

DIRBTINIO INTELEKTO ATPAŽINIMO METODŲ ANALIZĖ IR TAIKYMAI RANKA RAŠY TAM TEKSTUI ATPAŽINTI

Gediminas Kavaliauskas, Gražvydas Felinskas

Šiaulių universitetas, Matematikos ir informatikos fakultetas

Ivadas

Viena iš dažniausiai pasitaikančių dirbtinio intelekto, kaip mokslo šakos, probleminių sričių yra uždaviniai, susiję su įvairiais atpažinimo metodais (angl. *pattern recognition*). Ši probleminė sritis dar skyla į tokias atskiras dalis: spausdinto teksto atpažinimas (angl. *optical character recognition*), rašytinio teksto atpažinimas (angl. *handwriting recognition*), kalbos atpažinimas (angl. *speech recognition*) ir veidų atpažinimas (angl. *face recognition*). Programos, sukurtos remiantis atpažinimo metodo pagrindu, bendrauja su aplinka ir iš jos gautą informaciją lygina su duomenų bazėje esančiais šablonais. Tokių programų poreikis atsiranda tada, kai susiduriama su dideliais informacijos kiekiais, kuriems apdoroti gali prireikti daug laiko ar daugelio žmonių pastangų. Ši poreikį taip pat didina sparčiai populiarėjantys ir tobulėjantys įvairūs mobilieji įrenginiai, kurie šiandien gali įvykdyti tai, ką seniau galėjo atlikti tik galingi kompiuteriai su specialia ir brangiai kainuojančia programine įranga.

Šiame darbe susitelkta į vieną iš minėtų sričių – ranka rašyto teksto atpažinimą, kuri turi dvi panašias ir kartu skirtingas pagal savo taikymo sritį šakas. Viena iš šakų yra realiu laiku rašomo teksto atpažinimas (angl. *online recognition*). Įrankiai, kuriuose naudojami šie metodai, yra labai ištobulinti ir pasiekia aukštą patikimumo procentą. Tuo tarpu antroji šaka, minima gan retai, – tai ranka rašyto teksto atpažinimas iš statinio vaizdo (angl. *offline recognition*). Šioje srityje susiduriama su įvairiomis problemomis (Petter, Fragoso, Turk, Baur, 2011), pagrindinės iš jų – parašyto teksto išskaidymas atskirais simboliais, kompiuterinės sistemos apmokymas pagal tam tikrą raštą ir galų gale paties teksto atpažinimas. Šioms problemoms spręsti galima taikyti įvairius metodus.

Tyrimo tikslas – įvairių ranka rašyto teksto atpažinimo metodų bei algoritmų analizė ir jų pritaikymas, kuriant ranka rašyto teksto atpažinimo įrankį.

Uždaviniai: Išanalizuoti dirbtinio intelekto atpažinimo metodus bei jų taikymo sritis; išanalizuoti ir palyginti ranka rašyto teksto atpažinimo algoritmų efektyvumą; sukurti programinę įrangą, skirtą ranka rašyto teksto atpažinimui; pasitelkiant sukurtą įrankio testavimo rezultatus, atlikti probleminių sričių analizę; atlikti tokių metodų taikymų kitose srityse galimybių studiją.

Ranka rašyto teksto atpažinimo poreikio atsiradimas

Šiandien mus kiekviename žingsnyje lydi įvairios technologijos, žmonės vis dažniau vietoj ranka rašyto renkasi kompiuteriais spausdintus tekstus. Tačiau sutiksime ir daug tokių, kurie technologijų pilname pasaulyje nori informacija dalintis ranka parašyto teksto pavidalu. Dėl to atsiranda poreikis tą informaciją paversti kompiuteriui suprantamu simbolių rinkiniu. Žmonių noras panaudoti ranka rašytą tekstą įvairiuose įrenginiuose siekia net 1888 metus, kai buvo išduotas patentas įrenginiui (Gray), pavadintam teleautografu (angl. *Teleautograph*). Jis buvo skirtas ant specialios planšetės ranka parašytam tekstui perduoti į kitą tokį patį įrenginį, kuris sugebėtų atkurti tą patį vaizdą. 1915 metais buvo užpatentuotas pirmasis įrenginys, galėjęs atpažinti rašomus simbolius, analizuojant atliekamus veiksmus (Goldberg).

Rašytinio teksto atpažinimo galimybės šiuolaikiniuose įrenginiuose

Ranka rašyto teksto atpažinimas išlieka aktuali problema ir dabartiniuose įrenginiuose. Šį funkcionalumo trūkumą sprendžia operacinių sistemų kūrėjai, kuriems į pagalbą ateina įvairūs programinės įrangos gamintojai. Tačiau tiek vieni, tiek kiti susiduria su rašytinio teksto atpažinimo sistemos apmokymo poreikiu, t. y. iš pradžių sistema turi susieti kiekvieną rašomą simbolį su kompiuteriniu atitikmeniu.

