzije ulazne matrice $5 \times 5 \times 3$, filtra $2 \times 2 \times 3$, a pomak 1.

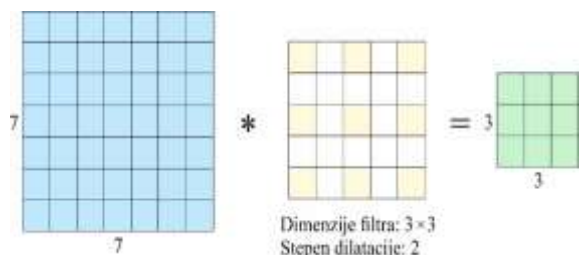


Slika 1 - Klasična konvolucija

Proširena konvolucija

Kod proširene konvolucije (eng. dilated convolution) koriste se filteri koji imaju umetnute nule između težinskih koeficijenata. Broj polja sa umetnutim nulama određen je stepenom dilatacije (eng. dilation rate), koji predstavlja dodatni hiper-parametar konvolucionog sloja u odnosu na mreže sa klasičnom konvolucijom. Prednost proširene konvolucije je što se pokriva veće receptivno polje, pa se mogu ostvariti dobre performanse sa manje slojeva u mreži, čime se smanjuju računski troškovi [18].

Sa druge strane, u [19] navode se i mane primene proširenih filtara. Mane se ogledaju u tome što vrednosti ulaznih podataka koje se uzimaju u obzir mogu biti loši predstavnici okruženja, naročito ako je prisutan šum, kao i što može doći do gubljenja značajnih informacija ako su susedne vrednosti ulaza međusobno nezavisne. CNN sa proširenim konvolucijom našle su primenu u klasifikaciji zvukova [18], segmentaciji slika [19], za uklanjanje šuma sa slika [20], itd. Slika 2 je primer ovog tipa konvolucije gde je dimenzija filtra $3 \times 3 \times 1$, stepen dilatacije 2, pomak 1, a dimenzije ulazne matrice $7 \times 7 \times 1$.



Slika 2 - Proširena konvolucija

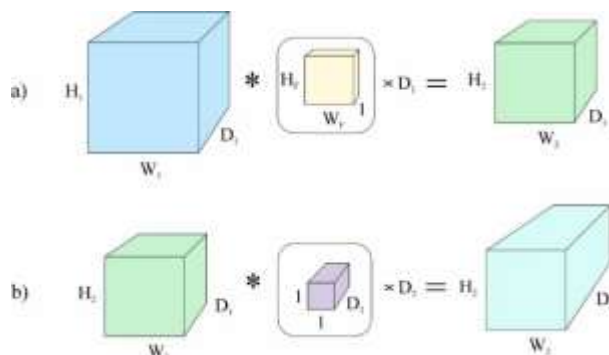
Prostorno odvojiva konvolucija

Prostorno odvojiva konvolucija (eng. spatial separable convolution) podrazumeva zamenu matričnog filtra sa druga dva, čijim se matričnim množenjem dobija polazni filter. Na taj način, obezbeđuje se dobijanje istog rezultata, ali uz primenu manjeg broja parametara, što smanjuje računsku zahtevnost. Smatra se da je upotreba prostorno razdvojive konvolucije prvi put predložena u [21] kada je korišćenje uprošćenih filtara smanjilo vreme trajanje obuke mreže do pet puta, pri čemu je sposobnost klasifikacije i prepoznavanja ostala gotovo ista.

Dubinski odvojiva konvolucija

Dubinski odvojiva konvolucija (eng. depthwise separable convolution) faktorizuje regularnu konvoluciju upotrebom dubinske konvolucije (eng. Depthwise convolution) i tačkaste konvolucije (eng. Pointwise convolution), sledstveno [22]. Slojevi sa dubinskom konvolucijom sadrže filtre kojih ima onoliko koliko ulazni podatak ima kanala, ali čija je dubina 1. Svaki filter vrši operaciju samo nad jednim kanalom ulaza. Na taj način se čuva dubina u aktivacionoj mapi. Nakon ovoga, slede slojevi tačkaste konvolucije, gde se koriste filtri dimenzije $1 \times 1 \times D_1$, gde je D_1 dubina ulaznog podataka, pa nova aktivaciona mapa zadržava dimenzije širine i visine, ali dubina iznosi 1. Naravno, povećanje dubine izlaza vrši se povećanjem broja filtara. Smisao dubinske odvojive konvolucije je smanjenje složenosti proračuna, što ubrzava obučavanje jer je potrebno podesiti manji broj parametara [22]. Broj potrebnih operacija množenja kod ovog tipa konvolucije je $D_2 \cdot H_F \cdot W_F / (D_2 + H_F \cdot W_F)$ puta manji od broja potrebnih množenja kod klasične konvolucije, za dobijanje istog rezultata, pri čemu je D_2 dubina izlazne aktivacione mape tj. broj filtara tačkaste konvolucije, a H_F i W_F visina i širina filtra dubinske konvolucije, sledstveno.

