

PRIMENA VEŠTAČKE INTELIGENCIJE ZA IDENTIFIKACIJU STANJA ELEKTRIČNOG BROJILA**APPLICATION OF ARTIFICIAL INTELLIGENCE TO IDENTIFY THE STATE OF THE ELECTRIC METER**

Aleksandar Đurić, *Fakultet tehničkih nauka, Novi Sad*

Oblast – ELEKTROTEHNIKA I RAČUNARSTVO

Kratka sadržaj – Skeniranje slika i pretvaranje skeniranih informacija u digitalni format je aktivna oblast istraživanja. Skeniranje je automatizovan, brz i efikasan proces u poređenju sa tradicionalnim unosom podataka. Prepoznavanje cifara sa slika je izazovan zadatak. Tradicionalni pristupi za rešavanje ovog problema obično razdvajaju korake lokalizacije, segmentacije i prepoznavanja. Ovaj rad predstavlja jedinstven pristup koji integriše ova tri koraka korišćenjem duboke konvolucione neuronske mreže koja radi direktno na pikselima slike.

Ključne reči: Veštačka inteligencija, Mašinsko učenje, Duboko učenje, Neuronske mreže

Abstract – Scanning images and converting the scanned information into digital format is an active research area. Scanning is an automated, fast and efficient process as compared to traditional data entry. Recognizing digits from images is a challenging task. Traditional approaches to solve this problem typically separate the localization, segmentation, and recognition steps. This paper presents a unified approach that integrates these three steps via the use of a deep convolutional neural network that operates directly on the image pixels.

Keywords: Artificial intelligence, Machine learning, Deep learning, Neural networks

1. UVOD

Problem prepoznavanja oblika je ljudima veoma blizak, sa njime se sreću tokom celog života, počevši od najranijeg detinjstva. Prepoznavanje oblika nije urođena osobina nego se stiče učenjem. Sposobnost prepoznavanja oblika ljudi razvijaju i usavršavaju gotovo celog svog života. Do nedavno su samo živi organizmi imali sposobnost prepoznavanja oblika. U poslednje vreme, zahvaljujući razvoju informacionih tehnologija i veštačke inteligencije, ostvareni su konkretni rezultati u oblasti prepoznavanja oblika. Pojam teorije prepoznavanja oblika podrazumeva matematičke metode, prvenstveno namenjene automatskom klasifikovanju konkretnih objekata. Teorija prepoznavanja oblika oslanja se uglavnom na upotrebu metoda veštačke inteligencije kao što su: neuronske mreže, rasplinuta logika (eng. fuzzy logic), matematička logika, statističke metode, itd.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio dr Aleksandar Selakov, docent.

U ovom radu opisano je rešenje za identifikaciju stanja starih (nedigitalnih) brojila električne energije tj. identifikaciju tarifnih brojeva sa fotografija nedigitalnih električnih brojila. Tema ovog rada je opisivanje sistema koji funkcioniše na principu najnovijih tehnologija i koji se uz jednostavnu pripremu ulaznog skupa podataka i konfiguraciju samih parametara sistema može primeniti za rešavanje različitih slučajeva optičkog prepoznavanja karaktera. Cilj ovog rada jeste implementacija softvera koji pomaže popisivačima električne energije tako što ne moraju da prepisuju stanje brojila, već samo fotografiju brojila, a softver prepozna stanje. Takođe, pomoću ovog softvera nedigitalna brojila je moguće pretvoriti u poludigitalna brojila, tako što bi se pored brojila stavila kamera koja na svakih 15 minuta snima stanje, koje zatim softver prepoznaje i prosleđuje dalje rezultate. Na ovaj način se simulira ponašanje modernih digitalnih brojila, a istovremeno i eliminiše potreba za popisivačima električne energije.

2. BAZA ULAZNIH PARAMETARA MREŽE

Na fotografijama u bazi prikazana su stara (nedigitalna) brojila električne energije (*Slika 1*). Cilj sistema je da za datu fotografiju brojila detektuje tarife i prikaže brojeve sadržane u okviru tarifa. Kako ne postoji javno dostupna baza podataka koja u sebi sadrži slike starih brojila, pre same realizacije sistema bilo je neophodno prikupiti odgovarajući broj fotografija. Uspešno je prikupljeno oko 2.500 fotografija. Nakon procesa prikupljanja fotografija, važno je bilo odraditi i proces filtriranja fotografija tj. uklanjanje fotografija sa šumovima (mutne fotografije ili fotografije kod kojih nisu jasno vidljive cifre na jednoj od tarifa). Posle procesa filtriranja, bazu podataka sačinjava skup od 2.330 fotografija.



