

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 1029

**SUSTAV ZA DETEKCIJU I
RASPOZNAVANJE PROMETNIH ZNAKOVA**

Ivan Kovaček

Zagreb , siječanj 2010

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 1029

SUSTAV ZA DETEKCIJU I RASPOZNAVANJE PROMETNIH ZNAKOVA

Ivan Kovaček

Zagreb , siječanj 2010

SADRŽAJ

1	Uvod	1
2	Umjetne neuronske mreže	2
2.1	Motivacija.....	2
2.2	Umjetni neuron.....	3
2.3	Arhitektura unaprijedne neuronske mreže.....	4
2.4	Učenje.....	5
2.4.1	Gradijentni spust.....	5
2.4.2	Propagacija greške unatrag.....	7
2.4.3	Ubrzanje konvergencije.....	8
2.4.4	Validacija.....	9
3	Značajke slike	10
3.1	Obrada slike.....	10
3.1.1	Ispravljanje smetnji uzrokovanih preplitanjem slike.....	10
3.1.2	Filtriranje slike.....	14
3.2	Sustavi boja	15
3.3	Gradijent slike	16
3.4	Primjena histograma	17
4	Detekcija i klasifikacija prometnih znakova	19
4.1	Cilj	19
4.2	Izvedba umjetnih neuronskih mreža.....	20
4.3	Detekcija znaka	21
4.3.1	Odabir ulaznih značajki	21
4.3.2	Rezultati.....	23
4.4	Klasifikacija znakova	24
4.4.1	Odabir ulaznih značajki	24
4.4.2	Rezultati.....	25
4.5	Integracija u postojeći sustav.....	27
4.6	Moguća poboljšanja.....	28
5	Zaključak	30
6	Literatura	31
7	Naslov, sažetak i ključne riječi.....	32
8	Title, abstract and keywords.....	33

1 Uvod

Današnja računala imaju veliku moć obrade podataka. Sve računske probleme, poput množenja matrica, računalo rješava neusporedivo bolje, i brže od čovjeka. Ipak, upotreba računala se pokazala lošom u rješavanju problema koje čovjek svakodnevno lakoćom rješava. Primjeri takvih problema su razumijevanje jezika, prepoznavanje objekata, sposobnost logičkog zaključivanja itd. Umjetna inteligencija je grana računarke znanosti koja proučava probleme vezane za inteligenciju, dok je računalni vid grana umjetne inteligencije koja se bavi problemima vezanim za sliku i video snimke. Uz sve dosadašnje napretke na području računalnog vida, ne postoji generalno rješenje za problem detekcije objekata u slici, te se svakom podproblemu, poput detekcije znaka ili lica, pristupa pojedinačno, uzimajući u obzir svojstva objekta.

U ovome radu rješava se problem detekcije i raspoznavanja znakova u slici. Osnovni cilj je za dani dio slike ocijeniti da li se radi o prometnom znaku, te o kojem se prometnom znaku radi, korištenjem umjetnih neuronskih mreža. Naposljetku je takvo rješenje integrirano u sustav za detekciju znakova u video snimci, temeljen na algoritmu Viole i Jonesa. Zbog jednostavnosti, oba sustava su ograničena na znakove upozorenja, to jest znakove oblika trokuta, jer raspoznavanje svakih od pojedinačnih tipova znakova predstavlja zasebni problem.

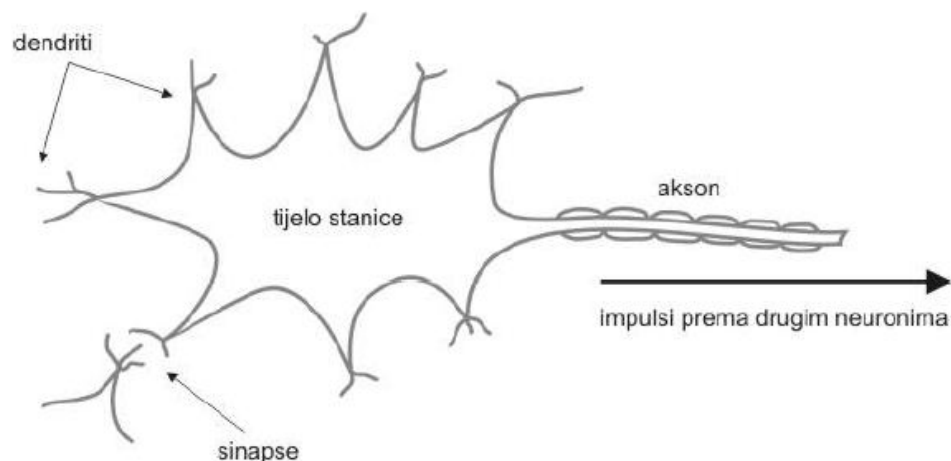
U nastavku je pojašnjen princip rada umjetnih neuronskih mreža, predložena je metoda za ispravljanje smetnje na slici uzrokovane preplitanjem te su razrađeni načini izlučivanja značajki iz slike. Naposljetku je predstavljena konkretna izvedba, usporedba rezultata u ovisnosti o odabiru značajki, te konačni rezultati.

2 Umjetne neuronske mreže

2.1 Motivacija

Probleme poput detekcije i klasifikacije slika, rijetko je moguće rješavati pomoću tradicionalnih algoritama. Pokazalo se kako je računalo vrlo loše u rješavanju mnogih problema u čijoj je srži sposobnost generalizacije. Za razliku od računala, čovjek takve probleme rješava brzo i efikasno. Upravo je ljudski mozak inspiracija za razvoj umjetnih neuronskih mreža.

Ljudski mozak sastoji se od međusobno povezanih neurona. Osnovni princip rada mozga je prolazak električnih impulsa kroz veze među neuronima.



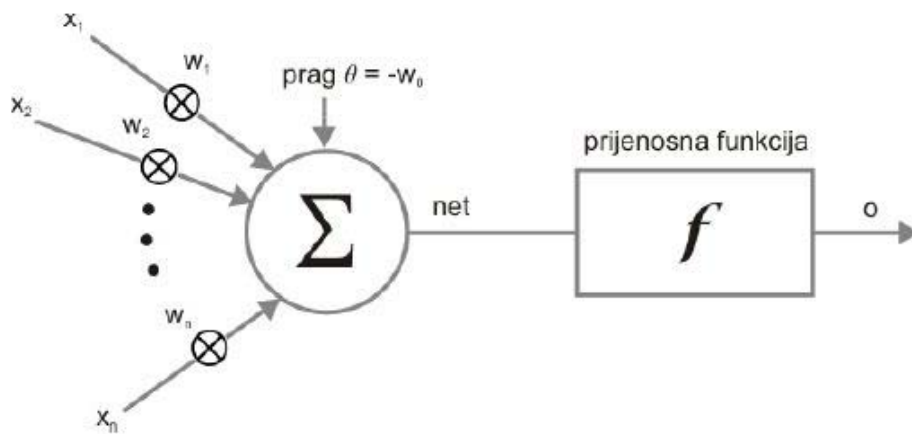
Slika 1. Pojednostavljeni model neurona [1]

Slika 1 prikazuje osnovne dijelove pojednostavljenog modela neurona. Električni impulsi iz susjednih neurona preko sinapsa i dendrita dolaze u tijelo stanice, te čine pobudu neurona. Odziv na pobudu se šalje preko aksona prema drugim neuronima.

Umjetne neuronske mreže [1] su matematički modeli zasnovani na pojednostavljenom modelu ljudskog mozga.

2.2 Umjetni neuron

1943. godine McCulloch i Pitts predložili su matematički model umjetnog neurona [1], prikazan na slici 2.