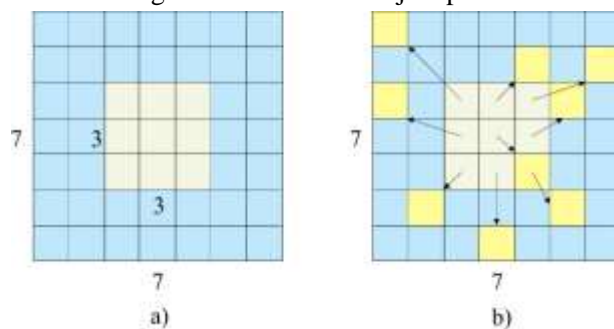
Dubinski odvojivu konvoluciju koriste mreže MobileNets, koje predstavljaju mreže kreirane za upotrebu u aplikacijama mobilnih uređaja za detekciju objekata, klasifikaciju atributa lica, geolokalizaciju velikih razmera, itd [22]. Slika 3 prikazuje princip dubinske odvojive konvolucije, gde su H_1 i H_2 visina ulaza i izlazne aktivacione mape, a W_1 i W_2 širina ulaznog podataka i izlazne aktivacione mape.



Slika 3 - Dubinski odvojiva konvolucija: a) dubinska konvolucija i b) tačkasta konvolucija

Deformabilna konvolucija

Jedan od potencijalnih nedostataka klasičnih CNN je njihova ograničenost u domenu geometrijskih transformacija, s obzirom na činjenicu da je njihova geometrijska struktura fiksna [23]. Ovaj problem se prevazilazi upotrebom mreža sa deformabilnom konvolucijom (eng. deformable convolution). Ona omogućava 2D promenu pozicija težinskih koeficijenata filtra, koja omogućava da filter zauzme slobodnu formu prilikom izdvajanja karakteristika ulaznog podataka, za razliku od klasične konvolucije, gde su težinski koeficijenti filtra uvek u kvadratnoj ili pravougaonoj formi. Pomeranje težinskih koeficijenata se odvija u okviru polja pomeranja (eng. offset field) koje je istih dimenzija kao i ulazni podatak. U toku obučavanja mreže se istovremeno podešavaju težinski koeficijenti filtra, kao i njihova pomeranja [23]. Mreže sa deformabilnom konvolucijom su se odlično pokazale u zadacima detekcije objekata i semantičke segmentacije [23]. Na slici 4 je prikazano poređenje klasičnog i deformisanog filtra na aktivacionoj mapi.

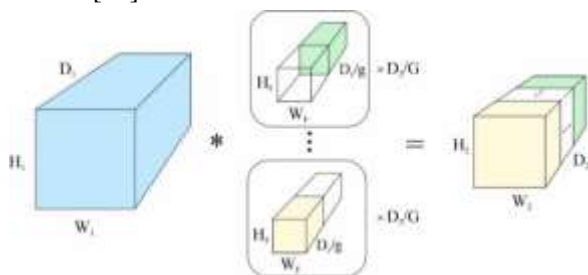


Slika 4 - a) Pozicija filtra kod klasične konvolucije, b) Pozicija filtra kod deformabilne konvolucije

Grupna konvolucija

Grupna konvolucija (eng. group convolution) je prvi put primenjena u AlexNet [9] mreži kako bi se prevazišao problem memorijskih ograničenja računara [9]. Ona podrazumeva korišćenje različitih grupa filtara čija dubina pokriva samo deo dubine ulaznog podataka, u zavisnosti koliko različitih grupa filtara se koristi. Svaka grupa vrši konvoluciju nad ulaznim

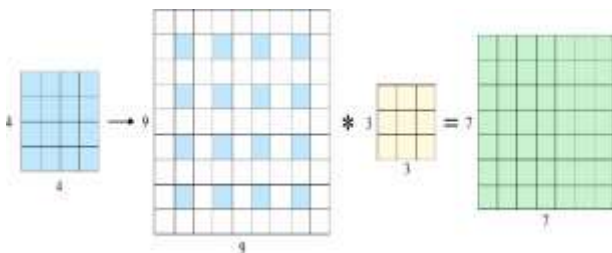
podatkom zasebno, a aktivaciona mapa, koja se dobija na izlazu je sastavljena od rezultata svih grupa filtara [24]. Samim tim, dubina dobijene aktivacione mape je jednaka zbiru primenjenih filtara svake grupe. Broj parametara i izračunavanja je G puta manji u odnosu na klasičnu konvoluciju, gde je G broj različitih grupa filtara. Iako je smanjenje broja potrebnih parametara prednost grupne konvolucije, u radu [24] se navodi i njena mana. Naime, kako je svaki izlazni kanal u vezi sa posebnom grupom kanala ulaznog podatka, a ne sa svim, informacije između kanala se ne prenose što može dovesti do gubitka značajnih informacija. Slika 5 ilustruje grupnu konvoluciju, gde simboli H , W i D označavaju visinu, širinu i dubinu, a indeksi 1, 2 i F se odnose na ulazni podatak, izlaznu aktivacionu mapu i filter, sledstveno. Grupna konvolucija implementirana je u ResNeXt [25] arhitekturu namenjena za klasifikaciju slika i zadatke prepoznavanja, ali je svoju primenu našla i u identifikaciji i klasifikaciji zlonamernih softvera [26].