Šiuo metu labiausiai išsiskirianti, galimybių atpažinti tekstą teikianti programinė įranga yra „MyScript Stylus“, sukurta kompanijos „Vision Objects“. Ši programinė įranga veikia visose populiariausiose operacinėse sistemose, ranka rašytą tekstą sugeba atpažinti 87 kalbomis, naudojasi teksto atpažinimo metodu, kai yra sekamas kiekvienas simbolio rašymo metu atliekamas judesys. Dar viena „MyScript Stylus“ savybė yra jos nuolatinis mokymasis – ilgiau pasinaudojus šia sistema, galima pasiekti beveik 100% teksto atpažinimo patikimumą (MyScript Stylus). Kompanijos „Microsoft“ operacinėje sistemoje „Windows 7“, skirtoje planšetiniams kompiuteriams, vartojama sistema, kuri taip pat grįsta apmokymu (Hui Ma, 2009).

Rašytinio teksto atpažinimo realiu laiku metodas taikomas dažnai, tuo tarpu rasti efektyvų įrankį,

kuris atpažinimą vykdytų iš paveikslėlio, sudėtinga. To priežastis – daug laiko reikalaujantis tokių sistemų realizavimas, nes reikia priimti nemažai sudėtingų sprendimų (Shridar, Houle, Kimura, 2009). Vienas iš įrankių, kuris siūlytų atpažinimą tokiu metodu, – tai „ABYY Finereader 11“, iki šiol naudotas tik spausdintam tekstui atpažinti. Tačiau paskutinėse šios programinės įrangos versijose buvo pridėtas papildomas funkcionalumas, leidžiantis programą apmokyti pagal raštą ir taip atlikti ranka rašyto teksto atpažinimo užduotis. Šio įrankio testavimas parodė, kad realizuoti atpažinimo mechanizmus dar reikia labai tobulinti.

Teksto atpažinimo metodai

Spausdinto ir rašytinio teksto atpažinimas yra vienos dirbtinio intelekto atpažinimo metodų šakos probleminės sritys. Nemažai rašytinio teksto atpažinimo metodų ir technologijų perimta, būtent, iš spausdinto teksto atpažinimo šakos. Pagrindinis šių technologijų skirtumas – kitaip nei spausdintame tekste, rašytiniame tekste kiekvienas simbolis gali turėti labai daug variacijų net tada, kai rašo tas pats asmuo (Jagadeesk, Kannan, Prabhakar, 2008). Programos, turinčios apdoroti tokį tekstą, turi gebėti atpažinti šimtus ar net tūkstančius vienos raidės vaizdavimo būdų. Rašytinio teksto atpažinimo atveju dažniausiai susiduriama su teksto atpažinimu iš paveikslėlio bei teksto atpažinimu realiu rašymo laiku, išsaugant kiekvieno judesio koordinatas (Kala, Vazirani, Shukla, Tiwari, 2010).

Realiu laiku rašomo teksto atpažinimo metodas pasižymi tuo: rašant simbolį, sistema renka papildomus duomenis apie visą rašomo simbolio procesą ir tai padeda lengviau atpažinti ir išskirti, kas buvo parašyta. Išskiriamos penkios pagrindinės tokios sistemos savybės (Singh, Sharma, 2007):

- Viskas vyksta realiu laiku ir įvesti simboliai iškart atpažįstami.
- Galima teksto atpažinimo korekcija realiu laiku – rašantis asmuo iškart mato, kokį simbolį atpažino sistema ir neteisingo atpažinimo atveju gali jį pakeisti.
- Teksto įvedimo metu renkama papildoma informacija apie rašiklio koordinatas, kiekvieno judesio eiliškumą, kryptį ir greitį.
- Nereikia atlikti daug papildomų veiksmų, prieš vykdant simbolio atpažinimo procesą.
- Lengva atskirti simbolius vieną nuo kito, remiantis informacija apie rašiklio atitraukimą nuo rašymo lentos.

Metodo privalumas: kadangi nereikia papildomai apdoroti įvestą simbolį, sutaupoma daug laiko ir atminties teksto atpažinimo procesui atlikti. Pagrindinis šio metodo trūkumas – tekstas rašomas po vieną

simbolį, o tai gali sukelti diskomfortą, jei žmogus įpratęs rašyti suliejant vieną su kitu. Be to, atsisakius papildomų teksto atpažinimo operacijų, smarkiai apribojamos teksto efektyvaus atpažinimo galimybės.