Slika 2. Umjetni neuron [1]

Dijelovi umjetnog neurona su:

- $x_{1..n}$ - ulazni signali
- $w_{0..n}$ - težinski faktori
- zbrajalo
- f - prijenosna funkcija

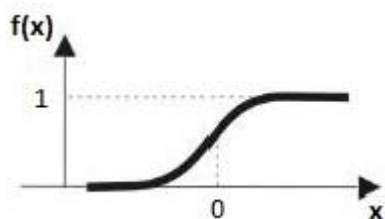
Analogija između umjetnog neurona i ljudskog neurona je sljedeća: x_i su ulazni signali iz drugih neurona, w_i su težine koje predstavljaju jakost sinapse, zbrajalo je tijelo stanice, a prijenosna funkcija predstavlja akson.

Vrijednost izlaza neurona određena je formulom 1:

$$o = f\left(\sum_{i=0}^n x_i w_i\right) \quad (1)$$

gdje je $x_0 = 1$.

Odabir prijenosne funkcije je proizvoljan, te ovisi o problemu. Pogodna prijenosna funkcija za klasifikaciju je sigmoidalna funkcija (2).



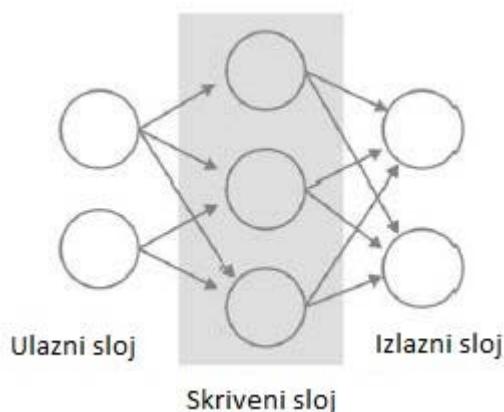
$$f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

Slika 3. Sigmoidalna funkcija [1]

Bitno svojstvo sigmoidalne funkcije je nelinearnost, koja omogućava razdvajanje linearno nerazdvojivih klasa pomoću neuronske mreže. U okviru ovog rada, neuronske mreže se koriste samo za klasifikaciju, te se u nastavku podrazumijeva da je prijenosna funkcija neurona funkcija (2).

2.3 Arhitektura unaprijedne neuronske mreže

Neuronska mreža nastaje povezivanjem neurona. Unaprijedna neuronska mreža (eng. *feedforward network*) je osnovna arhitektura mreže (Slika 4).



Slika 4. Arhitektura unaprijedne neuronske mreže [1]

Ulazni sloj predstavlja ulazne podatke, dok su izlazi neurona izlaznog sloja rezultati. Skriveni sloj se sastoji od jednog ili više podslojeva neurona. Veze postoje samo između susjednih slojeva i podslojeva: izlazi prethodnog sloja su ulazi sljedećeg, promatrajući od strane ulaznog sloja.

2.4 Učenje

Svaki od izlaza neuronske mreže je funkcija ulaznih podataka. Općenito je cilj konstruirati traženu funkciju. To se postiže postavljanjem težinskih faktora na adekvatne vrijednosti. Ručno određivanje težinskih faktora je teško već i za vrlo nizak broj neurona, a u konkretnim problemima je nemoguće. Određivanje težinskih faktora postiže se učenjem.

Jedan primjerak za učenje sastoji se od ulaznih podataka i traženog izlaza. Skup primjera za učenje je diskretni uzorak tražene funkcije, te je cilj učenja što bolje rekonstruirati funkciju iz njenih diskretnih uzoraka. Osnovna ideja je minimizirati funkciju pogreške u ovisnosti o težinskim faktorima. Odabrana funkcija pogreške E je polovina srednje kvadratne pogreške:

$$E(\vec{w}) = \frac{1}{2N} \cdot \sum_{i=1}^N (t_i - o_i)^2 \quad (3)$$

gdje su:

$\vec{w}$ - težinski faktori mreže

N - broj primjera za učenje

t_i - ciljana vrijednost i -tog primjerka

o_i - vrijednost izlaza mreže za i -ti primjerak.

Osnovna ideja je odrediti metodu traženja minimuma pogreške za jedan neuron, te pomoću propagacije pogreške unazad proširiti traženje minimuma na cijelu mrežu.

2.4.1 Gradijentni spust

Promatramo funkciju pogreške izlaza zbrajala, zanemarujući trenutno aktivacijsku funkciju neurona. Odabrana metoda za traženje minimuma funkcije pogreške jednog neurona je gradijentni spust [1]. Gradijent je smjer najbržeg rasta funkcije, te se težine mijenjaju u suprotnom smjeru:

$$\vec{w}(k+1) = \vec{w}(k) - \eta \cdot \nabla_{\vec{w}} E(\vec{w}) \quad (4)$$

gdje je η stopa učenja, $\vec{w}$ faktori težine neurona, a k korak učenja.

Gradijent funkcije pogreške iznosi:

$$\nabla_w E(\vec{w}) = \frac{dE(\vec{w})}{d\vec{w}} = \frac{d}{d\vec{w}} \left(\frac{1}{2N} \sum_{i=1}^N \varepsilon_i^2 \right) = \frac{1}{N} \sum_{i=1}^N \varepsilon_i \frac{d}{d\vec{w}} \varepsilon_i \quad (5)$$

gdje je N broj primjeraka za učenje, a ε_i iznos pogreške na i -tom primjeru.

Primjetimo da izračun gradijenta pomoću izraza (5) omogućuje samo grupno učenje. U praksi se pokazalo boljim pojedinačno učenje. Da bi se omogućilo pojedinačno učenje, funkcija pogreške se aproksimira koristeći samo jedan primjerak za učenje:

$$E(\vec{w}) \approx \frac{1}{2} \varepsilon^2 \quad (6)$$

Tada je aproksimacija gradijenta funkcije pogreške:

$$\nabla_w E(\vec{w}) \approx \varepsilon \frac{d}{d\vec{w}} \varepsilon \quad (7)$$

Ograničimo se trenutno na izlaz zbrajala neurona, za kojeg vrijedi:

$$o_{net} = \vec{w} \cdot \vec{x} \quad (8)$$

Tada vrijedi:

$$\varepsilon = t_{net} - o_{net} = t_{net} - \vec{w} \cdot \vec{x} \quad (9)$$

Iz formula (9) i (7) zaključujemo:

$$\nabla_w E(\vec{w}) \approx -\varepsilon \cdot \vec{x} \quad (10)$$

Promotrimo sada derivaciju sigmoidalne aktivacijske funkcije (2) :

$$\frac{d}{dx} f(x) = \frac{d}{dx} \left(\frac{1}{1 + e^{-x}} \right) = f(x)(1 - f(x)) \quad (11)$$

Veza između greške na izlazu zbrajala, i greške na izlazu prijenosne funkcije je:

$$\varepsilon = (t - o) \cdot f'(o_{net}) = o(1 - o)(t - o) \quad (12)$$

Iz (12), i (10) zaključujemo:

$$\nabla_w E(\vec{w}) \approx -o \cdot (1 - o) \cdot (t - o) \cdot \vec{x} \quad (13)$$

Iz (13) i (4) slijedi konačan izraz za traženje minimuma funkcije pogreške jednog neurona pojedinačnim učenjem gradijentnim spustom:

$$\vec{w}(k+1) = \vec{w}(k) + \eta \cdot o \cdot (1 - o) \cdot (t - o) \cdot \vec{x} \quad (14)$$

2.4.2 Propagacija greške unatrag

Pomoću izraza (14) moguće je gradijentnim spustom ugađati težine izlaznih neurona, ali da bi se pravilo primjenilo na neurone skrivenog sloja, potrebno je definirati pogrešku na izlazu neurona ($t - o$). Upravo to radi propagacija greške unatrag [1] (*eng. backpropagation*).

Veza greške δ na izlazu skrivenog neurona i greške ε (10) na izlazu zbrajala izlaznog neurona je:

$$\delta = w \cdot \varepsilon \quad (15)$$

gdje je w težinski faktor između dva neurona,

Greška na izlazu skrivenog neurona δ_h s obzirom na sve izlazne neurone, je:

$$\delta_h = t_h - o_h = \sum_{i=1}^K w_{hi} \varepsilon_i \quad (16)$$

gdje je K broj neurona izlaznog sloja, w_{hi} težinski faktor između h -tog neurona skrivenog sloja i i -tog neurona izlaznog sloja, a ε_i greška na izlazu iz zbrajala i -tog neurona izlaznog sloja.