Slika 5 - Grupna konvolucija

Transponovana konvolucija

Transponovana konvolucija (eng. transposed convolution) se u literaturi često izjednačava sa dekonvolucijom (eng. deconvolution) [27, 29], što nije potpuno tačno jer je dekonvolucija matematički definisana kao inverzna konvolucija, što nije slučaj sa transponovanom konvolucijom [28, 29]. Njen smisao je generisanje aktivacione mape većih dimenzija od ulazne aktivacione mape [29], što se odvija na sledeći način. Najpre se između svih redova i kolona ulaznog podatka dodaju nule, a potom se oko, na taj način proširenog ulaznog podatka primeni tehnika dodavanja nula. Broj dodatih nula se određuje na osnovu željene dimenzije izlazne aktivacione mape. Nakon ovoga se sprovodi klasičan proces konvolucije, sa pomakom jednakim jedinici. Ovaj proces je ilustrovan na slici 6.

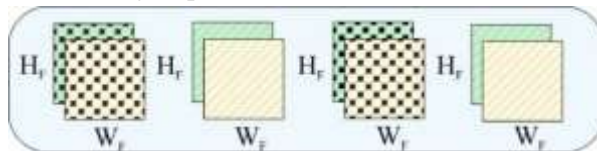


Slika 6 - Transponovana konvolucija

Transponovana konvolucija je korišćena u hibridnim modelima za predikciju [28], proširivanje ulaznog skupa podataka [30] za postizanje veće tačnosti i bolje generalizacije modela, itd.

Popločana konvolucija

Klasične konvolucione neuronske mreže koriste tzv. vezane težinske koeficijente (eng. tied weights), odnosno svaki neuron istog sloja deli iste težinske koeficijente, što doprinosi smanjenju parametara koji se uče i unosi translacionu invarijantnost u model [31]. Ipak, klasične mreže su veoma ograničene kada se radi o učenju drugih invarijantnosti, kao što su invarijantnosti razmere i rotacije, pa u [31], autori razvijaju novi model, koji koristi popločanu konvoluciju (eng. tiled convolution) sa ciljem da se baš te invarijantnosti uspešno nauče. Kod mreža sa popločanom konvolucijom, aktivacione mape se multiplikuju i grupišu kako bi se naučile invarijantne osobine uzoraka, a u okviru istog sloja se podešavaju filteri sa različitim težinskim koeficijentima [27], pri čemu se težinski koeficijenti filtra vezuju, odnosno ponavljaju sa određenim korakom, k . Parametar k određuje veličinu pločice čiji filteri neće imati iste težinske koeficijente, što znači da vrednost $k = 1$ određuje klasičnu CNN [31]. Slika 7 predstavlja konvolucioni sloj sa popločanom konvolucijom. Isti dizajn označava iste težinske koeficijente, dok ista boja ukazuje na pripadnost istoj aktivacionoj mapi.



Slika 7 - Konvolucioni sloj sa popločanom konvolucijom

Nakon što se mreža sa popločanom konvolucijom uspešno pokazala na testiranju sa CIFAR-10 setom podataka za klasifikaciju i NORB setom za detekciju objekata [31], formiran je i hibridni model upotrebljen za klasifikaciju teksta [32], dok je 3D popločana konvolucija primenjena za segmentaciju volumetrijskih slika [33].

2.2. Sloj sažimanja

U okviru sloja sažimanja odvija se nelinearna operacija kojom se grupa vrednosti na određenoj lokaciji aktivacione mape predstavlja jednom vrednošću [34]. Svrha ovog sloja je smanjenje dimenzija njegovog ulaza, čime se smanjuje računaska zahtevnost, a proces obučavanja pojednostavljuje i ubrza. Benefiti sažimanja aktivacionih mapi ogledaju se i u tome što se izdvajaju najvažnije karakteristike, dok se manje važne zanemaruju, pa se postiže određeni stepen translatorne invarijantnosti, a sposobnost generalizacije povećava [35]. Sloj sažimanja u arhitekturi CNN

najčešće sledi nakon konvolucionog sloja. Dva najčešća načina kojima se redukuje dimenzija aktivacionih mapa u ovom sloju su sažimanje maksimalnom vrednošću (eng. max pooling) kod koga se svaka zahvaćena grupa vrednosti predstavlja najvećom od njih i prenosi na izlaz, i sažimanje srednjom vrednošću (eng. average pooling), pri čemu se na izlaz prenosi srednja prosečna vrednost zahvaćene grupe. U radu [34] se osim ovih, detaljno predstavljaju još neke metode sažimanja koje se javljaju u CNN i vrši se njihovo poređenje. Izveden je opšti zaključak da su performanse različitih metoda sažimanja visoko zavisne od same arhitekture mreže i aktivacione funkcije kojom se unosi nelinearnost u model.

