

Dirbtinių neuronų tinklų taikymas automobilių registracijos numeriams atpažinti

R. Laptik, D. Navakauskas

*Elektroninių sistemų katedra, Vilniaus Gedimino technikos universitetas,
Naugarduko g. 41, LT-03227 Vilnius, Lietuva, tel. +370 5 2744756, el. p. raimond.laptik@el.vtu.lt*

Įvadas

Daugėjant mašinų, vis sunkiau kontroliuoti jų srautą. Gausėja automobilių vagysčių, eismo įvykių, kitokių kelių eismo taisyklių pažeidimų. Vienas svarbiausių uždavinių – sukurti ir įdiegti sistemą, galinčią automatiškai atpažinti ne tik atskirus automobilius, bet ir konkretaus automobilio registracijos numerį. Tokią sistemą galima būtų naudoti automatinei automobilių srauto kontrolei, pvz., šviesoforų darbo režimas galėtų būti keičiamas, atsižvelgiant į esamą automobilių srautą. Numerių atpažinimo sistema taip pat galėtų būti naudojama ne tik kaip savarankiškas įrenginys, bet ir kaip dalis bendros sistemos, skirtos vagysčių ir kelių eismo taisyklių pažeidimo prevencijai.

Šiuo metu pasaulyje veikia nemažai specializuotų automobilių numerių atpažinimo sistemų: „VisiCar“ [1], „Carmen“ [2], „Iris“ [3]. Šių sistemų simbolių atpažinimo algoritmai iš dalies yra panašūs į komercinių simbolių atpažinimo programų, tokių kaip „Fine Reader“, „CuneForm“, „Text Bridge“, arba teksto atpažinimo įrenginių, tokių kaip „Siemens Pocket Reader“, „C pen“, algoritmus. Automobilių numerių atpažinimo sistemų veikimo algoritmai iš esmės išsiskiria dideliu atpažinimo greičiu – daugelis šių sistemų veikia realiu laiko masteliu ir geba laiku priimti sprendimus. Šiame darbe dėmesį koncentruosime į automobilių registracinių numerių atpažinimo modulį, siekdami pasiūlyti naują simbolių atpažinimo algoritmą. Atpažinimo greičiui bei tikslumui

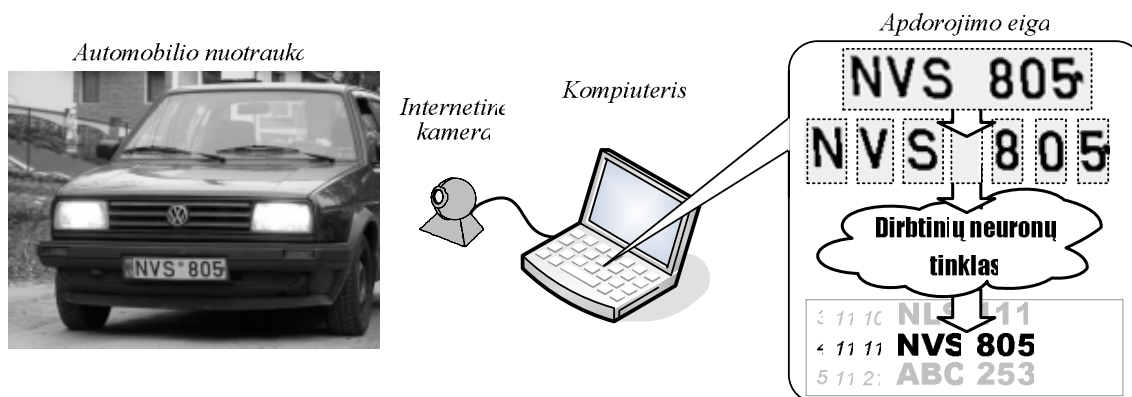
užtikrinti taikysime dirbtinių neuronų tinklų (toliau DNT) teorija pagrįstus modelius [4, 5]. Išskirtinis nagrinėjamų modelių bruožas, lyginant su anksčiau aptartomis sistemomis, yra galimybė dirbti su mažesnės raiškos nuotraukomis. Tai savo ruožtu sumažina reikalavimus vaizdo kamerei ir sudaro prielaidas atpiginti visą atpažinimo sistemą.