Iz (16) i (14) slijedi pravilo ugađanja skrivenog neurona gradijentnim spustom:

$$\vec{w}_h(k+1) = \vec{w}_h(k) + \eta \cdot o_h \cdot (1 - o_h) \cdot \delta_h \cdot \vec{x}_h \quad (17)$$

Pomoću gradijentnog spusta i propagacije greške unazad, moguće je učiti kompletnu mrežu. Slijedi algoritam za učenje:

```
Inicijaliziraj težinske faktore slučajne vrijednosti
Dok nije ispunjen uvjet zaustavljanja radi:
    Za svaki primjer za učenje radi:
        Izračunaj izlaz mreže o
        Ugodi faktore izlaznog sloja po (14)
        Ugodi faktore skrivenog sloja po (17)
    Kraj
Kraj
```

2.4.3 Ubrzanje konvergencije

2.4.3.1 Adaptacija stope učenja

Modifikacijom algoritma učenja često se može postići brža konvergencija. Jedna od modifikacija je adaptacija stope učenja [2].

Nakon jedne epohe učenja, to jest nakon jednog učenja svakog od primjera, stopa učenja se adaptira u odnosu na promjenu funkcije greške: ako se greška smanjila, stopa se povećava, ako se greška povećala, promjene težina se zanemaruju i stopa se smanjuje:

$$\begin{aligned} \text{ako : } E(k) - E(k-1) < 0 &\rightarrow 0 < q < 1 \\ \text{inace, } q &> 1 \end{aligned} \quad (18)$$

Adaptacija stope učenja:

$$\eta = \eta \cdot q \quad (19)$$

Odabir faktora povećanja i faktora smanjenja stope učenja, uvelike utječu na performansu. Ako su promjene premale, neće imati efekta, ako su prevelike, može doći i do usporenja konvergencije zbog prevelikog broja odbačenih epoha.

2.4.3.2 Moment učenja

Negativna strana metode gradijentnog spusta je rizik pronalaženja lokalnih minimuma. Dodavanjem momenta učenju [2] postiže se mogućnost bijega iz malih lokalnih minimuma, kao i nastavljanje učenja na ravnim plohama funkcije pogreške.

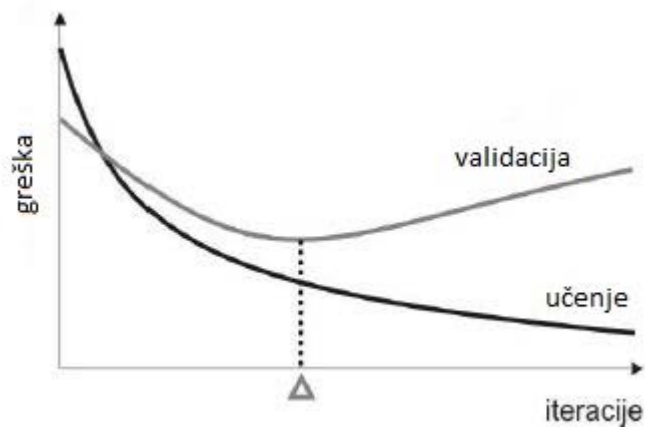
$$\vec{w}(k+1) = \Delta \vec{w} + m \cdot (\vec{w}(k) - \vec{w}(k-1)) \quad (20)$$

Izraz (20) predstavlja učenje sa momentom. Faktor m je faktor momenta i predstavlja udio prošle promjene težina u trenutnoj promjeni. Član $\Delta \vec{w}$ je promjena težina gradijentnim spustom.

2.4.4 Validacija

Neuronska mreža pretjeranim učenjem ulaznog skupa može naučiti šumove skupa za učenje, i izgubiti svojsvo generalizacije. Tako naučena neuronska mreža će dati loše rezultate u primjeni.

Sprječavanje pretjeranog učenja postiže se uvođenjem skupa za validaciju [1]. Mreža uči na primjerima skupa za učenje, a kao najbolja mreža se uzima mreža sa najboljim performansama na skupu za validaciju. Uz oprezan odabir skupova za učenje i validaciju, ovim pristupom se postiže očuvanje svojstva generalizacije mreže.



Slika 5. Korištenje validacije [1]

Važno je napomenuti da rezultati nad skupom za validaciju nisu rezultati testiranja. Iako mreža nije učila na primjerima za validaciju, samim odabirom mreže po performansama validacije, mreža je pod utjecajem tog skupa. Za testiranje je potrebno odvojiti još jedan skup primjeraka.

Osim što se po validaciji odabire najbolja mreža, po njoj se također i odabiru ulazne značajke, što se manifestira u još većem utjecaju validacije na mrežu.

3 Značajke slike

Da bi klasifikacija slika neuronskom mrežom bila efikasna, potrebno je odabrati značajke slike koje dobro opisuju i razdvajaju klase. Odabrane značajke se zatim postavljaju na ulaz neuronske mreže. Značajke također moraju omogućiti i dobru generalizaciju, jer često se događa da mreža nauči skup za učenje, dok se na skupu za validaciju postižu loši rezultati.

3.1 Obrada slike

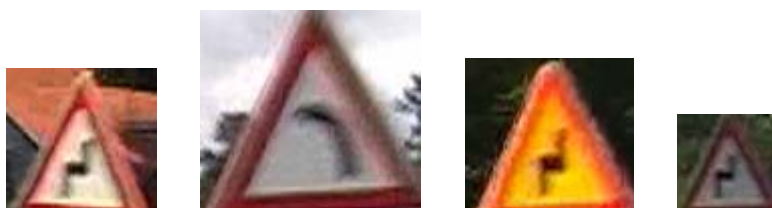
3.1.1 Ispravljanje smetnji uzrokovanih preplitanjem slike

Slike nad kojima se vrši klasifikacija su dijelovi video snimke. Kamera pri snimanju koristi preplitanje (*eng. interlacing*) : prvo snimi svaki drugi redak slike, a zatim preostale retke. Kao rezultat toga, javljaju se smetnje na slici pri brzim pokretima kamere. Na slici 6 su primjeri slika sa smetnjom pri utjecaju preplitanja.



Slika 6. Preplitanje slike

U skupu slika postoji okvirno 10% slika sa smetnjom prouzročenom preplitanjem, te bi adekvatna obrada takvih slika poboljšala rezultate klasifikacije. Jednostavna metoda sa dobrim rezultatima je odbacivanje polovice linija, te interpolacija linija koje nedostaju.



Slika 7. Slike ispravljene odbacivanjem linija

Slika 7 prikazuje slike sa slike 6 ispravljene metodom odbacivanja linija. Slike su vidno bolje nakon obrade, ali ipak nisu kvalitetne.

Problem u odbacivanju linija je gubitak informacije o slici, te kada bi se provodila obrada slika u tome obliku, informacija bi se izgubila i na kvalitetnim slikama.

Očuvanje informacije je motivacija za sljedeći pristup ispravljanju smetnji uzrokovanih preplitanjem. Promatrana smetnja se očituje u pomaknutim horizontalnim linijama, te bi aproksimacija inverza smetnje bio pomak tih linija natrag. Problem je otkrivanje najboljeg pomaka. U frekvencijskom spektru ispravne slike u pravilu se ističu niske frekvencije, dok su visoke relativno male. Promotrimo li sliku sa smetnjom uzrokovanom preplitanjem, zamijetiti ćemo da su istaknutije vertikalne frekvencije slike. Takva priroda smetnje omogućuje odabir udjela visokih vertikalnih frekvencija spektra slika za mjeru kvalitete.