Rašytinio teksto atpažinimo iš nuotraukos metodas, priešingai nei realiu laiku rašomo teksto atpažinimo atveju, vartojamas sistemoje, dirbančiose su statiniu vaizdu Jos, prieš pradėdamos simbolio atpažinimo procedūrą, neturi jokios papildomos informacijos apie parašytus simbolius. Norint pradėti teksto atpažinimą, turi būti atlikta tam tikrų veiksmų seka, kuri pradėdama išankstinio apdorojimo operacija (angl. *preprocessing*). Jos metu dažniausiai atliekami tokie veiksmai:

- Teksto atvaizdas išvalomas nuo „triukšmo“ – pašalinami atsitiktiniai pikseliai, kurie galėjo atsirasti skenuojant parašytą tekstą.
- Atliekama vaizdo aštrėjimo (angl. *sharpening*) procedūra, leidžianti išryškinti simbolių kontūrus.
- Atliekamas teksto plonėjimas, kurio metu parašytame tekste siekiama palikti tik pagrindinius pikselius, reikalingus tekstui atpažinti.

Šių atliekamų veiksmų paskirtis, priklausomai nuo vėliau planuojamo taikyti teksto atpažinimo metodo, įvairi. Vienu atveju tiesiog siekiama palengvinti ir pagerinti simbolių atpažinimo tikimybę, kitu atveju – sumažinti duomenų kiekį, kurį vėliau turės apdoroti teksto atpažinimo algoritmas. Kita operacija, reikalinga tekstui atpažinti, yra *segmentacija*. Jos metu siekiama iš vientiso teksto išgauti atskiras jo eilutes, iš jų – atskirus žodžius, galiausiai – atskirus simbolius, kuriuos jau galėtų apdoroti teksto atpažinimo algoritmas. Pagrindinė teksto atpažinimo proceso operacija – bruožų išskyrimas. Jos metu nustatoma, kurie algoritmai bus taikomi parašytiems simboliams atpažinti ir kaip efektyviai tai bus atliekama. Galų gale lieka klasifikavimo operacija – pats simbolių palyginimas su duomenų bazėje turimais šablonais ir jų tinkamas priskyrimas.

Visos išvardytos operacijos būtų bevertės, jei nebūtų taikomi sprendimo priėmimo modeliai. Populiariausi ir dažniausi iš jų yra *genetiniai*, *paslėpti Markovo* (angl. *hidden Markov*), *Bajeso tinklų*, *dirbtinių neuroninių tinklų* modeliai. Pasirinktas modelis turi lemti įtakos tam, kaip bandoma atpažinti ranka rašytą tekstą. Vis dažniau šie modeliai derinami tarpusavyje, siekiant efektyvesnio teksto atpažinimo. Nepriklausomai nuo to, kuris teksto atpažinimo modelis vartojamas, visos teksto atpažinimo sistemos turi būti „apmokytos“. Teksto atpažinimo atveju apmokymo duomenys yra simbolių šablonai (tam tikro asmens), pagal kuriuos bandoma nustatyti analizuojamą raidę.

Priklausomai nuo to, *kokį* tekstą norime atpažinti, ar nuo to, *kaip* mes norime vykdyti teksto

atpažinimą iš paveikslėlio, šis dar skaidomas į 3 kategorijas:

- OCR (Optical character recognition) – spausdintų simbolių atpažinimas.
- ICR (Intelligent character recognition) – ranka rašytų simbolių atpažinimas.
- IWR (Intelligent word recognition) – ranka rašytų žodžių atpažinimas.

Teksto atpažinimo ir vertimo į redaguojamus kompiuterinius dokumentus uždaviniai sprendžiami jau daugiau kaip 30 metų, tačiau net šiandien dar nėra tokios programos (ar algoritmo), kuri teksto atpažinimą atliktų 100 % tikslumu. OCR yra labiausiai išstobulinta; atpažįstant geros kokybės nuskenuotą dokumentą, pasiekiamas 99,8 % tikslumas (Holley, 2009).