2.3. Potpuno povezani sloj

U strukturi CNN nakon poslednjeg konvolucionog ili sloja sažimanja najčešće sledi jedan ili više potpuno povezanih slojeva. Kao što im samo ime kaže, svaki neuron ulaznog sloja je povezan sa svakim neuronom narednog sloja, pri čemu svaku vezu karakterišu različiti težinski koeficijenti, odnosno, nema deljenja zajedničkih težinskih koeficijenta, kao što je to slučaj sa koeficijentima konvolucionog sloja. U okviru ovog sloja se formira izlazna veličina celog modela. Ipak, prisustvo ovog sloja nije neophodno u CNN jer on može biti zamenjen sa 1×1 konvolucionim slojem [27]. Uticaj broja potpuno povezanih slojeva i broja neurona u njima na performanse CNN u klasifikaciji slika je istraživano u [36], gde je pokazano da CNN sa manjim brojem podesivih parametara zahtevaju upotrebu više neurona u potpuno povezanom sloju, kao i veći broj slojeva za klasifikaciju kada se mreža obučava na skupu podataka sa više klasa i manje uzoraka po svakoj klasi, nego kad se obučavajući skup sastoji od većeg broja slika, ali manje klasa. Pokazano je i da je kod CNN sa većim brojem podesivih parametara potrebno manje neurona u ovom sloju, bez obzira na tip obučavajućeg skupa.

3. PODELA KONVOLUCIONIH MREŽA PREMA DIMENZIONALNOSTI I NJIHOVA PRIMENA

3.1. Jednodimenzionalne konvolucione mreže - 1D CNN

Jednodimenzionalne CNN su vrsta CNN kod kojih se filter sa težinama kreće samo duž jedne ose, formirajući jednodimenzionalni niz na svom izlazu. Koristi se kod ulaznih podataka koji su dvodimenzionalni, odnosno, kada se mogu predstaviti vektorom određene dubine. Shodno tome, najčešće se koriste u zadacima identifikacije, klasifikacije i obrade signala, predikcije vremenskih nizova i slično.

Ovaj tip mreža našao je naročitu primenu u sistemima čovek-mašina. Konkretan primer može se naći u [37], gde je zadatak mreže bio da klasifikuje

signale dobijene elektroencefalogramom (EEG) kako bi se postavio željeni cilj mobilnom robotu. U upravljačkoj šemi između čoveka i robota, zadatak čoveka bio je vizualizuje rotirajući objekat kad se robot pozicionira u pravcu cilja, što EEG zabeležio, a 1D CNN je uspešno razlikovala te signale od signala stanja opuštenosti, odnosno onih kad robot nije bio adekvatno pozicioniran i to sa tačnošću od 92.04%, što je daleko više od tačnosti postignute upotrebom drugih klasifikatora. Istraživanje [38] koristi 1D CNN kako bi se na osnovu elektromiografskog signala, koji prikazuje električni potencijal mišića čoveka, izvršila predikcija potrebne sile za hvatanja predmeta. Potom, vrši se dinamička rekonstrukcija sile na robotskom zahvatnom mehanizmu, tako što se kao željena vrednost na ulazu PID kontrolera dovodi izlaz iz konvolucione neuronske mreže. Postignutom tačnošću od 99% u offline testiranju i 83% u online testiranju, nadmašene su metode linearne regresije i 2D CNN korišćene u istu svrhu.

Sistemi upravljanja temperaturom u pametnim zgradama, usredsređeni na svakog korisnika pojedinačno, takođe mogu da se oslanjaju na 1D CNN, kako bi se optimizovao toplotni komfor i potrošnja energije. U radu [39] jednodimenzionalna CNN se koristi kao deo upravljačkog sistema za automatizovano prepoznavanje fizičke aktivnosti stanara na osnovu akcelerometrijskih podataka sa narukvice, što se koristi kao ulaz u model podržanog učenja (eng. reinforcement learning model) za upravljanje temperature u prostoriji u realnom vremenu. Primenom predložene metode upravljanja toplotna nelagodnost stanara se smanjila za 10,9% uz zadržavanje energetske potrošnje, dok ostvarena tačnost 1D CNN modela iznosi od 95,8% do 100% za svakog od korisnika.

Još jedna značajna uloga koju mogu imati 1D CNN u sistemima upravljanja je detekcija kvara i rekonstrukcija signala sa mernih uređaja. Rad [40] pokazuje implementaciju 1D CNN u kombinaciji sa dugom kratkoročnom memorijom (eng. long short-term memory) za rešavanje ovog zadatka na primeru izlaznih signala Holovih senzora DC motora bez četkica. Model za detekciju kvara dostiže tačnost od 98%, dok model predložen za rekonstrukciju signala ostvaruje minimalnu tačnost od 95%, čime je potvrđena njihova efikasnost.