Dvodimenzionalna diskretna kosinusna transformacija (21) je oblik diskretne Fourierove transformacije:

$$X_{k_1, k_2} = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x_{n_1, n_2} \cos \left[\frac{\pi}{N_1} \left(n_1 + \frac{1}{2} \right) k_1 \right] \cos \left[\frac{\pi}{N_2} \left(n_2 + \frac{1}{2} \right) k_2 \right] \quad (21)$$

Pojašnjenje:

X_{k_1, k_2} - vrijednost spektra za prostorne frekvencije k_1 i k_2

N_1, N_2 - dimenzije slike (prostorna domena)

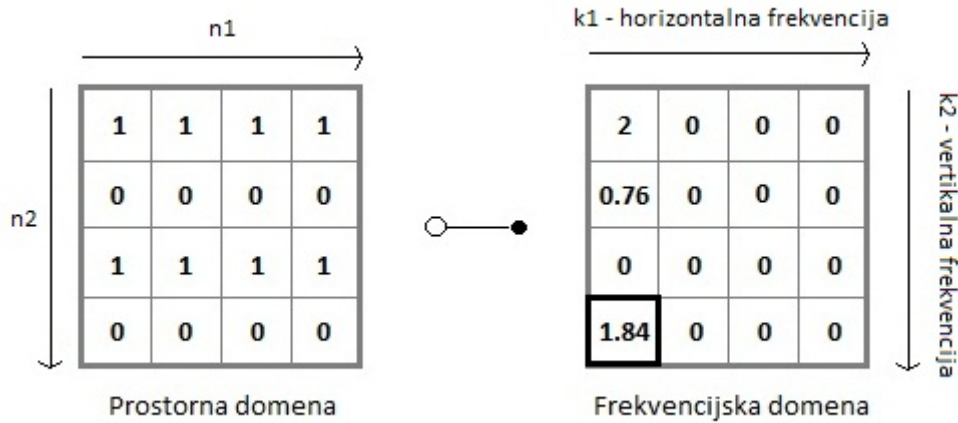
x_{n_1, n_2} - vrijednost slike na poziciji n_1, n_2

Dvodimenzionalna diskretna kosinusna transformacija (2D DCT) je kombinacija dvaju jednodimenzionalnih diskretnih kosinusnih transformacija (22).

$$X_k = \sum_{n=0}^{N-1} x_n \cos \left[\frac{\pi}{n} \left(n + \frac{1}{2} \right) k \right], k = 0..N-1 \quad (22)$$

Provodi se 1D DCT nad vertikalnim linijama slike, te se zatim provodi 1D DCT nad horizontalnim linijama vertikalnog spektra.

Osnovna svojstva 2D DCT transformacije povoljna su za izdvajanje vertikalnih frekvencija slike. Veličina spektra jednaka je veličini slike, te je spektar razapet dimenzijama horizontalne i vertikalne frekvencije. Tako se u 0-tom stupcu i n-tom retku nalazi faktor vertikalne prostorne frekvencije slike.



Slika 8. Dvodimenzionalna diskretna kosinusna transformacija

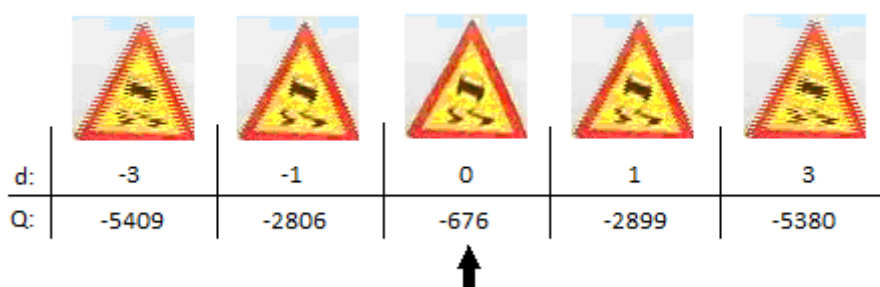
Slika 8 prikazuje jednostavan primjer svojstva izdvajanja vertikalne frekvencije. Uzorak u prostornoj domeni je izmjenjena crno-bijelih horizontalnih linija, dok u horizontalnom smjeru ne postoje nikakve oscilacije. Zbog toga su vrijednosti horizontalnih frekvencija u spektru 0, osim za nultu frekvenciju, koja predstavlja istosmjernu komponentu.

Odabir mjere kvalitete slike Q svodi se na odabir funkcije visokog vertikalnog spektra slike. Odabrana mjera je negativna suma 0.25% spektra, za frekvencije najbliže najvećoj vertikalnoj frekvenciji:

$$Q = - \sum_{k_1=0}^H \sum_{k_2=V}^{N_2-1} X_{k_1, k_2}, \quad \begin{matrix} H = \lceil 0.05 \cdot N_1 \rceil \\ V = \lceil 0.95 \cdot N_2 \rceil \end{matrix} \quad (23)$$

Nakon što je odabrana mjera kvalitete slike, možemo pristupiti procesu ispravljanja slike pomakom horizontalnih linija.

Osnovna ideja je pomakom parnih (ili neparnih) horizontalnih linija za različite iznose, odrediti pomak za koji je slika imala najveću mjeru kvalitete Q . Diskretna kosinusna transformacija se računa nad luminentom komponentom slike, koja se razmatra u nastavku rada. Promotrimo ispravljanje slike na kojoj nema pojave preplitanja.



Slika 9. Ispravljanje preplitanja pomakom

Slika 9. prikazuje primjer ispravljanja preplitanja metodom pomaka. Veličina d je pri tome horizontalni pomak linija, dok je Q mjera kvalitete predložena izrazom (23). Očito je da je početna slika, s pomakom 0, najkvalitetnija od ponuđenih, a za tu sliku je i mjera kvalitete Q najveća. To odražava bitno svojstvo ovog pristupa ispravljanju smetnje: očuvanje informacije i kvalitete slika bez smetnje uzrokovane preplitanjem. Pokazalo se da se ovom metodom ispravne slike ne mijenjaju.



Slika 10. Usporedba metoda ispravljanja preplitanja

Na slici 10 nalazi se usporedba dvaju metoda popravljjanja. Metoda ispravljanja pomakom popravlja sliku, ali se ipak pokazuje lošijom od metode ispravljanja. Uzrok tome je priroda smetnje. Pri preplitanju, podaci na zakašnjelim linijama nisu samo horizontalno pomaknuti od željenih podataka, nego su pomaknuti ovisno o smjeru gibanja kamere.

Ispravljanje pomakom može poslužiti za detekciju preplitanja: ako je najbolji pomak različit od 0, slika sadrži smetnje uzrokovane preplitanjem.

Kombinacija dvaju metoda daje najbolje rezultate. Ispravljanje pomakom služi kao detektor preplitanja. Ako je detektirana smetnja, slika se ispravlja metodom odbacivanja linija.

3.1.2 Filtriranje slike

Za neke primjene pogodno je sliku obraditi filtriranjem. Filtriranje se provodi pomoću diskretne konvolucije (24) pojedinačno po x i z dimenziji slike:

$$(f * g)(n) = \sum_{k=-\infty}^{\infty} f(k) \cdot g(n-k) \quad (24)$$

gdje je f signal koji se filtrira, a g je filtar. Sumu je potrebno računati samo za onoliko članova kolika je veličina filtra, po elementu rezultata.

Filtri se često koriste za izgladivanje slike, ili za naglašavanje određenih značajki. Slika 11 prikazuje primjer filtriranja slike sa filtrom $g = [0.2 \quad 1 \quad 0.2]$.



Slika 11. Primjer filtriranja slike.

Osim filtriranja slike, ponekad je korisno normalizirati sliku. Normalizacijom se maksimalna i minimalna vrijednost postavljaju na unaprijed određene vrijednosti, a ostale vrijednosti se skaliraju.

3.2 Sustavi boja

Boje su važne značajke slike, te je potrebno dobro odabrati reprezentaciju boja pri rješavanju problema klasifikacije.

Osnovni zapis slike je u RGB(eng. *red, green, blue*) formatu. Slika je razložena na tri komponente: crvenu, plavu i zelenu, za svaki slikovni element. Intuitivno je da su podaci o tri konkretne boje loše značajke za predstavljanje klase.