Slika 1. Primer fotografije za prepoznavanje

Korišćene su tri metode preprocesiranja: konvertovanje kolor fotografija u monohromatske, smanjivanje dimenzija i augmentacija.

Za augmentaciju je korišćena skripta [1] pomoću koje je nad svakom ulaznom fotografijom izvršeno pet operacija: rotiranje u levo, rotiranje u desno, uveličavanje fotografije, iskrivljenje u levu ili desnu stranu i iskrivljenje unazad ili unapred, te je stoga i skup podataka uvećan pet puta. Proces filtriranja je pimenjen na fotografije dobijene nakon augmentacije zbog ponovne pojave fotografija sa šumom. Konačno, skup podatak za obuku i validaciju je imao 13.177 monohromatski fotografija starih brojala veličine 180x180 piksela.

Model obučen prethodno opisanim skupom podataka nije generisao zadovoljavajuće rezultate prepoznavanja.

Detaljnijom analizom ulaznih fotografija došlo se do zaključka da one sadrže, za svrhe ovog rada, veliki broj bespo- trebnih informacija. Na osnovu slike (*Slika 1*) vidimo da pored tarifnih cifara, postoje delovi ulaznih fotografija (uglavnom je to donji levi ugao i/ili donji desni ugao) koji takođe sadrže određen niz cifara. Sistem sam vrši skeniranje u potrazi za karakterima za prepoznavanje, nema predefini- sane regione. Ulazne fotografije koje pored tarifnih cifara poseduju dodatne regije u okviru kojih se nalaze određeni nizovi cifara, vrlo verovatno, će zbuniti model prilikom obuke tj. pomeriće fokus modela sa tarifnih cifara na druge delove fotografije, a to dalje rezultira lošim prepoznavanjem.

Kako bi povećali preciznost prepoznavanja potrebno je, na neki način, ukloniti suvišne informacije sa ulaznih fotografija. Ovaj problem je rešen izdvajanjem regija od interesa sa ulaznih parametara mreže. Ideja izdvajanja regija od interesa je preuzeta iz rada [2].

Originalni skup podataka ima 818 fotografija na kojima se nalazi jednotarifno brojilo i 1.512 fotografija na kojima su dvotarifna brojila. Na 8 fotografija jednotarifnih brojila, zbog pozicija iz kojih su snimljene, nije bilo izvodljivo izdvojiti tarifu. Nakon završetka izdvajanja regija od interesa sa originalnih slika (*Slika 1*), dobijen je novi skup podataka koji sadrži 3.834 fotografije na kojima se nalaze samo tarife (*Slika 2*).



Slika 2. Primer izdvojene tarife

2.1 Iteracija 1: 128x128 bez augmentacije

Prvobitnom izmenom, fotografije na kojima se nalaze tarife su sa originalne veličine smanjene na veličinu 128x128 piksela, a potom konvertovane u monohromatske fotogra- fije. Značajan napredak je postignut u kvalitetu prepozna- vanja. Ovo je razlog zbog kojeg proces augmentacije nije realizovan nad ovim skupom podataka. Pozitivan pomak u kvalitetu prepoznavanja bio je samo dodatni motiv za dalje eksperimentisanje sa dimenzijama ulaznih parametara mreže.

2.2 Iteracija 2: 96x96 bez augmentacije

Kako bi se našao kompromis između neophodnih resursa i kvaliteta prepoznavanja, originalne fotografije, na kojima se nalaze tarife, u ovoj iteraciji su smanjene na veličinu 96x96 piksela. Fotografije su nakon modifikacije dimenzija konvertovane u monohromatske. Rezultati prepoznavanja ove iteracije su poprilično dobri i iz tog razloga proces augmentacije nije primenjen.

2.3 Iteracija 3: 64x64 bez augmentacije

Kvalitetni rezultati prethodne iteracije bili su samo dodatni motiv za dalje eksperimentisanje sa dimenzijama ulaznih parametara mreže. Kako su rezultati prethodne iteracije poboljšani smanjenjem dimenzija fotografija, doneta je odluka da se nastavi sa istom paradigmom. Dakle, originalne fotografije, na kojima se nalaze tarife, su u okviru ove iteracije smanjene na veličinu 64x64 piksela. Osim smanjena veličine, smanjen je i kvalitet fotografija, što je uticalo na kvalitet same obuke. Fotografije su nakon korekcije dimenzija konvertovane u monohromatske. Odluka da se nastavi istom paradigmom se ispostavila kao pogrešna, jer su rezultati prepoznavanja ove iteracije za nijansu lošiji u odnosu na prethodnu iteraciju. Rezultati prepoznavanja ove iteracije, sami za sebe, su zadovoljavajući, te stoga nije bilo potrebe za realizovanjem procesa augmentacije.