3.2. Dvodimenzionalne konvolucione mreže – 2D CNN

Klasične CNN, koje se najčešće sreću u literaturi su upravo 2D CNN i one su objašnjene u prethodnom poglavlju.

Primena 2D CNN u sistemima upravljanja je sve više u porastu, naročito kad se radi o navigaciji mobilnih robota, gde postoji mnogo načina za njeno

ostvarenje. Tako, na primer, mreža može biti obučena za predviđanje uglova točkova, kao u [41], gde je postignuta tačnost od 85,03%. Takođe, ustanovljeno je da je mreža pogodna za učenje i obavljanje raznovrsnih zadataka na putevima sa i bez obeleženih traka, kao i to da se njenom primenom može zameniti upotreba konvencionalnog PID kontrolera. Istraživanje [42] takođe koristi 2D CNN kao kontroler, ali specijalno obučenu za izbegavanje prepreka na putu. Mreža obučena za izbegavanje prepreka u laboratoriji postiže tačnost od 95% do 97%, dok druga mreža, za izbegavanje prepreka u hodniku ostvaruje tačnost od 93-96%. Prepoznata je i mana upotrebe neuro kontrolera u ovu svrhu, koja se ogleda u nemogućnosti pamćenja prethodno preduzetih radnji, što može dovesti do kolizije sa preprekama, u zavisnosti od njihove pozicije, ali rad nudi i potencijalna rešenja za navedeni problem. Mreža specijalizovana za detekciju prepreka korišćena je u [43], gde se nakon identifikacije jedne od prepreka, pristupa utvrđivanju distance između nje i robota kako bi se primenio odgovarajući algoritam upravljanja za njeno izbegavanje, što je u ovom slučaju bilo reaktivno upravljanje. Studija [44] nudi nov pristup stereo vizuelnog upravljanja neholomnim mobilnim robotom zasnovan na semantičkim informacijama o okruženju robota. Za semantičku segmentaciju korišćene su potpune konvolucione neuronske mreže (eng. fully convolutional networks), koje su ostvarile srednju tačnost od 98,72%. Kako se navodi, značajne prednosti semantičke segmentacije u stereo vizuelnom upravljanju su bolje predstavljanje informacija o okruženju robota, kao i činjenica da se ona može koristiti i u svrhu obavljanja drugih zadataka, kao što su izbegavanje prepreka, prepoznavanje objekata, lokalizacija, itd.

Postoje mnoge tehnike upravljanja robotskim manipulatorima koje uključuju 2D CNN. Tako je u [45] osmišljen sistem upravljanja gde neuronska mreža konstantno otkriva i vrši procenu trenutnog položaja objekta od interesa, kako bi manipulator zauzeo odgovarajuću poziciju u odnosu na njega. Greška u poziciji stimuliše promenu vrednosti upravljanja brzine zglobova, koja dovodi do smanjenja greške i na taj način se postiže pravilno pozicioniranje manipulatora. Studija [46] za praćenje putanje manipulatora kombinuje 2D CNN, čija je uloga online identifikacija nepouzdanosti sistema i FOTSM (eng. fractional-order terminal sliding-mode) kontroler zadužen za održavanje stabilnosti manipulatora. Predloženom strategijom upravljanja postiže se velika robusnost, eliminiše pojava zveketanja, uspešno se kompenzuju poremećaji i rešen je problem singularnosti.

Osim za upravljanje u domenu robotike, dvo-dimenzionalne CNN mogu se koristiti prilikom upravljanja procesa proizvodnje 3D štampom, kako bi

poboljšala geometrijska tačnost izrađenih delova, što je pokazano u [47]. Naime, mreža je deo zatvorene sprege sa dva PID kontrolera za upravljanje brzinom dodavanja materijala i temperaturom i njena uloga je da klasifikuje ispravno urađen gornji sloj prilikom štampe. Upravljanja se menjaju u skladu sa parametrima PID kontrolera sve dok izlaz iz mreže ne pokaže da je sloj ispravno odštampan. Tačnost modela iznosi 90.08%.

Autori [48] prepoznaju da je problem u korišćenju linearnih kvadratnih regulatora (eng. linear quadric regulator, LQR) u upravljanju vibracijama taj što fiksni parametri ne mogu da daju zadovoljavajuće rezultate kada postoji više različitih tipova vibracija u okruženju. Kao rešenje problema, predlaže se 2D CNN za identifikaciju vrste vibracije, na osnovu čega se postavljaju parametri LQR, prethodno optimizovani metaheurističkim algoritmom za svaku vrstu vibracije. Mreža na svom ulazu koristi pikselizovanu matricu talasnih koeficijenata, a prosečna tačnost prelazi 98%.