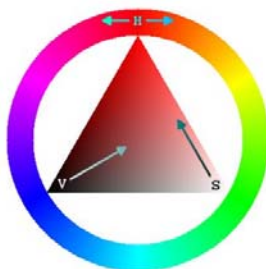
Korisna informacija o slici je luminantna komponenta, te odgovara komponenti Y iz YUV sustava boja. U i V su krominantne komponente. Formula (25) je izraz za izračun luminante komponente iz pomoću komponenti RGB sustava. Slika 12 prikazuje odnos slike i njene luminantne komponente.

$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \quad (25)$$



Slika 12. Luminantna komponenta slike

HSV(eng. *hue, saturation, value*) sustav boja sastoji se od komponenti nijanse (eng. *hue*), zasićenja (eng. *saturation*) i vrijednosti (eng. *value*). Uzimanjem nijanse kao značajke, stiče se otpornost na promjenu osvjetljenja.



Slika 13. HSV sustav boja [3]

3.3 Gradijent slike

Rubovi slike sadrže bitnu informaciju o obliku. Gradijent luminantne komponente relativno dobro izdvaja rubove slike, pošto je gradijent smjer najbržeg rasta funkcije.

Gradijent funkcije dvije varijable:

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} \vec{i} + \frac{\partial f(x, y)}{\partial y} \vec{j} \quad (26)$$

Pošto je slika diskretna, koristimo aproksimacije derivacija po x i y:

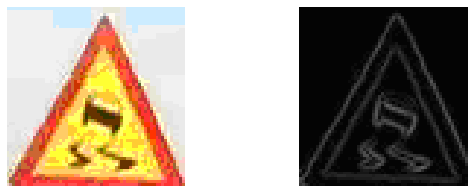
$$dx = \frac{\partial f(x, y)}{\partial x} = \frac{f(x+1, y) - f(x-1, y)}{2} \quad (27)$$

$$dy = \frac{\partial f(x, y)}{\partial y} = \frac{f(x, y+1) - f(x, y-1)}{2} \quad (28)$$

Pomoću izraza (27) i (28) gradijent se aproksimira na temelju vrijednosti horizontalnih i vertikalnih susjeda slikovnog elementa. Prikaz gradijenta pomoću kuta i iznosa često je prikladniji:

$$r = \sqrt{dx^2 + dy^2}, \quad \phi = \arctg\left(\frac{dy}{dx}\right) \quad (29)$$

Slika 14 prikazuje iznos gradijenta luminantne komponente slike. Može se zamijetiti sposobnost izdvajanja rubova.

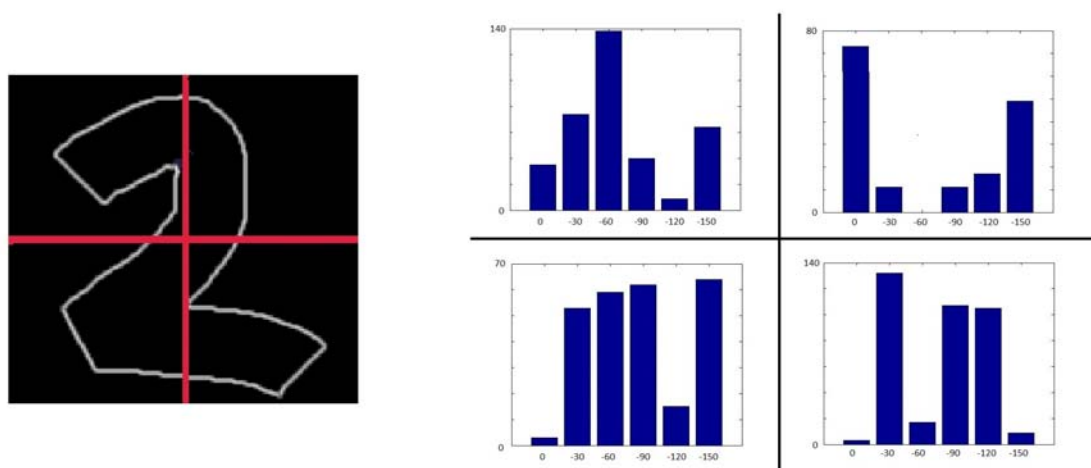


Slika 14. Gradijent luminantne komponente slike

3.4 Primjena histograma

Upotreba histograma je jednostavan način izlučivanja značajki nad dijelovima slike. Adekvatnim odabirom veličine i položaja segmenata slika nad kojima se rade histogrami značajki, postiže se određena otpornost na pomake slike.

Posebno zanimljiv je histogram orijentacije gradijenata [4]. Nakon odabira segmenata slike nad kojim se rade histogrami, za svaki slikovni element u segmentu promatra se kut gradijenta. Histogram kutova gradijenata daje opis orijentacije rubova za dotični segmenti slike, te time daje informaciju o obliku koji se nalazi u segmentu.

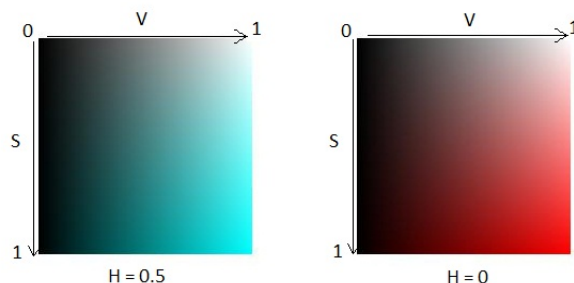


Slika 15. Histogram orijentacije gradijenata

Slika 15 prikazuje gradijent slike podijeljen na segmente, te histograme smjera gradijenata. Vrijednosti kutova linearno su raspoređene od 0 do -150 stupnjeva. Histogrami vjerno prikazuju mogućnost izdvajanja informacije o smjeru rubova slike.

U navedenom primjeru histogram se temelji na smjeru, to jest, na rasponu od 180 stupnjeva. Detaljniju informaciju o rubovima može se dobiti proširanjem histograma na orijentaciju, to jest na 360 stupnjeva, te proširenjem na dvodimenzionalni histogram dodavajući dimenziju iznosa gradijenta.

Histogram nijanse boja daje informaciju o zastupljenosti boja na dijelu slike. Promotrimo prvo karakteristiku HSV sustava prikazanu na slici 16.

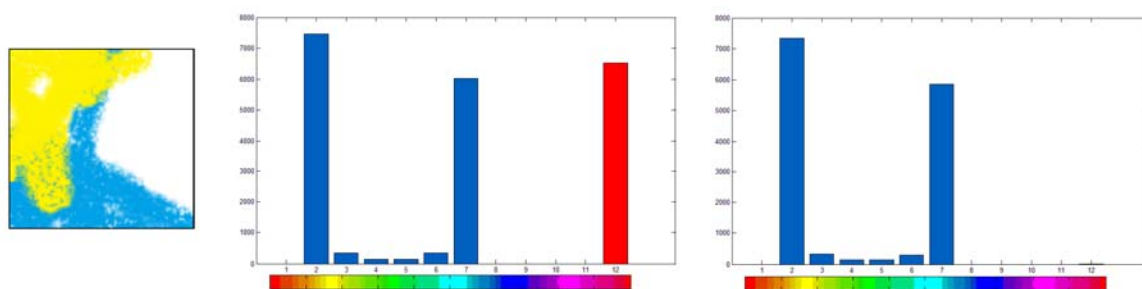


Slika 16. HSV sustav boja

Prikazane su vrijednosti slikovnih elemenata u ovisnosti o vrijednosti (V komponenta) i zasićenju (S komponenta). Niski iznosi V i S komponenti, zastupaju crnu, odnosno bijelu boju, bez obzira na nijansu, to jest na iznos H komponente. Nijansa je nedefinirana za crnu i bijelu boju, te se za njih pri pretvorbi iz drugih sustava nijansa postavlja na 0, što odgovara crvenoj boji.