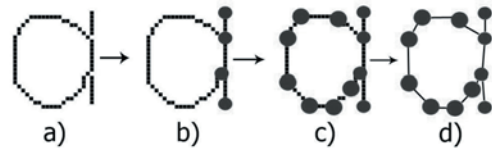
Segmentacija

Bene didžiausia probleminė sritis, atpažįstant rašytinį tekstą, yra segmentacija. Nuo jos priklauso, ar teisingai buvo atskirti simboliai, perduoti tolimesniam apdorojimui. Priešingu atveju, net ir geriausias teksto atpažinimo algoritmas būtų nieko vertas. Dėl šios priežasties, būtent, segmentacija kelia didelį mokslininkų susidomėjimą. Atliekant teksto segmentaciją, iškyla nemažai problemų, priklausančių nuo žmogaus rašymo stiliaus. Jas galima padalinti į tokias dalis (Shridhar, Houle, Kiaura, 2009):

- Teksto eilutės išskyrimas iš teksto – ši dalis sukelia sunkumų tuomet, kai teksto eilutės yra labai arti viena kitos ir raidžių dalys papuola į kitas eilutes (sprendimui pritaikomi specialūs filtrai, kurie aptinka tokias raidžių liekanas ir jas pašalina iš atskirtų eilučių).
- Žodžio atskyrimas iš eilutės – problemų išskirti žodžius atsiranda tada, kai to paties žodžio raidės ne visada yra sujungiamos (toku atveju paprasto sprendimo neužtenka ir segmentavimo operacijos metu reikia panaudoti papildomus mechanizmus (Bhattacharyya, Sarma, 2010).
- Didžiųjų ir mažųjų raidžių skirtumo atpažinimas – problemų gali kilti rašant tam tikras raides.
- Sakinio pradžios ir pabaigos aptikimas.

Dirbtinio intelekto metodai

Aštuonetainio grafo metodo esmė yra apdorojamo simbolio vertimas aštuonetainiu grafu, kur kiekvienas simbolio pikselis atitinka grafo viršūnę. Toks sudarytas grafas leidžia pasiekti, kad parašytas simbolis įgautų tam tikrą formą, nepriklausančią nuo rašyenos stiliaus (1 pav.).

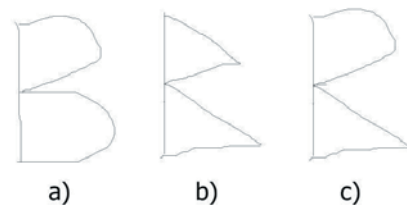


1 pav. „a“ raidės vertimas į aštuonetainį grafą

Siekiant atlikti atpažinimo procedūrą, grafo viršūnėms ir viršūnės jungiančioms briaunoms suteikiami svoriai. Taip pat būtina sužymėti visus grafo bruožus – kilpas, vertikalias linijas, horizontalias linijas, vingiuotas linijas, jų svorius. Atliekant atpažinimo procedūrą, lyginami visi gauto grafo bruožai su duomenų bazėje esančiais raidžių šablonų bruožais, o lyginimo rezultatas išreiškiamas panašumo koeficientu. Po kelių modifikavimų, keičiant bruožų svorius, galiausiai gaunamas reikiamas panašumo koeficientas ir parašyto simbolio grafiui priskiriama raidė. Remiantis šiuo metodu atlikto eksperimento rezultatais, 82 % atvejų parašyti simboliai buvo atpažinti teisingai (Jagadeesh, Kannan, Prabhakar, 2008).

Genetinio algoritmo panaudojimas leidžia išspręsti vieną iš didžiausių rašytinio teksto atpažinimo problemų. Šiuo metodu paremtose sistemose nebūtina didžiulė šabloninių duomenų bazė, mat, genetinio algoritmo principas suteikia galimybę iš kelių turimų duomenų variacijų gauti kitus duomenų rinkinius. Taip atsiranda galimybė kurti dar kitas sprendinių variacijas (2 pav.).

Rašytinio teksto atpažinimo atveju, genetinio algoritmo sugeneruoti nauji sprendimai tarsi imituoja kitų žmonių rašyseną, taip žymiai padidindami tikimybę teisingai nustatyti parašytą raidę. Tai patvirtina R. Kala, H. Vazirani, A. Shukla, B. Tiwari (2010) atliktas tyrimas, kurio rezultatai rodo, jog genetinio algoritmo metodas leido pasiekti 98,44% teisingai nustatytų atvejų.



2 pav. „B“ raidės šablonai
(a ir b – apmokymo duomenys; c – genetinio algoritmo sugeneruotas vaizdas)

Lašelio aptikimo (angl. *Blob detection*) metodas seniai žinomas, jo pagrindu kuriami nauji algoritmai (Danker, 1981). Realizuoti lašelio aptikimo algoritmai dažniausiai veikia sparčiai ir nereikalauja didelių kompiuterinės technikos resursų. Tai ypač

aktualu kuriant programinę įrangą mobiliesiems įrenginiams, kuriuose, panaudojus tokius algoritmus, įmanomi objektų aptikimo uždavinių sprendimai realiuoju laiku. Pavyzdžiui, automobilių numerių aptikimas, teksto vertimas (Petter, Fragoso, Turk, Baur, 2011), judesio sekimo sistemos, panašių objektų paieška ir kiti uždaviniai. Šis metodas pasižymi tuo, kad įvairių taškų ar jų sankaupų (objektų) aptikimas atliekamas pagal jų savybių skirtumą, lyginant su to objekto aplinka. Tos savybės gali būti spalva, ryškumas, šviesumas ar kontrastas. Matematikos požiūriu, šis objektas (ar jo savybės) atvaizde būtų aptinkamas kaip lokalus minimumas ar maksimumas. Šis metodas pasirinktas ranka rašyto teksto atpažinimo įrankiui kurti, panaudojant jį teksto segmentacijos uždavimams atlikti.