3.3. Trodimenzionalne konvolucione mreže – 3D CNN

Kod 3D CNN, filter se kreće u tri pravca, dok ulazni podaci mogu imati i određenu dubinu, kao četvrtu dimenziju koja ih određuje, a to su najčešće videi, snimci magnetne rezonance, itd. Rezultat ovakve konvolucije zauzima trodimenzionalni prostor, najčešće u obliku kvadra. Zbog prirode ulaznih podataka, 3D CNN se najviše primenjuju na robotskim sistemima ili kod autonomnih vozila. Kada se radi o upravljanju robotskih manipulatora, rad [49] prepoznaje da 2D CNN ograničavaju smer hvatanja, čineći ga fiksnim i predlažu 3D CNN prilikom hvatanja nepoznatih objekata pomoću mekih zahvatnih uređaja jer se omogućava procena pozicije hvatanja iz više pravaca i sa različitom orijentacijom zglobova. Ovakav sistem postigao je uspešnost izvršenja zadatka od 87%, a takođe je i robusan u radu sa šumovitim podacima.

Razumevanje semantičkih scena je veoma značajno za navigaciju i upravljanje mobilnim robotima, pa autori [50] predlažu 3D CNN za prepoznavanje i klasifikaciju 3D objekata kako bi se prevazišao nedostatak 2D CNN da se ne razmatra reprezentacija objekta sa različitih pozicija posmatranja robota. Zanimljivo je da se nakon prikupljanja podataka kamere pristupa 3D modeliranju objekta tako da se on predstavi kao volumetrijski prikaz u cilindričnoj mreži, što odstupa od uobičajenog prikaza u formi 3D matrice, nakon čega se uči klasifikacija. Ovakav pristup pokazao je bolje performanse od klasičnog pristupa.

Na kraju, 3D CNN su pokazale tačnost od 80% u zadatku predviđanja verovatnoće najkraće trajektorije

kretanja autonomnog vozila [51], što ukazuje na to da se mogu koristiti za razvoj modela upravljanja kretanjem takve vrste vozila.

4. ZAKLJUČAK

U okviru ovog rada dat je pregled istraživanja u oblasti CNN. S obzirom na to da je istraživanje u vezi sa temom veoma opširno, rad je ograničen isključivo na tumačenje strukture mreža i to naročito u domenu konvolucionog sloja. Predstavljeno je nekoliko različitih varijacija u pogledu konvolucije koja se odvija u tom sloju, a koje su načešće u upotrebi. Takođe, izneti su benefiti modifikovanih konvolucija u odnosu na klasičnu prilikom izvršenja određenih zadataka.

Drugi deo rada odnosi se na primenu koju CNN imaju u sistemima automatskog upravljanja, kao neizostavnog dela industrije 4.0. Prikom razmatranja naučnih radova, autori su izvršili podelu primene na osnovu dimenzionalnosti mreža, kako bi se stekao jasniji uvid u vrstu zadataka koji se mogu rešiti primenom odgovarajućeg tipa mreže.

Kao rezultat istraživanja, može se izvesti zaključak da je primena CNN u automatskom upravljanju sve zastupljenija i da njihov značaj na tom polju raste, naročito kada se radi o upravljanju robotskim sistemima. Ipak, s obzirom da je potencijal ovih mreža veliki, postoji još mnogo prostora za ispitivanje njihove primene u različitim sistemima.

5. ZAHVALNICA

Ovde prikazani rezultati su rezultat istraživanja podržanog od strane Ministarstva nauke, tehnološkog razvoja i inovacija RS po Ugovoru 451-03-47/2023-01/200105 od dana 03. 02. 2023. godine.

LITERATURA

- [1] Chun K. W, Kim H, Lee K. A study on research trends of technologies for industry 4.0; 3D printing, artificial intelligence, big data, cloud computing, and internet of things, in *Advanced Multimedia and Ubiquitous Engineering: MUE/Future Tech 2018*, Springer, Singapore, pp. 397-403, 2019.
- [2] Sharma N, Sharma R, Jindal N. Machine learning and deep learning applications – a vision, *Global Transitions Proceedings*, Vol. 2, No. 1, pp. 24-28, 2021.
- [3] Xin Y, Kong L, Liu Z, Chen Y, Li Y, Zhu H, Gao M, Hou H, Wang C. Machine learning and deep learning methods for cybersecurity, *IEEE Access*, Vol. 6, pp. 35365-35381, 2018.
- [4] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity, *The bulletin of mathematical biophysics*, Vol. 5, pp. 115-133, 1943.
- [5] Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biological Cybernetics*, Vol. 36, No. 4, pp. 193-202, 1980.
- [6] Hubel D. H, Wiesel T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *The Journal of Physiology*, Vol. 160, No. 1, pp. 106-154, 1962.
- [7] LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation applied to handwritten zip code recognition, *Neural Computation*, Vol. 1, No. 4, pp. 541-551, 1989.
- [8] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278-2324, 1998.
- [9] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks, *Communications of the ACM*, Vol. 60, No. 6, pp. 84-90, 2017.
- [10] He K, Zhang S, Ren S, Sun J. Deep residual learning for image recognition, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 770-778, 2016.
- [11] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition, *arXiv preprint*, arXiv:1409.1556, 2014.
- [12] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1-9, 2015.
- [13] Kamilaris A, Prenafeta-Boldu FX. A Review of the Use of Convolutional Neural Networks in Agriculture, *The Journal of Agricultural Science*, Vol. 156, No. 3, pp.312-322, 2018.
- [14] Anwar S. M, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical image analysis using convolutional neural networks: a review, *Journal of Medical Systems*, Vol. 42, pp. 1-13, 2018.
- [15] Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: A scoping review, *Journal of Dentistry*, Vol. 91: 103226, 2019.
- [16] Hosaka T. Bankruptcy prediction using imaged financial ratios and convolutional neural networks, *Expert Systems with Applications*. Vol. 117, pp. 287-299, 2019.
- [17] Hijazi S, Kumar R, Rowen C. Using convolutional neural networks for image recognition, *Cadence Design Systems Inc.:* San Jose, CA, USA, 2015.