Zbog prikazane karakteristike HSV sustava, histogram H komponente ne daje vjeran opis boja slike. Uzimanjem u obzir samo slikovnih elemenata čije su H i S komponente veće od određene granice, eliminira se utjecaj bijele i crne boje.

Slika 17 prikazuje histograme boja za danu sliku. Prvi histogram je izrađen bez diskriminacije na temelju vrijednosti i zasićenja. Na tome histogramu pojavljuje se veliki udio crvene boje, utjecane zastupljenošću bijele boje u slici. Drugi histogram je izrađen nad elementima čije su V i S komponente veće od 0.1, te je time odbačen utjecaj bijele i crne boje.



Slika 17. Histogram boja

4 Detekcija i klasifikacija prometnih znakova

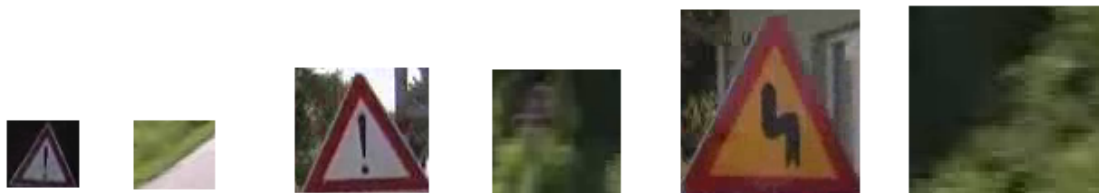
4.1 Cilj

Osnovni cilj sustava je detekcija znakova u video snimci, te klasifikacija detektiranih znakova. Slika 18 prikazuje umanjeni primjer isječka iz video snimke u kojem je potrebno izdvojiti i klasificirati znakove.



Slika 18. Isječak iz video snimke

Ovaj rad fokusira se na klasifikaciju dijelova slike. Cilj je pomoću jedne neuronske mreže odrediti da li se radi o znaku, te ako se radi o znaku, pomoću druge neuronske mreže odrediti o kojem se znaku radi. Slika 19 prikazuje primjere dijelova slika koje treba klasificirati. U okviru ovog rada, promatraju se samo znakovi upozorenja, to jest znakovi oblika trokuta.



Slika 19. Primjeri dijelova slike iz video snimke

4.2 Izvedba umjetnih neuronskih mreža

Sve korištene neuronske mreže su unaprijedne (poglavlje 2.3, slika 4). Sastoje se od ulaznog sloja, skrivenog sloja sa jednim podslojem, te izlaznog sloja. Prijenosna funkcija svih neurona skrivenog i izlaznog sloja je sigmoidalna funkcija (2) razmatrana u poglavlju 2.2.

Ulazni sloj predstavlja ulazne značajke, te je broj neurona ulaznog sloja jednak broju ulaznih značajki. Broj neurona u skrivenom sloju mora biti dovoljno velik da bi mreža mogla naučiti traženu funkciju, a ipak je potrebno da bude dovoljno nizak da bi brzina konvergencije bila prihvatljiva. U pravilu je u izvedenim neuronskim mrežama veličina skrivenog sloja približno jednaka trećini veličine ulaznog sloja neurona.

Broj izlaznih neurona u mrežama za detekciju znaka je 1. Zbog odabira sigmoidalne prijenosne funkcije, izlaz iz mreže je realan broj između 0 i 1. Ako je izlaz veći od određene granice k (primjerice $k = 0.5$), promatranu sliku proglašujemo znakom. Izlazne vrijednosti primjera za učenje postavljamo na 1, ako je primjerak znak, a na 0 inače.

Broj izlaznih neurona u mrežama za klasifikaciju znaka je jednak broju klasa. Odabrana klasa znaka je redni broj onog izlaznog neurona čija je izlazna vrijednost najveća. Izlazne vrijednosti primjera za učenje se postavljaju u skladu s time: izlazna vrijednost neurona koji predstavlja klasu primjerka postavlja se na 1, dok se izlazne vrijednosti ostalih neurona postavljaju na 0.

Zbog prirode sigmoidalne funkcije, očekivanja vrijednosti ulaznih neurona normiraju se oko 0, na temelju skupa za učenje. Nakon toga vrijednosti se skaliraju u ovisnosti o zadanom rasponu.

Početne vrijednosti težina neuronske mreže se postavljaju na slučajne brojeve između -1 i 1. Ovakvo postavljanje početnih vrijednosti težina i normiranje ulaza osiguravaju rad prijenosne funkcije (2) u prijelaznom području.

4.3 Detekcija znaka

4.3.1 Odabir ulaznih značajki

Za rješavanje problema detekcije znaka, poželjno je odabrati ulazne značajke po kojima se znak može lako razlikovati od okruženja. Prometni znakovi se ističu svojom bojom i oblikom, te su ulazne značajke odabrane u skladu s time.

4.3.1.1 Histogrami boje i zasićenja

Prvi način odabira ulaznih značajki temelji se na svojstvima znakova vezanim za boju. Slika se prije izlučivanja značajki skalira na veličinu 24x24.

Informaciju o boji slike dobro opisuje histogram boja obrađen u poglavlju 3.4. Pošto je u znakovima udio bijele boje također značajan, koristi se i histogramom S komponente (zasićenja) HSV sustava koji daje informaciju o udjelu bijele boje u slici. Oblik znaka se također uzima u obzir, te se histogrami izrađuju samo na dijelovima slike.



Slika 20. Odabir slikovnih elemenata

Slika 20 prikazuje područja nad kojima se izrađuju histogrami. Histogram boje se izrađuje nad slikovnim elementima unutar vanjskog ruba, dok se histogram zasićenja izrađuje na manjem području, na slikovnim elementima unutar unutrašnjeg ruba.

Veličina histograma boje i histograma zasićenja je 20. Time je veličina ulaznog sloja neuronske mreže 40, dok je veličina skrivenog sloja na 32.

4.3.1.2 Histogrami orijentacije gradijenata

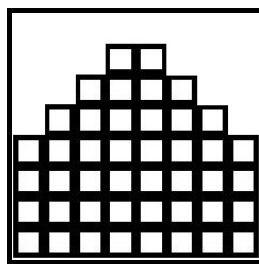
Drugi način izlučivanja značajki oslanja se na trokutasti oblik znakova. Histogrami orijentacije gradijenata (poglavlje 3.4) luminantne komponente dobro izdvajaju informaciju o rubovima slike, koji su istaknuti u slikama prometnih znakova.

Prije izlučivanja značajki, slika se skalira na veličinu 48x48, zatim se vrijednosti luminantne komponente normaliziraju na raspon 0 – 255. Nakon normalizacije, luminantna komponenta se filtrira Gausovim filtrom:

$$f = \frac{1}{15} \cdot \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (30)$$

Filtriranjem slike Gausovim filtrom postiže se ugađivanje rubova slike, te time smjer gradijenta postaje ravnomjerno raspoređen na rubovima, što je bitno pri izradi histograma orijentacije gradijenata.

Histogrami se izrađuju nad blokovima veličine 6x6, te su raspoređeni uzevši u obzir poziciju znaka u slici. Slika 21 prikazuje raspored blokova.



Slika 21. Raspored blokova

Histogrami se izrađuju za 6 kuteva te uzimaju u obzir samo smjer, to jest, 180 stupnjeva.

Veličina ulaznog sloja neuronske mreže je 264, dok je veličina skrivenog sloja 79.