Dirbtinių neuronų tinklų metodas – tai mėginimas imituoti žmogaus smegenų darbą. Žinoma, kad žmogaus smegenis sudaro daugybė tarpusavyje sujungtų neuronų, kurie sudaro savitą tinklą, o jame apdorojama ir kartu perduodama įvairi informacija (Laptik, Navakauskas, 2005). Panašiu principu veikia ir *dirbtinių neuronų tinklų* algoritmai (DNT). Neuronų tinklui galima pateikti tam tikrą informaciją, kuri, perėjusi per tinklą, grįžta kaip tam tikras rezultatas. Toks veikimas patogus atliekant teksto atpažinimą – sistemai pateikiamas simbolio atvaizdas, jį tą atvaizdą apdoroja ir gauna rezultatą, kuris išsaugomas. Tai vadinama neuronų tinklo apmokyimu. Tokiai apmokytai sistemai vėliau davus apdoroti kitą simbolį ir nurodžius tam tikrą paklaidą, ji gali nustatyti, koks tai simbolis (Al-Dinari, Sumari ir kt., 2009). Dėl šio metodo universalumo ir gebėjimo įvertinti įvairias paklaidas, būtent, jis pasirinktas sukurtojo įrankio teksto atpažinimo procedūrai atlikti.

Kompiuterinio vaizdo apdorojimo bibliotekų pasirinkimas

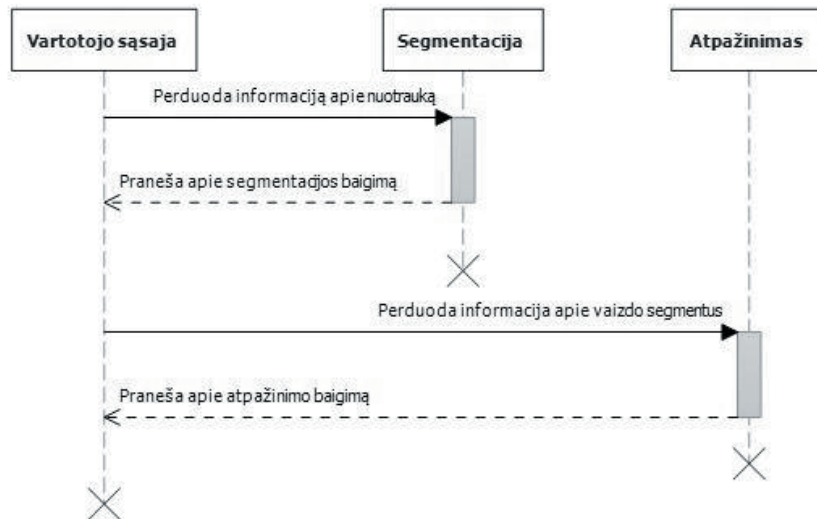
Atlikus rašytinio teksto atpažinimo metodų bei modelių analizę, buvo nuspręsta tolimesnius darbus vykdyti taikant „*offline*“ rašytinio teksto atpažinimo metodus. Siekiant sukurti įrankį, kuriame būtų pritaikyti minėti metodai, atlikta įvairių kompiuterinio vaizdo apdorojimo (angl. *computer vision*) programavimo bibliotekų apžvalga. Iš visų bibliotekų labiausiai išsiskyrė dvi. Jų teikiamų funkcijų prita-

kymo galimybės buvo labai plačios ir pristatytos su detalio dokumentacija, – „OpenCV“ ir „Aforge.NET“. „OpenCV“ biblioteka platinama BSD licenzijos pagrindu. Šią biblioteką nemokamai galima naudoti tiek mokslo, tiek ir komercijos tikslais. Jos branduolys parašytas naudojant C programavimo kalbą, tačiau kartu joje yra realizuotos pilnos C++ ir Python programavimo kalbų sąsajos. „Aforge.NET“ biblioteka platinama LGPL v3 licenzijos pagrindu. Tai taip pat atvirojo kodo biblioteka, kurią gali naudoti įvairių programų kūrėjai bei mokslininkai, dirbantys „Computer vision“ ir dirbtinio intelekto srityse. Ši biblioteka orientuota į .NET karkasą (angl. framework) ir yra parašyta C# programavimo kalbos pagrindu.