3. ARHITEKTURE MREŽE

U cilju ostvarivanja što boljih rezultata procesi pretpro- cesiranja, učenja i testiranja su ponavljani u više navrata sa različitim parametrima i različitim arhitekturama neuronske mreže. Od svih izmena napravljenih tokom pomenutih procesa izdvajaju se dve arhitekture modela.

Svaka korišćena arhitektura je posedovala konvolucionu neuronsku mrežu (KNM) čiji je zadataka da vrši segmentaciju fotografija i izdvajanje ključnih osobina koje sadrže karaktere za prepoznavanje, a zatim da ih prosledi, zavisno od arhitekture, potpuno povezanoj neuronskoj mreži (PPNM) ili povratnoj neuronskoj mreži sa dugotrajnom kratkoročnom memorijom (PNMDKM) koje za zadatak imaju klasifikaciju dobijenih ključnih osobina i da kao rezultat daju karaktere sa fotografije u digitalnoj formi.

3.1 KNM–PPNM arhitektura

Konvoluciona neuronska mreža (KNM) je izgrađena pomoću gradivnih blokova koji imaju sledeću strukturu:

- Konvolucioni sloj
- Sloj aktivacione funkcije
- Konvolucioni sloj
- Sloj aktivacione funkcije
- Sloj sažimanja maksimalnom vrednošću po obe dimenzije

Arhitektura konvolucione neuronske mreže se sastoji od tri prethodno opisana gradivna bloka.

Potpuno povezana neuronska mreža kao ulaz zahteva vektor podataka. Iz tog razloga se izlazni tenzor konvolucione neuronske mreže dalje propušta kroz sloj za izravnavanje (eng. Flatten), koji od izlaza konvolucione mreže pravi jednodimenzionalni vektor.

Potpuno povezana neuronska mreža (PPNM) je izgrađena pomoću sledećih slojeva:

- Potpuno povezani sloj
- Sloj aktivacione funkcije
- Potpuno povezani sloj
- Sloj aktivacione funkcije
- Sloj za sprečavanje preobučavanja (eng. Dropout)

Skup podataka pored fotografija sa petocifrenim tarifama sadrži i nekoliko desetina fotografija na kojima se nalaze šestocifrene tarife, te iz tog razloga mreža ima šest

klasifikacionih slojeva. Broj klasa koje neuronska mreža treba da raspozna je 11, brojevi iz opsega [0,10]. Za identifikaciju tarifnih cifara koriste se brojevi iz opsega [0,9], a broj 10 služi kao indikator na osnovu kog se zna da li se radi o petocifrenoj ili šestocifrenoj tarifi.

3.2 KNM-PPNM arhitektura

Konvoluciona neuronska mreža i klasifikacioni slojevi su identični kao kod KNM-PPNM arhitekture. Prema tome, ovde se fokus stavlja na povratnu neuronsku mrežu.

Izlazni oblik (eng. output shape) tenzora konvolucione neuronske mreže nije upotrebljiv kao ulaz u povratnu neuronsku mrežu sa dugotrajnom kratkoročnom memorijom. Izlazni tenzor konvolucione neuronske mreže ima 4 dimenzije, a povratna neuronska mreža sa dugotrajnom kratkoročnom memorijom kao ulaz očekuje 3 dimenzije. Ovaj problem je rešen upotrebom sloja za promenu dimenzija (eng. Reshape).

Povratna neuronska mreža sa dugotrajnom kratkoročnom memorijom (PNMDKM) je izgrađena pomoću sledećih slojeva:

- Sloj povratne neuronske mreže sa dugotrajnom kratkoročnom memorijom
- Sloj aktivacione funkcije
- Sloj povratne neuronske mreže sa dugotrajnom kratkoročnom memorijom
- Sloj aktivacione funkcije

Izlaz povratne neuronske mreže sa dugotrajnom kratkoročnom memorijom se prosleđuje klasifikacionim slojevima.

4. PROCES UČENJA SISTEMA

U mašinskom učenju, cilj je postići dobru generalizaciju modela, odnosno da se model uspešno izvršava nad podacima koji nisu viđeni ranije. Ukoliko je trenirani model previše dobro prilagođen podacima kojima je treniran, može doći do stanja previše prilagođenog modela (eng. overfitting) [3-5]. S obzirom da se algoritam prostiranja greške unazad u praksi pokazao pouzdanim i računski isplativim, korišćen je za treniranje mreža u ovom radu. Prilikom treniranja mreža, računa se funkcija gubitka (eng. loss) te preciznost (eng. accuracy) upotrebom kategoričke unakrsne entropije (eng. categorical crossentropy). Takođe, tokom treniranja se koristi optimizacioni algoritam Adam.