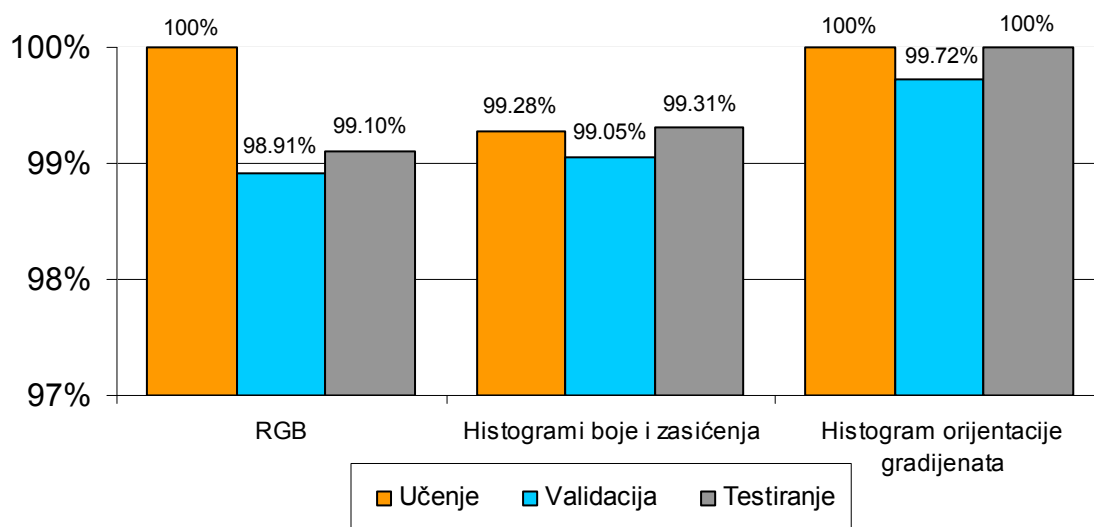
4.3.2 Rezultati

Skupovi za učenje, validaciju i testiranje sastoje se od dvije vrste slika, znakova, i okoline. Sve slike su izabrane iz jednog skupa slika metodom slučajnog odabira. Početni skup se sastoji od slika izdvojenih iz video snimaka. Tablica 1 prikazuje podatke o veličini skupova za učenje, validaciju i testiranje.

	Učenje	Validacija	Testiranje
Znakovi	1426	353	348
Ostalo	1223	384	339

Tablica 1. Veličine skupova za učenje, validaciju, i testiranje

Učenje je provedeno nad 3 načina odabira ulaznih značajki: RGB format, histogrami boje zasićenja, te histogrami orijentacije gradijenata.



Slika 22. Usporedba rezultata za različit odabir značajki

Slika 22. prikazuje rezultate na skupovima za učenje, validaciju i testiranje, u ovisnosti o odabiru ulaznih značajki. Odabir ulaznih značajki utječe na performansu, ali manje nego što je očekivano. Razlog tomu leži u velikom skupu za učenje, te neuronska mreža može naučiti i generalizirati, iako je odabir značajki loš. Najbolji rezultati se postižu na histogramima orijentacije gradijenata.

4.4 Klasifikacija znakova

4.4.1 Odabir ulaznih značajki

Za razliku od detekcije znaka, pri klasifikaciji znakova poželjno je odabrati ulazne značajke koje dobro opisuju razliku između različitih klasa znakova. Informacija o razlici između klasa se nalazi na području središta znaka.

4.4.1.1 Luminantna komponenta

Prvi način odabira značajki je luminantna komponenta središta znaka. Slike se skaliraju na veličinu 24x24, zatim se izdvajaju središnji elementi slike, na način prikazan na slici 23.



Slika 23. Izdvajanje središnjeg dijela slike

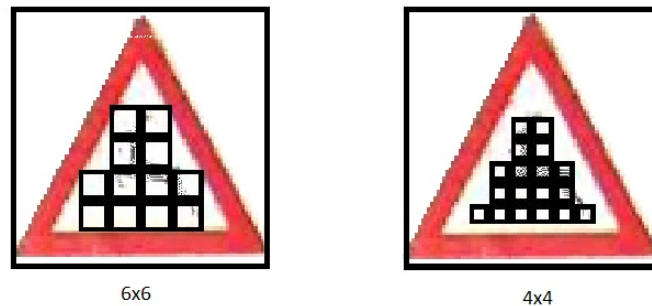
Dobivena luminantna komponenta dijela slike se zatim normalizira na raspon od 0 do 255.

Ukupan broj promatranih slikovnih elemenata je 112, te je toliki i broj neurona ulaznog sloja neuronske mreže. Broj neurona skrivenog sloja je 34.

4.4.1.2 Histogrami orijentacije gradijenata

Razlika između različitih klasa znakova je u oblicima koji se nalaze u središtu, te se histogram orijentacije gradijenata koji nosi informaciju o rubovima na središtu slike nameće kao logičan odabir ulaznih značajki.

Slike se skaliraju na veličinu 48x48, te se luminantna komponenta normalizira na raspon od 0 do 255. Nakon toga slijedi filtriranje Gausovim filtrom (30). Zatim se izrađuju histogrami orijentacije gradijenata nad blokovima smještenim u središtu slike. Postoje dva skupa blokova, jedni veličine 4x4, te drugi veličine 6x6. Slika 24 prikazuje raspored blokova.



Slika 24. Raspored blokova

Histogrami nad blokovima veličine 6x6 se rade za 4 kuta i 180 stupnjeva, dok se histogrami nad blokovima veličine 4x4 rade za 7 kuteva i 360 stupnjeva. Broj neurona ulaznog sloja neuronske mreže je 174, te je broj neurona u skrivenom sloju jednak 52.

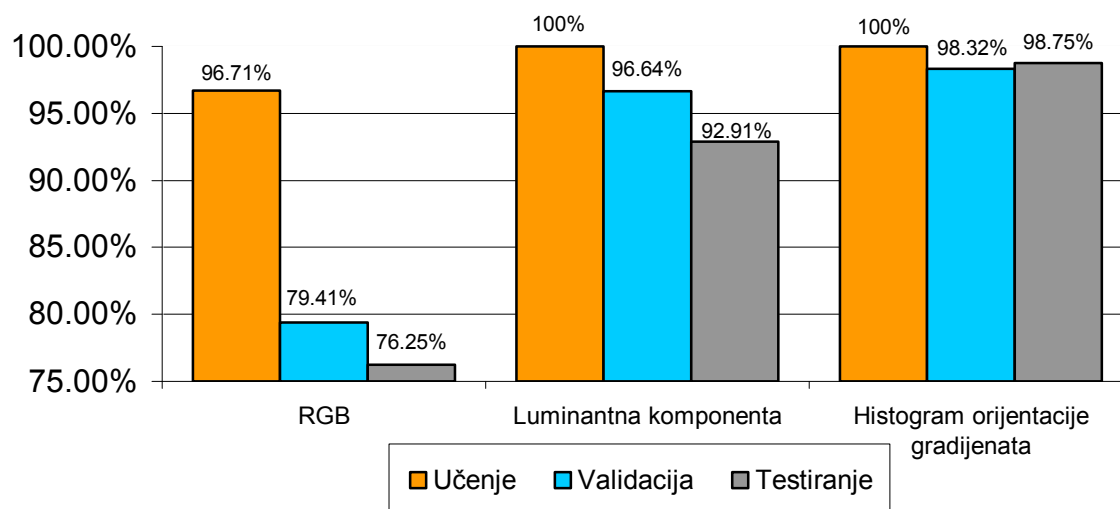
4.4.2 Rezultati

Skupovi za učenje, validaciju i testiranje odabrani su na isti način kao i skupovi za detekciju znakova. Odabrane su klase znakova čijih je primjeraka na raspolaganju više od 140. Takvih je klasa ukupno 8. Korištene klase prikazane su na slici 24. Skup za učenje sastoji se od 80 primjeraka po klasi, skup za validaciju od 30 primjeraka po klasi, te skup za testiranje od 30 primjeraka po klasi.



Slika 24. Korištene klase znakova.

Učenje je provedeno za tri načina odabira ulaznih značajki: RGB format, luminantna komponenta središta slike, te histogrami orijentacije komponenata opisani u prethodnom poglavlju.



Slika 25. Rezultati klasifikacije u ovisnosti o odabiru značajki

Slika 25 prikazuje rezultate učenja. Za razliku od rezultata kod detekcije, odabir ulaznih značajki značajno utječe na rezultate klasifikacije. Uzrok tome je relativno mali broj primjeraka za učenje po klasi. Histogram orijentacije gradijenata daje najbolje rezultate, od 98.75% ispravno klasificiranih primjeraka na skupu za testiranje.