- [18]Zhang X, Yuexian Z, Wei S. Dilated convolution neural network with LeakyReLU for environmental sound classification, in *2017 22nd International Conference on Digital Signal Processing DSP*, IEEE, London, UK, pp.1-5, 23-25 August 2017.
- [19]Ziegler T, Fritsche M, Kuhn L, Donhauser K. Efficient smoothing of dilated convolutions for image segmentation, *arXiv preprint*, arXiv:1903.07992, 2019.
- [20]Wang Y, Wang G, Chen C, Pan Z. Multi-scale dilated convolution of convolutional neural network for image denoising, *Multimedia Tools and Applications*, Vol. 78, pp. 19945-19960, 2019.
- [21]Mamalet F, Garcia C. Simplifying ConvNets for faster learning, in *Proc. 22nd International Conference on Artificial Neural Networks*, Lausanne, Switzerland, pp. 58-65, 11-14 September 2012.
- [22]Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Anam H. Mobilenets: Efficient convolutional neural networks for mobile vision applications, *arXiv preprint*, arXiv:1704.04861, 2017.
- [23]Dai J, Qi H, Xiong Y, Li Y, Zhang G, Hu H, Wei Y. Deformable convolutional networks, in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, pp. 764-773, 22-29 October, 2017.
- [24]Zhou Y, Chen S, Wang Y, Huan W. Review of research on lightweight convolutional neural networks, in *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, Chongqing, China, pp. 1713-1720, 12-14 June, 2020.
- [25]Xie S, Girshick R, Dollar P, Tu Z, He K. Aggregated residual transformations for deep neural networks, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, Hawaii, pp. 1492-1500, 21-26 July, 2017.
- [26]Albahar M. A, ElSayed M. S, Jurcut A. A modified ResNeXt for Android malware identification and classification, *Computational Intelligence and Neuroscience*, Vol. 2022, 2022, doi: 10.1155/2022/8634784.
- [27]Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J, Chen T. Recent advances in convolutional neural networks, *Pattern Recognition*, Vol. 77, pp. 354-377, 2018.
- [28]Bukka S. R, Gupta R, Magee A. R, Jaiman R. K. Assessment of unsteady flow predictions using hybrid deep learning based reduced-order models, *Physics of Fluids*, Vol. 33, No. 1, 013601, 2021.
- [29]Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning, *arXiv preprint*, arXiv:1603.07285, 2016.
- [30]Cui Z, Zhang M, Cao Z, Cap C. Image data augmentation for SAR sensor via generative adversarial nets, *IEEE Access*, Vol. 7, pp. 42255-42268, 2019.
- [31]Nigam J, Chen Z, Chia D, Koh P, Le Q, Ng A. Tiled convolutional neural networks, in *Proceedings of the Advances in Neural Information Processing Systems*, Vancouver, Canada, Vol. 23, pp. 1279-1287, 6-9 December, 2010.
- [32]Trusca MM, Spanakis G. Hybrid tiled convolutional neural networks (HTCNN) text sentiment classification, in *Proceedings of the 12th International Conference on Agents and Artificial Intelligence*, Vellela, Malta, Vol. 2, pp. 506-513, 22-24 February, 2020.
- [33]Zeng G, Zheng G. 3D tiled convolution for effective segmentation of volumetric medical images, in *22nd International Conference on Medical Image Computing and Computer Assisted Intervention*, Shenzhen, China, pp. 146-154, 13-17 October, 2019.
- [34]Zafar A, Aamir M, Nawi NM, Arshad A, Riaz S, Alruban A, Dutta AK, Almotairi S. A comparison of pooling methods for convolutional neural networks, *Applied Sciences*, Vol. 12, No. 7, 8643, 2022.
- [35]Wu H, Gu X. Towards dropout training for convolutional neural networks, *Neural Networks*, Vol. 71, pp. 1-10, 2015.
- [36]Basha SHS, Dubey SR, Pulabaigari V, Mukherjee S. Impact of fully connected layers on performance of convolutional neural networks for image classification, *Neurocomputing*, Vol. 378, pp. 112-119, 2020.
- [37]Izzuddin TA, Safri NM, Othman MA. Mental imagery classification using one-dimensional convolutional neural network for target selection in single-channel BCI-controlled mobile robot, *Neural Computing and Applications*, Vol. 33, pp. 6233-6246, 2021.
- [38]Xue J, Lai KWC. Dynamic gripping force estimation and reconstruction in EMG-based human-machine interaction, *Biomedical Signal Processing and Control*, Vol. 80, 104216, 2023.
- [39]Jung S, Jeoung J, Hong T. Occupant-centered real-time control of indoor temperature using deep learning algorithms, *Building and Environment*, Vol. 208, 108633, 2022.
- [40]Chu KSK, Chew KW, Chang YC. Fault-diagnosis and fault-recovery system of Hall sensors in brushless DC motor based on neural networks, *Sensors*, Vol. 23, No. 9, 4330, 2023.