Rašytinio teksto atpažinimo programos projekto realizacija

Realizuotą ranka rašyto teksto atpažinimo sistemą sudaro trys pagrindinės dalys: teksto segmentaciją atliekanti paprogramė, teksto atpažinimą atliekanti paprogramė bei už vartotojo sąsają ir visų sistemos dalių veikimą atsakinga pagrindinė programa. Galutinę sistemą sudaro kelios paprogramės su jas valdančia pagrindine programa. Sukurtos sistemos veikimo schema pavaizduota 3 pav. Šioje programoje įvestas valdymo mechanizmas, kuris kitas paprogrames paleidžia tik tam tikromis sąlygomis, o visos sistemos darbas pagrįstas sinchronizuotų pranešimų apsikeitimu.

Segmentacijos paprogramės veikimą galima išskaidyti į tokius pagrindinius žingsnius: priimti informaciją apie paveikslėlį, paveikslėlyje atlikti visų objektų aptikimą (pasinaudojant „OpenCV“ bibliotekos „Lašelio aptikimo“ funkcija) ir informacijos apie rastus objektus išsaugojimą, paskaičiuoti vidutinį vieno objekto aukštį, išskaidyti pateiktą paveikslėlį į eilutes (remiantis vidutinio objekto aukščio parametru), analizuoti kiekvieno objekto eilutėje koordinates x ir y ašyje ir taip sudaryti sąrašą, kokia eilės tvarka turėtų eiti objektai vienas po kito, nustatyta eilės tvarka apie kiekvieną objektą sukurti virtualų keturkampį ir tame keturkampyje atsідūrusį vaizdą išsaugoti naujo paveikslėlio pavidalu, pranešti pagrindinei programai, jog segmentavimas baigtas. Programos veiksmų rezultatas pateiktas 4 ir 5 pav.



3 pav. Sistemos komponentų sąveika



4 pav. Teksto segmentacija į eilutes

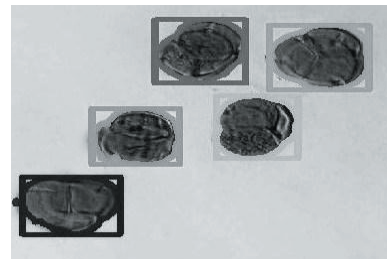


5 pav. Teksto segmentacija į simbolius

Teksto atpažinimą atliekančios paprogramės veikimas: priimti informaciją apie po segmentavimo sudėtus duomenis, normalizuoti kiekvieną iš turimų atvaizdų į tam tikrą vieno dydžio šabloną (pasinaudojant „Encog“ bibliotekos funkcija), gautus šablonus versti į bitų masyvus, kiekvieno objekto bitų masyvą perduoti į „Encog“ bibliotekos siūlomą neuronų tinklą, rezultatą iš neuronų tinklo lyginti su iš anksto sistemos apmokytomis duomenimis ir taip išrinkti labiausiai tinkantį variantą, rezultatus išsaugoti į .csv ir .txt formato rinkmenas, pranešti apie darbo pabaigą.

Atliekant teksto atpažinimo programos realizavimą ir testavimą, pastebėta, kad pasirinktos priemonės, atlikus tam tikras korekcijas, gali būti vartojamos ir kitiems uždaviniams spręsti. Viena pagrindinių priežasčių, leidžiančių taikyti sukurtą sistemą kitoms sritims, tai „lašelio aptikimo“ algoritmo taikymas. Šis algoritmas leidžia aptikti ne tik teksto simbolius, bet ir kitus objektus, kurie koku nors požymiu išsiskiria iš aplinkos. Buvo testuota sistema

su žiedadulkių nuotraukomis. Eksperimentai parodė (6 pav.), jog analizuoti ir realizuoti metodai ateityje gali būti sėkmingai tobulinami ir taikomi kuriant žiedadulkių atpažinimo ir jų koncentracijos automatinio nustatymo sistemą.



6 pav. Rododendro žiedadulkių vaizdas po lašelio aptikimo metodo pritaikymo

Rezultatų analizė

Atpažinimo sistema testuota lotyniškos abėcėlės su didžiosiomis ir mažosiomis raidėmis. Siekiant paprastumo, sistema apmokyta vienu duomenų rinkiniu, iš kurio kiekviena gauta raidė normalizuojama į 40 pikselių plotį ir 60 pikselių aukštį. Atliekant pirminius sistemos rašto atpažinimo paveikimo testus, pastebėta, kad rezultatų efektyvumas gan žemas. Lygiagrečiai „dirbtinių neuronų tinklų“ algoritmui buvo sukurtas papildomas metodas. Šio metodo veikimo principas grįstas gauto objekto, paversto bitų masyvu, analize. Šį metodą realizuoti buvo labai patogu, nes dirbtinis neuronų tinklas, kurio pagalba vyksta teksto atpažinimas, taip pat naudoja panašią duomenų struktūrą. Apibendrinti testavimo rezultatai pateikti 1 lentelėje.