Kitoje šio straipsnio dalyje glaustai apžvelgsime kuriamos atpažinimo sistemos struktūrą bei jos veikimą. Toliau nagrinėsime parinktus DNT modelius – Kohoneno [6] ir daugiasluoksnius perceptronų [7] tinklus bei jų taikymo simboliams atpažinti būdus. Paskutinėje straipsnio dalyje pateiksime automobilių numerių atpažinimo, taikant siūlomus modelius, eksperimentų rezultatus. Šių eksperimentų tikslas – rasti tinkamus DNT struktūrų bei jų mokymo taisyklių parametrus.

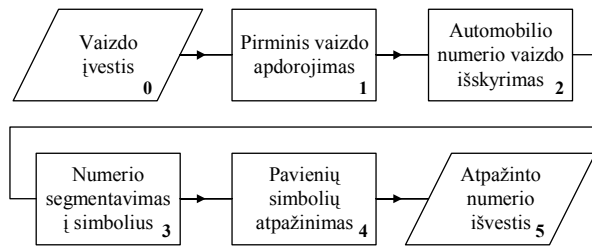
Kuriamos atpažinimo sistemos sudėtis

Siekiant pigumo aparatinis kuriamos sistemos įdiegimas yra gana paprastas (žr. 1 pav.). Šią sistemą sudaro: a) vaizdo įvesties įrenginys – internetinė kamera su 640×480 taškų raiška, b) asmeninis kompiuteris (jei reikia mobilumo, nešiojamasis kompiuteris).

Specializuota programinė įranga, skirta vaizdams apdoroti bei atpažinti, sudaro atpažinimo sistemos esmę. Detaliau aptarkime atpažinimo programinio modulio veikimą (žr. 2 pav.).



1 pav. Automobilių registracijos numerių atpažinimo sistemos įdiegimas



2 pav. Atpažinimo programinio modulio struktūra

Pirmasis žingsnis po vaizdo įvesties yra jo pirminis apdorojimas. Pirminio vaizdo apdorojimo metu spalvotas vaizdas paverčiamas pilku, taikant įprastą išraišką $p = (r + g + b)/3$; čia r , g , ir b yra raudonos, žalios ir mėlynos spalvų dedamosios; p – skaitis. Vėliau, taikant adaptyviąją slenksčio funkciją, pilkas vaizdas yra verčiamas juodai baltu, t . Slenksčio T vertė konkrečiam vaizdo taškui yra skaičiuojama analizuojant aplink jį 10×10 dydžio lange esančius taškus. Eksperimentuojant nustatyta, kad tinkamiausi toliau apdoroti rezultatai gaunami, kai slenksčio vertė randama iš išraiškos $T = (p_{\min} + p_{\max})/2$; čia p_{\min} ir p_{\max} yra ribinės skaisčio vertės nagrinėjamame lange.

Antruoju žingsniu, analizuojant vaizdą po pirminio apdorojimo, nustatoma automobilio registracijos numerio vieta. Laikoma, kad numeris yra tam tikro, standartų nustatyto, dydžio kontūras su iš anksto žinomo dydžio simboliais. Taikant rekurentinį užpildymo algoritmą, vaizde randami kontūrai bei jų ribinės koordinatės. Rastų ribinių koordinatėjų skirtumas leidžia įvertinti kontūro dydį. Jeigu rasto kontūro dydis yra panašus į automobilio registracijos numerio dydį, tai tikrinama, ar kontūre esantys objektai atitinka standartizuoto automobilio registracijos numerio simbolių dydį. Jei atitinka ir šie dydžiai, nusprendžiama, jog rastas automobilio numerio kontūras, todėl toliau apdoroti perduodamas tik šiam kontūre esantis vaizdas.