4.5 Integracija u postojeći sustav

Video u kojem je potrebno detektirati i klasificirati znakove, je rezolucije 720x576. Ispitati sve dijelove slike različitih veličina bilo bi relativno sporo. Algoritam Viole i Jonesa [5] je algoritam za detekciju objekata u video snimci prilagođen za brzu obradu, te omogućava detekciju objekata u video snimci u realnom vremenu. Neuronske mreže obrađene u ovom radu su integrirane u postojeći sustav [6] temeljen na algoritmu Viole i Jonesa.

Osnovna ideja ovog završnog rada je isječke slike koji su proglašeni znakovima od strane algoritma Viole i Jonesa, dodatno provjeriti neuronskom mrežom za detekciju, te ako je isječak znak, odrediti njegovu klasu.

Slika 26 prikazuje primjer detekcije znakova u slici. Isječci označeni plavom bojom su detektirani algoritmom Viole i Jonesa, dok su okviri označeni crvenom bojom dodatno detektirani neuronskom mrežom.



Slika 26. Prepoznavanje znakova u video snimci

Ovakvom kombinacijom klasifikatora moguće je poboljšati rezultate detekcije. Ipak, rezultati klasifikacije među klasama znakova su vrlo loši. Razlog tomu je loša centriranost znakova u isječcima pronađenim od strane algoritma

Viole i Jonesa, jer znakovi na slikama iz skupova za učenje, validaciju i testiranje neuronskih mreža u ovom radu su veoma precizno centrirani. Slika 27 prikazuje primjere detektiranih isječaka od strane algoritma Viole i Jonesa



Slika 27. Detektirani isječki slike

4.6 Moguća poboljšanja

Važno je napomenuti karakteristiku skupa iz kojeg su izabrani skupovi za učenje, validaciju i testiranje neuronskih mreža u ovome radu. Slika 28 prikazuje niz slika prometnih znakova iz tog skupa. Svih pet slika prikazuje isti znak, ali u susjednim okvirima video snimke.



Slika 28. Prometni znakovi iz početnog skupa

Pošto su skupovi za učenje, validaciju i testiranje izabrani metodom slučajnog odabira iz početnog skupa, čest je slučaj da se različite slike istog znaka, poput onih na slici 28, pojave i u skupu za učenje, i u skupu za validaciju, i u skupu za testiranje. Iako se radi o različitim slikama, korelacija između njih je visoka. Zbog navedene karakteristike skupova, rezultati postignuti neuronskim mrežama nisu mjerodavni. Potrebno je provesti testiranje na odvojenom skupu slika. Skup za validaciju bi trebalo također odabrati tako da bude neovisan o skupu za učenje.

Algoritam Viole i Jonesa, čiji rezultati su prikazani na slici 26, treniran je za potrebe demonstracije u ovom završnom radu, pošto je prijašnja inačica bila

trenirana na manjem i lošijem skupu slika. Rezultati algoritma se mogu uvelike poboljšati.

Problem loše centriranosti okvira koji su detektirani algoritmom Viole i Jonesa, moguće je riješiti naknadnim centriranjem koristeći Houghovu transformaciju za detekciju trokuta. Također, moguće je proći susjedstvo i unutrašnjost okvira neuronskom mrežom, te pronaći okvir za koji je izlaz neuronske mreže najveći.

Praćenjem detektiranih okvira u video snimci, dobila bi se dodatna informacija o kontekstu, te bi se mogli uvelike poboljšati rezultati detekcije i klasifikacije na razini video snimke.

Za negativne primjerke znakova u skupu za učenje neuronske mreže za detekciju, dobro bi bilo odabrati krive detekcije algoritma Viole i Jonesa. Time bi neuronska mreža naučila ispravljati “mane” algoritma Viole i Jonesa.

5 Zaključak

Iako im je koncept poznat polovinu stoljeća, umjetne neuronske mreže pokazuju se moćnim alatom za primjenu na području računalnog vida. Primjena umjetne neuronske mreže na problemu detekcije i klasifikacije prometnih znakova polučila je relativno dobre rezultate, te time pružila osnovu za daljnje istraživanje.

Kroz prethodna poglavlja poseban naglasak je bio na odabiru ulaznih značajki slike. Pokazalo se da broj primjeraka za učenje više utječe na performansu, nego pažljivi odabir ulaznih značajki.

U dva različita problema, detekciji znaka, te klasifikaciji znakova, odabir histograma orijentacije gradijenata za ulazne značajke, u konačnici je dao najbolje rezultate, te se pokazao boljim od odabira ulaznih značajki specifičnih za problem. Time se histogram orijentacije gradijenata nameće kao potencijalni univerzalni odabir značajki za probleme klasifikacije na području računalnog vida.

6 Literatura

[1] Bojana Dalbelo Bašić, Marko Čupić, Jan Šnajder

“Umjetne neuronske mreže” , 08. lipnja 2009.

http://www.fer.hr/download/repository/UI_14_umjetne_neuronske_mreze.pdf

[2] Yann LeCun, Leon Bottou, Genevieve B. Orr, Klaus-Robert Müller

“Efficient BackProp“, 1998

<http://yann.lecun.com/exdb/publis/pdf/lecun-98b.pdf>

[3] “HSL and HSV”, 27. prosinca 2009.

http://en.wikipedia.org/wiki/HSL_and_HSV

[4] “Histogram of oriented gradients” , 27. prosinca 2009.

http://en.wikipedia.org/wiki/Histogram_of_oriented_gradients

[5] “Viola Jones object detection framework”, 20. siječnja 2010.

http://en.wikipedia.org/wiki/Viola-Jones_object_detection_framework

[6] Igor Bonači i ostali

“Detekcija prometnih znakova, Tehnička dokumentacija” , FER, 2009

7 Naslov, sažetak i ključne riječi

Naslov

Sustav za detekciju i raspoznavanje prometnih znakova

Sažetak

Osnovni cilj rada je detekcija i raspoznavanje znakova upozorenja u slici, korištenjem umjetnih neuronskih mreža. U radu je definiran koncept umjetnih neuronskih mreža, pojašnjen je njihov princip rada, te algoritam propagacije greške unazad kao način učenja neuronske mreže. Predložena je metoda za detekciju smetnje na slici uzrokovane preplitanjem, temeljena na prostornom frekvencijskom spektru i diskretnoj kosinusnoj transformaciji. Razrađeni su načini izlučivanja značajki slike pogodni za probleme detekcije i raspoznavanje prometnih znakova. Kao najbolji odabir ulaznih značajki pokazao se histogram orijentacije gradijenata. Naposljetku je predstavljena konkretna izvedba neuronskih mreža, usporedba rezultata u ovisnosti o odabiru značajki, te konačni rezultati. Takvo rješenje integrirano je u sustav za detekciju prometnih znakova u video snimci temeljen na algoritmu Viole i Jonesa.

Ključne riječi

Prometni znakovi, umjetne neuronske mreže, detekcija, raspoznavanje, preplitanje slike, histogram orijentacije gradijenata, algoritam Viole i Jonesa.

8 Title, abstract and keywords

Title

System for traffic signs detection and recognition

Abstract

The main goal of this thesis is the detection and classification of traffic signs in images, by using artificial neural networks. The concept of artificial neural network is defined, and backpropagation algorithm is observed. A method for the detection of interlacing artifacts in images, based on discrete cosine transform, is proposed. Methods for extracting features from images are discussed, taking into account the characteristics of the problem of detection and classification of traffic signs. For both problems, histogram of oriented gradients was the best choice. Finally, the neural networks were integrated into system for detection traffic signs in video, based on Viola-Jones algorithm.

Keywords

Traffic signs, artificial neural networks, detection, recognition, interlacing, histogram of oriented gradients, Viola – Jones algorithm.