1 lentelė. *Teksto atpažinimo sistemos efektyvumo tyrimo rezultatai*

| Algoritmas | Dirbtinių neuronų tinklas | | Bitų masių analizė | | Iš viso simbolių |
|----------------------|---------------------------|------------|--------------------|------------|------------------|
| | Teisingų vnt. | Teisingų % | Teisingų vnt. | Teisingų % | |
| Abėcėlė | 45 | 15,96 | 203 | 71,99 | 282 |
| Tvarkingas tekstas | 20 | 5,93 | 231 | 68,55 | 337 |
| Atsitiktinis tekstas | 31 | 8,20 | 247 | 65,34 | 378 |
| Iš viso: | 96 | 9,63 | 681 | 68,30 | 997 |

Dirbtinių neuronų tinklo atveju bandant atpažinti tuos pačius simbolius, su kuriais prieš tai sistema buvo apmokyta, tikslumo procentas svyravo nuo 5 % iki 35 % ir tik nustačius raidės dydį 40 x 60 pikselių, rezultatai tapo stabilesni – apie 30 %. Jau minėtas ir analizuotas normalizacijos procesas, kurio metu iš parašytos raidės padaromas 1 pikselio storio objektas, turėtų padidinti dirbtinių neuronų tin-

klo metodo efektyvumą. Bitų analizės metodas būtų kur kas efektyvesnis, jei būtų galima įtraukti parametras, kuriuo galima nustatyti, ar rašoma raidė buvo didžioji, ar mažoji. Tai padėtų išvengti didžiosios dalies tokių klaidų. Dažnai pasitaiko ir kitų atpažinimo klaidų dėl kai kurių raidžių rašymo panašumų. 2 lentelėje pateikti tokių klaidingai atpažintų raidžių derinių dažnumai.

2 lentelė. *Dažniausiai pasitaikę neteisingai atpažintų simbolių deriniai*

| Derinys | a-u | A-N | c-C | d-M | e-l | h-k | I-S | Y-T | m-x | n-Q | o-O | p-P | s-l | u-d |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Kiekis | 7 | 6 | 16 | 6 | 33 | 21 | 7 | 5 | 5 | 8 | 6 | 7 | 7 | 5 |

Išvados

1. Teorinė analizė parodė, kad ranka rašytam tekstui atpažinti efektyviai gali būti taikomi šie dirbtinio intelekto metodai: „lašelio aptikimo“ ir „dirbtinių neuronų tinklų“. Šie metodai buvo realizuoti sukurtoje programinėje įrangoje.
2. Pagrindinės teksto atpažinimo sistemų probleminės sritys yra teksto skaidymas į simbolius ir teisingas panašiai rašomų simbolių atpažinimas.
3. Sukurtosios programinės įrangos testavimas parodė, jog, vartojant bitų masių analizės algoritmą, simboliai teisingai atpažįstami efektyviau nei taikant dirbtinių neuronų tinklų algoritmą, kurio efektyvumui padidinti reikalingas papildomas normalizacijos procesas.
4. Kiekvienam teksto atpažinimo algoritmui reikalingos skirtingos paruošiamosios procedūros, nuo kurių priklauso algoritmo veikimo efektyvumas.
5. Programinės įrangos testavimas su žiedadulkių vaizdais parodė, jog realizuoti algoritmai ateityje gali būti sėkmingai tobulinami ir taikomi kuriant programinę įrangą, skirtą įvairių žiedadulkių atpažinimui ir jų koncentracijos automatinei analizei.

Literatūra

1. Al-Omari S. A. K., Sumari P., Al-Taweel S. A., Husain A. J. A., 2009, Digital Recognition using Neural Network. *Journal of Computer Science*. Vol. 5. P. 427–434.