Trečiuoju žingsniu išskirtas automobilio registracijos numerio vaizdas segmentuojamas į atskirų simbolių vaizdus. Tam naudojami išankstiniai duomenys apie simbolių dydžius bei vaizdo skaisčio histogramos.

Ketvirtuoju apdorojimo žingsniu pavienių simbolių vaizdai yra atpažinami, o vėliau registruojami. Simbolių atpažinimas yra vienas iš sudėtingiausių žingsnių, todėl šio žingsnio aptarimui skirsime daugiausia dėmesio.

Taikomi dirbtinių neuronų tinklai

Kuriamoje sistemoje dirbtinių neuronų tinklas yra taikomas simboliams atpažinti. DNT veikia kaip klasifikatorius, priskirdamas jam pateiktą normalizuotą simbolio vaizdą vienai ar kitai klasei.

Tinklui simbolių vaizdai yra pateikiami 100 elementų \mathbf{a} vektoriais, gautais stulpeliais išskaidant normalizuoto vaizdo \mathbf{A} matricą. Tam tikslui tikrasis simbolio dydis yra keičiamas į 10×10 elementų \mathbf{A} matricą. Koeficientas k_H nustatomas pagal simbolio aukščio santykį, t. y. iš išraiškos $k_H = H_i / H_n$; čia H_n – norimas simbolio aukštis (šiuo atveju – 10), o H_i – turimas simbolio aukštis. Taigi

ieškomas santykinis juodų taškų, esančių $k_H \times k_H$ dydžio kvadrate, skaičius:

$$q_{ij} = \left(\sum_{I=I_0+i-k_H/2}^{I_0+i+k_H/2} \sum_{J=J_0+j-k_H/2}^{J_0+j+k_H/2} t_{IJ} \right) / k_H^2; \quad (1)$$

čia t_{IJ} – juodai balto vaizdo taško vertė; I_0 ir J_0 – simbolio pradžios visame vaizde koordinatės; I ir J – taško vaizde koordinatės; i ir j – matricos indeksai. Kiekvienas \mathbf{A} matricos elementas randamas taip:

$$a_{ij} = \begin{cases} 0, & \text{kai } q_{ij} < 0,5; \\ 1, & \text{kai } q_{ij} \geq 0,5. \end{cases} \quad (2)$$

Pereikime prie konkrečių dirbtinių neuronų tinklų taikymo simboliams atpažinti – aptarkime parinktų tinklų struktūras ir jų mokymą.

Kohoneno tinklas [6] parinktas dėl jo gerų apibendrinimo (kai tinklas geba teisingai klasifikuoti nematytus pavyzdžius) savybių. Tinklo struktūrą sudaro vienmatis įėjimo ir dvimatis išėjimo sluoksniai (žr. 3 pav.). Įėjimo vektorius \mathbf{a} yra pateikiamas per atitinkamus neuronų ryšių svorius į kiekvieną iš neuronų, esančių išėjimo sluoksnyje. Atsako formavimo metu neuronu laimėtoju ($s_L = 1$) tampa tas neuronas, kurio ryšių svorių vektorius \mathbf{w}_L yra arčiausiai pateikto įėjimo vektoriaus \mathbf{a} , t. y. $\|\mathbf{a} - \mathbf{w}_L\| = \min_i \|\mathbf{a} - \mathbf{w}_i\|$. Atstumas tarp vektorių apskaičiuojamas taikant Euklido metriką:

$$\|\mathbf{a} - \mathbf{w}\| = \sqrt{\sum_j (a_j - w_j)^2}. \quad (3)$$

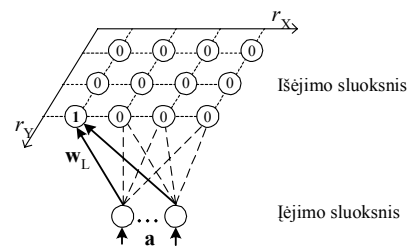
Kohoneno tinklas mokomas pagal taisyklę:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{L_i}(\mathbf{a}(t) - \mathbf{w}_i(t)); \quad (4)$$