Skup podataka je podeljen tako da je za obuku neuronske mreže odvojeno 95% od ukupnog skupa, dok je preostalih 5% služilo za validaciju mreže.

4.1 Iteracija 1: 128x128 bez augmentacije

Kao najbolji model u okviru ove iteracije pokazao se model sa KNM-PPNM arhitekturom. Za obuku je korišćeno 3.642 monohromatske fotografije veličine 128x128 piksela koje su podeljene u pakete (eng. batch) od 32 fotografije, a izabrani broj epoha je 50. Treniranje modela je trajalo 110 minuta.

4.2 Iteracija 2: 96x96 bez augmentacije

Za ovu iteraciju najbolji model je izgrađen pomoću KNM-PPNM arhitekture. Trening skup, od 3.642 monohromatske fotografije veličine 96x96 piksela, je podeljen na pakete od 32 fotografije, dok je izabrani broj epoha iznosio 100. Kriterijum zaustavljanja je postavljen na 100 epoha. Treniranje modela je trajalo 150 minuta.

4.3 Iteracija 3: 64x64 bez augmentacije

Za treću iteraciju skup za obučavanje, od 3.642 monohromatske fotografije veličine 64x64 piksela, je podeljen u pakete od po 32 fotografije, dok je definisan broj epoha bio 50. Trening faza je trajala 32 minuta. Za poslednju iteraciju, model sa KNM-PPNM arhitekturom se pokazao kao najkvalitetniji.

5. TESTIRANJE MREŽA I REZULTATI

Testiranje je izvršeno nad skupom fotografija koje su naknadno prikupljene i koje nisu učestvovala u trening fazi, a isto tako ni u validacionoj fazi. Korišćenjem sasvim novog skupa podataka za testiranje se proverava da mreža nije pretrenirana i da nije naučila neke karakteristike specifične samo za podatke u okviru trening skupa.

Za svrhe testiranja prikupljeno je 44 fotografije brojila električne energije. Skup od 44 fotografije se sastoji od 38 fotografija na kojima se nalaze dvotarifna brojila i 6 fotografija na kojima su jednotarifna brojila. U slučaju fotografija sa dvotarifnim brojičima, na 5 fotografija su brojila sa šestocifrenim tarifama, a na preostale 33 fotografije su brojila sa petocifrenim tarifama. Svaka fotografija sa jednotarifnim brojičima sadrži petocifrenu tarifu. Dakle, može se zaključiti da ukupan broj karaktera koje sistem treba da prepozna iznosi 420. Odnos prepoznatih karaktera i ukupnog broja za prepoznavanje uzet je kao parametar uspešnosti sistema. Pored ovoga analizirane su još neke karakteristike: uspešnost prepoznavanja kompletne tarife, prepoznavanje u odnosu na cifru koja treba da se prepozna i u odnosu na poziciju u tarifi.

Kada uporedimo rezultate tri iteracije treniranja i testiranja mreže (Tabela 1) primetno je da pretposlednja izmena doprinosi unapređenju obučenosti mreže, dok poslednja izmena za nijansu narušava obučenost mreže.

Tabela 1. Statistika uspešnosti prepoznavanja po iteraciji

	Očekivano	Prepoznato	Procenat uspešnosti
Iteracija 1	420	396	94.29%
Iteracija 2	420	404	96.19%
Iteracija 3	420	400	95.24%

Sistem je u iteraciji 1 obučen da bolje uočava cifre na fotografijama u odnosu na preostale dve iteracije. Smanjenjem dimenzija ulaznih parametara sistema pojavljuju se fotografije na kojima sistem prepoznaje jednu ili dve cifre zavisno od iteracije (Tabela 2).

Tabela 2. Uspešnost prepoznavanja kompletne tarife

Prepoznato cifara	Iteracija 1	Iteracija 2	Iteracija 3
0	0	0	0
1	0	0	1
2	0	1	3
3	4	5	1
4	15	3	5
5	54	63	62
6	9	10	10
Ukupno:	82	82	82

U testnom skupu najmanje se pojavljuje broj osam, samo 18 puta. Sve tri iteracije su uspele da prepoznaju broj osam sa stoprocentnom tačnošću. U tabeli (Tabela 3) je pokazano da su iteracije 2 i 3 u potpunosti prepoznale četiri cifre, 1-2-8-9 u slučaju iteracije 2 i 4-6-8-9 u slučaju iteracije 3, dok je iteracija 1 u potpunosti prepoznala 3 cifre, 1-4-8.