2. Bhattacharyya K., Sarma K. K., 2010, ANN-based Innovative Segmentation Method for Handwritten text in Assamese. *International Journal of Computer Science Issues (IJCSI)*. Vol. 7. Issue 4. P. 9–16.
3. Danker A. J., 1981, Blob Detection by Relaxation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 3. Issue 1. P. 79–92.
4. Goldberg H. E., Controller. *United States Patent US1117184* [interaktyvus] [žiūrėta 2012 m. spalio 1 d.]. Prieiga per internetą: <<http://www.freepatentsonline.com/1117184.pdf>>.
5. Gray E., Teleautograph. *United States Patent US0386815* [interaktyvus] [žiūrėta 2012 m. spalio 1 d.]. Prieiga per internetą: <<http://www.freepatentsonline.com/386815.pdf>>.
6. Holley R., 2009, Analysing and Improving OCR Accuracy in Large Scale Historic Newspaper Digitisation Programs. *D-lib Magazine*. Vol. 15. Number 3/4.
7. Hui Ma, 2009, Handwriting Recognition's Odyssey to Windows 7 [interaktyvus] [žiūrėta 2012 m. spalio 1 d.]. Prieiga per internetą: <http://research.microsoft.com/en-us/news/asia/features/ciw_handwriting-rec.aspx>.
8. Jagadeesh Kannan R., Prabhakar R., 2008, An Improved Handwritten Tamil Character Recognition System using Octal Graph. *Journal of Computer Science*, Vol. 4. Issue 7. P. 509–516.
9. Kala R., Vazirani H., Shukla A., Tiwari R., 2010, Offline Handwriting Recognition using Genetic Algorithm. *International Journal of Computer Science Issues (IJCSI)*. Vol. 7. Issue 4. P. 16–25.
10. Laptik R., Navakauskas D., 2005, Dirbtinių neuronų tinklų taikymas automobilių registracijos numeriams atpažinti. *Elektronika ir elektrotechnika*. Nr. 8 (64). P. 27–30.

11. MyScript Stylus. *Vision objects* [interaktyvus] [žiūrėta 2012 m. spalio 1 d.]. Prieiga per internetą: <<http://www.visionobjects.com/en/myscript/about-myscript/>>.
12. Shridhar M., Houle G. F., Kimura F., 2009, Recognition strategies for general handwritten text documents. *Integrated Computer-Aided Engineering*. Vol. 16. Issue 4. P. 299–314.
13. Singh H., Sharma R. K., 2007, Moment in Online Handwritten Character Recognition. *Proceedings of National Conference on Challenges & Opportunities in Information Technology* [interaktyvus] [žiūrėta 2012 m. spalio 1 d.]. Prieiga per internetą: <<http://www.rimtengg.com/coit2007/proceedings/pdfs/46.pdf>>.
14. Petter M., Fragoso V., Turk M., Baur Ch., 2011, Automatic text detection for mobile augmented reality translation. *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. P. 48–55.

THE ANALYSIS OF RECOGNITION METHODS BASED ON ARTIFICIAL INTELLIGENCE AND THEIR APPLICATION IN HANDWRITTEN TEXT RECOGNITION

Gediminas Kavaliauskas, Gražvydas Felinskas

Summary

The fields of applications of the recognition methods based on artificial intelligence increasingly expand. Handwriting text recognition is one of these fields. This article provides analysis of several recognition methods and algorithms based on artificial intelligence. Based on this analysis, an application was created for the purpose of recognising a handwritten text. The text segmentation was implemented using the blob detection algorithm. Text recognition was performed using the bit array analysis algorithm. During the implementation and testing stage of the application, the main problem areas were identified: text segmentation into correct symbols and recognition of symbols that are written in a similar way. This article also provides the review on the perspectives of the application of these recognition methods in other fields.

Key words: artificial intelligence, handwriting recognition, recognition methods and algorithms.

DIRBTINIO INTELEKTO ATPAŽINIMO METODŲ ANALIZĖ IR TAIKYMAI RANKA RAŠYTO TEKSTO ATPAŽINIMUI

Gediminas Kavaliauskas, Gražvydas Felinskas

Santrauka

Dirbtinio intelekto atpažinimo metodų taikymo sritys vis labiau plečiasi. Viena iš sričių yra ranka rašyto teksto atpažinimas. Šiame straipsnyje apžvelgti dirbtinio intelekto atpažinimo metodai, atlikta teksto atpažinimo algoritmų analizė. Remiantis analizės rezultatais, sukurta ranka rašyto teksto atpažinimo programa, kurioje teksto segmentavimo operacija atliekama „lašelio aptikimo“ algoritmu. Teksto atpažinimo operacijai atlikti vartojamas bitų masyvų analizės algoritmas. Realizuojant programą bei atliekant jos testavimą, nustatytos pagrindinės šio tipo programų probleminės sritys: taisyklingas teksto skaidymas į simbolius ir taisyklingas panašiai rašomų simbolių atpažinimas. Šiame straipsnyje taip pat apžvelgtos šių atpažinimo metodų taikymo perspektyvos kitose probleminėse srityse.

Prasminiai žodžiai: dirbtinis intelektas, rašytinio teksto atpažinimas, atpažinimo metodai ir algoritmai.

Įteikta 2012-09-26