čia $\mathbf{w}_i(t)$ yra i -tojo išėjimo neurono ryšių svorių vektorius t laiko momentu; $\mathbf{a}(t)$ – tuo pačiu laiko momentu pateikiamas automobilio registracijos numerio simbolio normalizuotas vaizdas; h_{L_i} – artumo funkcija, randama iš lygties

$$h_{L_i} = \alpha(t) \exp\left(-\|\mathbf{r}_i - \mathbf{r}_L\|^2 / 2\sigma^2(t)\right); \quad (5)$$

Šioje išraiškoje $\alpha(t)$ yra mokymo greičio koeficientas (tolygiai mažėjantis mokymo metu), $\mathbf{r}_i \in \mathfrak{R}^2$ ir $\mathbf{r}_L \in \mathfrak{R}^2$ – vektoriai, nustatantys esamo neurono (i) ir neurono



3 pav. Kohoneno tinklo struktūra

laimėtojo (L) padėtį dvimatėje erdvėje, $\sigma(t)$ – artumo funkcijos plotis (taip pat tolygiai mažėjantis su kiekvienu mokymo žingsniu). Norint priskirti simbolį konkrečiai klasei, visi neuronai, kurių artumo funkcijos vertė yra didesnė nei 0,01, priskiriami šiai klasei. Mokymo metu tinklui pateikiami įvairūs automobilio registracijos numerio simboliai, taigi mokymo pabaigoje Kohoneno tinklo neuronų ryšių svorių vektoriai pasiskirsto taip, kad į kiekvieną pateiktą unikalų simbolį aktyviai reaguoja tik vienas neuronas.

Daugiasluoksnius perceptronų tinklas [7] parinktas todėl, kad jis gali aproksimuoti bet kokio sudėtingumo netiesinį sąryšį. Nagrinėjamo tinklo struktūra yra pateikta 4 pav.; čia \mathbf{a} – automobilio registracijos numerio simbolio normalizuotas vaizdas; M^k – neuronų skaičius k -tajame sluoksnyje; \mathbf{s}^k – neuronų k -tojo sluoksnio atsakas (išėjimo signalų vektorius); \mathbf{W}^k – neuronų ryšių tarp $k-1$ ir k sluoksnių svorių matrica; K – neuronų sluoksnių skaičius.

Daugiasluoksniams perceptronų tinklui mokytį taikomas klasikinis atgalinio klaidos sklaidimo algoritmas:

$$\tilde{w}_{ji}^{k-1} = (1 - \mu) w_{ji}^{k-1} - \mu \lambda s_j^{k-1} \delta_i^k; \quad (6)$$

čia w_{ji}^{k-1} – ryšio, jungiančio j -tąjį neuroną sluoksnyje $k-1$ su i -tuoju neuronu k -tajame sluoksnyje, svoris; \tilde{w}_{ji}^{k-1} – to paties ryšio svorio nauja vertė; μ – inercijos konstanta; λ – mokymo greičio konstanta; s_j^{k-1} – j -tojo neurono, esančio $k-1$ sluoksnyje, išėjimo signalas; δ_i^k – i -tojo neurono, esančio k -tajame sluoksnyje, klaida, apskaičiuojama pagal lygtį

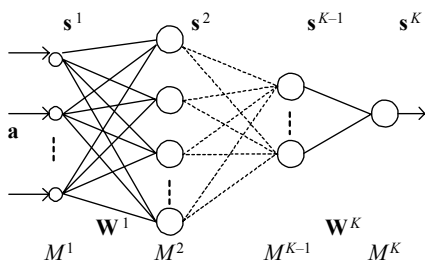
$$\delta_i^k = \begin{cases} f'(z_i^k)(s_i^k - \bar{s}_i), & \text{kai } k = K, \\ f'(z_i^k) \sum_j w_{ij}^k \delta_j^{k+1}, & \text{kai } 1 \leq k < K; \end{cases} \quad (7)$$