- [41] Dangskul W, Phattaravatin K, Rattanaporn K, Kijaidure Y. Real-time control using convolutional neural network for self-driving cars, in *2021 7th International Conference on Engineering, Applied Sciences and Technology*, online, pp. 125-128, 1-3 April, 2021.
- [42] Thomas A, Hedley J. FumeBot: A deep convolutional neural network controlled robot, *Robotics*, Vol. 8, No. 3, 62, 2019.
- [43] Dos Reis DH, Welfer D, De Souza Leite Cuadros MA, Gamarra DFT. Mobile robot navigation using an object recognition software with RGBD images and the YOLO algorithm, *Applied Artificial Intelligence*, Vol. 33, No. 14, pp. 1290-1305, 2019.
- [44] Jokić A, Petrović M, Miljković Z. Semantic segmentation based stereo visual servoing of nonholonomic mobile robot in intelligent manufacturing environment, *Expert Systems with Applications*, Vol. 190, 116203, 2022.
- [45] Al-Shanoon A, Lang H. Robotic manipulation based on 3-D visual servoing and deep neural networks, *Robotic and Autonomus Systems*, Vol. 152, 104041, 2022.
- [46] Zhou M, Feng Y, Xue C, Han F. Deep convolutional neural network based fractional-order terminal sliding-mode control for robotic manipulators, *Neurocomputing*, Vol. 416, pp. 143-151, 2020.
- [47] Lyu J, Manoochehri S. Online convolutional neural network-based anomaly detection and quality control for fused filament fabrication process, *Virtual and Physical Prototyping*, Vol. 16, No. 2, pp. 160-177, 2021.
- [48] Lu X, Liao W, Huang W, Xu Y, Chen X. An improved linear quadratic regulator control method through convolutional neural network-based vibration identification, *Journal of Vibration and Control*, Vol. 27, No. 7-8, pp. 839-853, 2021.
- [49] Choi C, Schwarting W, DelPreto J, Rus D. Learning object grasping for soft robot hands, *IEEE Robotics and Automation Letters*, Vol. 3, No. 3, pp. 2370-2377, 2018.
- [50] Moon J, Kim H, Lee B. View-point invariant 3D classification for mobile robots using a convolutional neural network, *International Journal of Control, Automation and Systems*, Vol. 16, No. 6, pp. 2888-2895, 2018.
- [51] Pardhi P, Yadav K, Shrivastav S, Sahu SP, Dewangan DK. Vehicle motion prediction for autonomus navigation system using 3 dimensional convolutional neural network, in *Proceedings of the 5th International Conference on Computing Methodologies and Communication*, Erode, India, pp. 1322-1329, 8-10 April, 2021.

SUMMARY

CONVOLUTIONAL NEURAL NETWORKS IN AUTOMATIC CONTROL SYSTEMS – THE STATE-OF-THE-ART

Convolutional neural networks are type of deep neural networks used for classification, identification, prediction and object detection. They are suitable for dealing with input data of various dimensions, such as signals, images and videos. Their importance is confirmed by the fact that they are used more than any other type of deep networks. This is the reason for constant development of new algorithms that improve existing models or creation of new models that accelerate or ameliorate learning process. They are utilized in a wide range of scientific and industrial fields due to their possibility of achieving high accuracy and simplicity of implementation. In this paper structure of convolutional networks is presented and, in particular, novelties in the study of convolutional layer are discussed, where different types of convolution are interpreted. Additionally, special attention has been paid to the use of these networks in control systems in recent years, as a result of the occurrence of Industry 4.0. During scientific work analysis, convolutional networks application are divided according to the dimensionality of input data, that is, according to the dimensionality of networks and the tasks that they can solve.

Key Words: Convolutional neural networks, convolution, convolutional layer, automatic control