Tabela 3. Uspešnost prepoznavanja u odnosu na cifru koja se prepoznaje

Očekivana cifra	Broj cifara	Iteracija 1	Iteracija 2	Iteracija 3
0	48	46	47	46
1	38	38	38	37
2	28	27	28	27
3	58	53	54	49
4	45	45	43	45
5	74	62	68	68
6	37	35	35	37
7	41	40	40	40
8	18	18	18	18
9	33	32	33	33
Ukupno:	420	396	404	400

Prilikom prepoznavanja stanja električnog brojlara važno je da bar prve tri tarifne pozicije u slučaju petocifrenih brojlara, a prve četiri tarifne pozicije u slučaju šestocifrenih brojlara budu dobro prepoznate. Na primer, sa slike (Slika 2) se vidi da je stanje tarife: 28246, a neka je model prepoznao: 29246. Dakle, model jeste pogrešio samo jednu cifru ali pozicija na kojoj je došlo do zabune je važnija. U ovom slučaju, greška koja se propagira na stvarni svet je reda 1000 kWh, što nije malo ni sa elektroenergetske ni sa finansijske tačke gledišta. Ovo je razlog zbog kojeg je uveden kriterijum uspešnosti prepoznavanja broja u odnosu na poziciju (Tabela 4).

Tabela 4. Uspešnost u odnosu na poziciju broja koji se prepoznaje

Pozicija	Broj cifara	Iteracija 1	Iteracija 2	Iteracija 3
1	82	70	80	75
2	82	81	81	81
3	82	80	80	82
4	82	79	76	76
5	82	76	77	76
6	10	10	10	10
Ukupno:	420	396	404	400

6. ZAKLJUČAK

Glavni cilj ovog rada bio je implementacija softvera za identifikaciju stanja električnog brojlara. Pod identifikacijom stanja se podrazumeva pretvaranje karaktera, u ovom slučaju brojeva, sa fotografija u njihove digitalne ekvivalente. Dakle, identifikacija stanja suštinski predstavlja problem koji je dosta rasprostranjen u modernom svetu – optičko prepoznavanje karaktera.

U radu su opisane metode za rešavanje ovog problema koje podrazumevaju pretprocesiranje ulaznih podataka, njihovu segmentaciju, zatim prepoznavanje ključnih osobina za detekciju i klasifikaciju karaktera sa fotografije.

Prikazane su najbolje metode za rešavanje datog problema iz perspektive neuronskih mreža.

Na osnovu rezultata testiranja najboljih arhitektura svake iteracije ustanovljeno je da se neuronska mreža, sa KNM-PNMDKM arhitekturom iz iteracije 2, istakla po svim segmentima prepoznavanja, te je iz tog razloga izabrana kao najkvalitetnija neuronska mreža. Rezultati testiranja izabrane neuronske mreže su sledeći, preciznost prepoznavanja individualnih cifara iznosila je 96.19%, a preciznost prepoznavanja kompletnog niza cifara je iznosila 89.02%, mreža je od 82 tarife uspešno prepoznala 73. Dobijeni rezultati govore o tome da je cilj u uvodu postignut. Razvijen je softver koji, korišćenjem najnovijih tehnologija, sa velikom tačnošću identifikuje stanje električnog brojlara.

7. LITERATURA

- [1] <https://github.com/mdbloice/Augmentor>, skripta za augmentaciju, poslednji pristup 14.12.2022.
- [2] Muhammad Asif, Maaz Bin Ahmad, Shiza Mushtaq *et al.* "Long Multi-digit Number Recognition from Images Empowered by Deep Convolutional Neural Network". The Computer Journal, Volume 65, Issue 10, October 2022, Pages 2815–2827. 13 September 2021.
- [3] Tetko, I. V., Livingstone, D. J. and Luik, A. I. „Neural network studies. 1. Comparison of overfitting and overtraining," Journal of chemical information and computer sciences, 1995, 35.5: 826-833.
- [4] Dietterich, T. „Overfitting and undercomputing in machine learning," ACM computing surveys (CSUR), 1995, 27.3: 326-327.
- [5] Caruana, R., Lawrence, S., Giles, CL. „Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping." In: Advances in neural information processing systems. 2001. p. 402- 408.

Kratka biografija:



Aleksandar Đurić rođen je u Šapcu 1998. godine. Osnovne akademske studije na Fakultetu tehničkih nauka Univerziteta u Novom Sadu upisao je 2017. godine. Diplomirao je 2021. godine i iste godine upisuje master akademske studije.

Kontakt: acadjuric98@gmail.com