čia \bar{s}_i – norimas i -tojo neurono išėjimo signalas; z_i^k – visų i -tąjį neuroną, esantį k -tajame sluoksnyje, patenkančių signalų suma; $f'(\cdot)$ – neurono aktyvinimo funkcijos išvestinė. Sigmoidinė ar hiperbolinio tangento funkcijos yra dažniausiai taikomos neurono aktyvinimo funkcijos.

Visų i -tąjį neuroną k -tajame sluoksnyje patenkančių signalų suma

$$z_i^k = \sum_{j=1}^{M^{k-1}} w_{ij}^{k-1} s_j^{k-1}; \quad (8)$$

čia i -tojo neurono $k-1$ sluoksnyje išėjimo signalas apskaičiuojamas pagal lygtį $s_i^{k-1} = f(z_i^{k-1})$.



4 pav. Daugiasluoksnių perceptronų tinklo struktūra

Skaitmenų atpažinimo dirbtinių neuronų tinklais tyrimo rezultatai

Tirsime, kokie parinktų DNT struktūrų bei mokymo taisyklių parametrai leidžia tiksliai atpažinti automobilių registracijos numeriuose esančius skaitmenis.

Kohoneno tinklui parenkame Gauso artumo funkciją, kai $\sigma(t) \in [10^{-1}, 3]$, $\alpha(t) \in [10^{-4}, 1]$, o maksimalus iteracijų skaičius yra 3 000. Keisdami tinklo dydį, tirsime įėjimo pavyzdžių normavimo ribas. Taikydami du skaičių rinkinius: vieną – mokymui, kitą – tikrinimui, randame simbolių atpažinimo tikslumą. Gauti eksperimentų rezultatai pateikti 1 lentelėje. Matyti, jog įėjimo pavyzdžius centruoti „0“ atžvilgiu yra tikslinga, taigi tolesniuose tyrimuose pavyzdžių amplitudes normuosime $[+1, -1]$ ribose.

Tinklo dydžio parinkimas yra vienas svarbiausių uždavinių, kadangi jis turi įtakos tinklo apibendrinimo savybėms. Kohoneno tinklo įėjimo sluoksnio neuronų skaičių lemia įėjimo pavyzdžio dydis, o išėjimo sluoksnio dydį galima parinkti. Šių eksperimentų rezultatai pateikti 2 lentelėje. Matyti, jog 20×20 neuronų išėjimo sluoksnio taikymas garantuoja didžiausią atpažinimo tikslumą.

Parinkto Kohoneno tinklo atpažinimo tikslumą galima dar labiau padidinti. Remiantis 3 lentelėje pateiktais eksperimentų rezultatais, matyti, jog, esant tinkamiausioms parametų vėrtėms $\alpha_{\max}(t) = 0,5$ bei $\sigma_{\max}(t) = 3$, atpažinimo tikslumas siekia 90%. Atpažinti visus tikrinimo rinkinio simbolius pavyksta tik padidinus mokymo rinkinių skaičių iki 5.

1 lentelė. Įėjimo pavyzdžių ribinių verčių parinkimo rezultatai

Išėjimo sluoksnis, $N_X \times N_Y$	Pavyzdžių vėrtės		Atpažinimo tikslumas, %
	max	min	
10×10	1	-1	70
	0	-1	80
	0	1	60
20×10	1	-1	70
	0	-1	60
	0	1	70
20×20	1	-1	80
	0	-1	50
	0	1	70

2 lentelė. Kohoneno tinklo dydžio parinkimo rezultatai

Išėjimo sluoksnis, $N_X \times N_Y$	Atpažinimo tikslumas, %
10×10	70
20×10	70
20×20	80
30×20	70
30×30	50

3 lentelė. $\alpha(t)$ ir $\sigma(t)$ parinkimo rezultatai

$\alpha(t)$		Atpažinimo tikslumas, %	$\sigma(t)$		Atpažinimo tikslumas, %
max	min		max	min	
1,0	10^{-4}	80	5	10^{-1}	80
0,7		80	4		80
0,5		90	3		90
0,3		70	2		60
0,1		50	1		50

Daugiasluoksniu perceptronu tinklo (DPT) mokymo metu parinktas inercijos koeficientas $\mu = 0,3$. Bandyta parinkti tinklo dydį, tačiau gautas nepakankamas atpažinimo tikslumas (žr. 4 lentelės dalį, kai $\gamma = 0$).

Nuspręsta *mažinti ryšių svorius*, nes tai gerina tinklo apibendrinimo savybes ir kartu turėtų padidinti atpažinimo tikslumą. Pasirinkta ryšių svorius mažinti kas 20-tąjį mokymo žingsnį pagal lygtį $\tilde{w}_{ji}^{k-1} = (1-\gamma)\tilde{w}_{ji}^{k-1}$, čia γ – mažinimo koeficientas, kurio vertė eksperimentiškai parinkta lygi 0,1. Rezultatai (pateikti 4 lentelės antroje dalyje) patvirtina, jog gautas geras atpažinimo tikslumas net esant mažesniai tinklo dydžiui.

Tinkamo *tinklo dydžio parinkimas* susijęs su atpažinimo kokybės įvertinimu. Turimo tikrinimo rinkinio atpažinimo kokybę vertinsime dviem požūriais: a) nustatydami mažiausią laimėtojo ir jam artimiausio neuronų išėjimo signalų skirtumą, įvertinsime blogiausią atvejį; b) nustatydami vidurkį skirtumų tarp laimėtojo ir

4 lentelė. γ įtaka DPT dydžiui

γ	Neuronų skaičius, $M^2-M^3-M^4$	Mokymo		Atpažinimo tikslumas, %
		Rinkiniai	trukmė	
0	30–30–10	1	230	60%
	100–30–10	5	146	80%
0,1	15–30–10	5	554	Arti 100%
	15–25–10	5	1620	Arti 100%
	15–20–10	5	554	Arti 100%
	10–15–10	5	596	90%

5 lentelė. DPT dydžio parinkimo rezultatai

Neuronų skaičius, $M^2-M^3-M^4$	Mokymo trukmė	Minimalus skirtumas	Skirtumų vidurkis
15–30–10	1239	0,077	0,598
15–25–10	2420	0,023	0,591
15–20–10	960	0,167	0,565

R. Laptik, D. Navakauskas. Dirbtinių neuronų tinklų taikymas automobilių registracijos numeriams atpažinti // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2005. – Nr. 8(64). – P. 27–30.

Nagrinėjamas automobilių valstybinių registracijos numerių atpažinimas taikant dirbtinių neuronų tinklus. Pristatoma kuriamos sistemos struktūra, susidedanti iš internetinės kameros, asmeninio kompiuterio ir programinio modulio. Trumpai aptariama atpažinimo programinio modulio sudėtis, t. y. esminiai apdorojimo etapai: pirminis vaizdo apdorojimas, automobilio registracijos numerio vaizdo išskyrimas, segmentavimas į atskirus simbolius, simbolių vaizdų normalizavimas, bei pavienių simbolių atpažinimas. Pateikiamos Kohoneno ir daugiasluoksnių perceptronų tinklų struktūros ir mokymo taisyklės, šių dirbtinių neuronų tinklų taikymas paviniams simboliams klasifikuoti. Pateikiami tinklo apibendrinimą gerinantys mokymo taisyklių parametrai. Eksperimentų rezultatais grindžiami tinkamiausi tinklų struktūrų bei jų mokymo taisyklių parametrai. Il. 4, bibl. 7 (lietuvių kalba; santraukos lietuvių, anglų ir rusų k.).

R. Laptik, D. Navakauskas. Application of Artificial Neural Networks for the Recognition of Car Number Plates // Electronics and Electrical Engineering. – Kaunas: Technologija, 2005. – No. 8(64). – P. 27–30.

The problem of recognition of car number plates by artificial neural networks is studied. The recognition system being created and consisting of personal computer with WEB camera and special software is presented. The composition of the recognition software module (preprocessing, extraction of the image of car number plate, image segmentation, recognition of individual symbols blocks) is briefly discussed. By the presentation of structures and training rules an analysis of Kohonen Maps and Multilayer Perceptrons for the symbol recognition is done. Results of experiments from the recognition of numerals from car number plates are provided. The estimation of the size of networks, parameters involved in their training rules is done. Ill. 4, bibl. 7 (in Lithuanian; summaries in Lithuanian, English, Russian).

Р. Лаптик, Д. Навакаускас. Использование искусственных нейронных сетей для распознавания автомобильных регистрационных номеров // Электроника и электротехника. – Каунас: Технология, 2005. – № 8(64). – С. 27–30.

Рассматривается применение искусственных нейронных сетей для распознавания автомобильных регистрационных номеров. Представлена структура системы распознавания номеров, состоящая из персонального компьютера и интернетной камеры. Обсуждается программный модуль распознавания символов, состоящий из блоков первоначальной обработки, определения позиции номера, выделения символов, распознавания отдельных символов. Распознавание символов рекомендуется выполнять с помощью искусственных нейронных сетей. Рассматриваются сеть Кохонена и многослойный перцептрон, их пригодность для классификации символов. Проведены эксперименты, найдены подходящие размеры сетей, параметры их обучения. Ил. 4, библи. 7 (на литовском языке; рефераты на литовском, английском и русском яз.).

jam artimiausio neuronų išėjimo signalų, įvertinsime vidutinę kokybę. Natūralu, kad esant didesnėms įverčių vertėms, atpažinimo kokybė yra geresnė. Šių eksperimentų rezultatai pateikti 5 lentelėje. Matyti, jog tinkamiausias DPT dydis yra 15–20–10. Papildomais eksperimentais nustatyta, jog šio tinklo tinkamiausias mokymo žingsnis yra 0,05.

Išvados

Trumpai pristatyta kuriamos automobilių registracijos numerių atpažinimo sistemos struktūra bei aptarta atpažinimo programinio modulio sudėtis. Parodyta, jog Kohoneno ir daugiasluoksniais perceptronų tinklais galima klasifikuoti pavinius automobilių registracijos numerių simbolius. Eksperimentiškai nustatyti nagrinėtų DNT struktūrų bei jų mokymo taisyklių tinkamiausi parametrai.

Literatūra

1. **Draghici S.** VisiCar System // Wayne State University, 1997.
2. Number Plate Recognition Software Engine – CARMEN.– Adaptive Recognition Hungary.– <http://www.anpr.net>
3. **Automatic** Number Plate Reader – IRIS // Roke Manor Research Limited, 1998.– <http://www.roke.co.uk>
4. **Navakauskas D., Paulikas Š.** Dirbtinių neuronų tinklai kalbos signalams apdoroti // Elektronika ir elektrotechnika.– Kaunas: Technologija, 1995. – Nr. 4. – P. 26–34.
5. **Paulikas Š., Navakauskas D.** Adaptyviosios signalų apdorojimo sistemos // Elektronika ir elektrotechnika.– Kaunas: Technologija, 1996. – Nr. 1(5).– P. 33–37.
6. **Kohonen T.** Self-Organizing Maps.– Springer-Verlag, 2001.
7. **Rumelhart D. E., Hinton G. E., Williams R. J.** Learning internal representations by error propagation // In Parallel Distributed Processing: Explorations in the Microstructures of Cognition.– MA: MIT Press, 1986.– Vol. 1.– P. 318–362.

Pateikta spaudai 2005 03 14