



BIOCHEMINIŲ VAIZDŲ CHARAKTERINGŲJŲ POŽYMIŲ SANDARA

Dovilė KURPYTĖ

Vilniaus Gedimino technikos universitetas, Vilnius, Lietuva
El. paštas: dovile.kurpyte@vgtu.lt

Santrauka. Biochemijoje itin intensyviai pradedamos naudoti trimačio vaizdinimo technologijos. Jos leidžia greičiau ir naudojant mažiau resursų, išskirti vaizduose esančius požymius, reikalingus ląstelių savybėms bei biocheminių procesų kilmei išaiškinti. Straipsnyje sprendžiama fluorescencinės mikroskopijos HeLa ląstelių vaizdų charakteringųjų požymių atrankos trimačio vaizdavimo, naudojant neraiškios logikos klasifikatorių, uždaviniui problema. Parodyta, kad tikslinga biocheminių vaizdų trimatei analizei taikyti objektų skaičio ir dydžio charakteringuosius požymius.

Reikšminiai žodžiai: biocheminiai vaizdai, trimatis vaizdavimas, vaizdų charakteringieji požymiai, požymių sistemos sandara, neraiškiosios logikos klasifikatorius.

Ivadas

„Vienas paveikslas vertas daugiau nei tūkstančio žodžių“ – šiandienai ypač aktualia sentencija tapęs posakis, kurio ištakos siekia praeito amžiaus pirmąjį dešimtmetį. Gerai žinoma, kad žmonės daug geriau orientuojasi, jei turi vaizdą, o ne detalų aprašymą. Biochemijoje, nagrinėjant epitelio, kraujo, kamienines ar kitas ląsteles, dažnai gaunami vaizdai, kuriuose matomi fluorescenciškai išryškinti cheminiai junginiai, proteinai ar ląstelės organoidai.

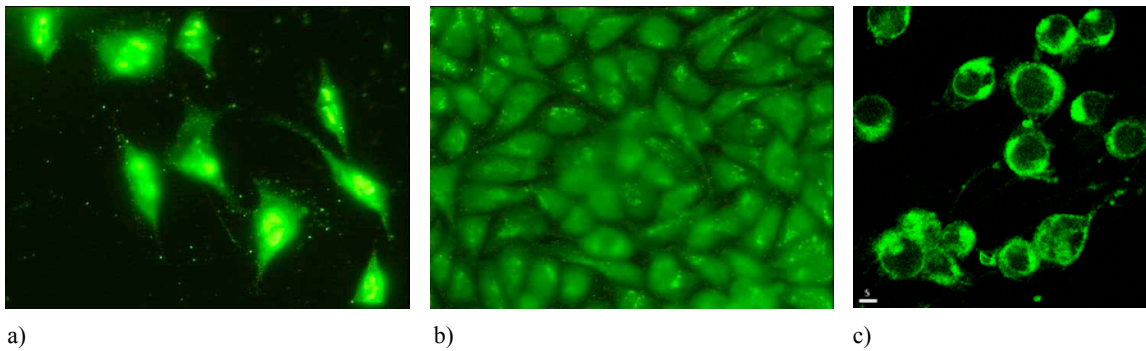
Analizuojant biomedicininis vaizdus, svarbu išskirti ląsteles bei jos organoidus, atskirti triukšmą bei foną. Vaizdo analizė pradedama jo segmentavimu, kurio metu išskiriami prasmingi objektai, ir mūsų atveju bandoma nustatyti ląstelių kontūrus, ląstelėje esančius organoidus, nufiltruoti triukšmą bei atskirti foną. Šie žingsniai yra itin svarbūs, nes rezultatas yra pagrindas tolimesnei ląstelių analizei, pvz., jų vietai, savybėms, judėjimui ar galiausiai ląstelės elgsenai nustatyti.

Dėl aiškesnės biomedicininis vaizdų analizės amė dešimtmetyje pradėtas taikyti trimatis vaizdavimas, kuris atveria naujas galimybes pamatyti tai, kas anksčiau buvo nematoma. Pilnai atkuriami erdvinė ląstelių struktūra leidžia geriau suprasti tiriamo biocheminio junginio funkcionalumą bei paskirtį. Todėl apdorojus ir atvaizdavus turimą pirminį vaizdą gaunami nauji trimačiai modeliai yra kur kas informatyvesni. O šiuos sujungus su papildoma biochemine informacija gaunami itin išsamūs vaizdų rinkiniai, kurie daug greičiau ir paprasčiau suprantami juos naudojančiam

specialistui. Toks ląstelių vaizdavimas gali padėti pamatyti dar ankstyvoje stadijoje prasidedančią ligą ar kitus ieškomus pokyčius.

Biochemijoje tyrimams *in vitro*, pvz., vėžio tyrimuose, dažniausiai naudojamos HeLa ląstelės (Masters 2002). Šios ląstelės buvo paimtos iš 1951 m. į John Hopkins ligoninę Baltimorėje, JAV patekusios Henriettos Lacks. Ji susirgo gimdos kaklelio vėžiu, nuo kurio dar tais pačiais metais mirė. Paėmus biopsiją iš jos neoplazmos ir pasėjus navikines ląsteles dirbtinėje terpėje, buvo pastebėta, kad jos dauginasi atskirai nuo žmogaus kūno. Tai tapo mokslo naujiena, nes iki to laiko niekam nebuvo pavykę dirbtinėje terpėje išauginti žmogaus ląstelių. Taip išvesta nauja ląstelių linija – HeLa, kuri naudojama kaip modelinės žmonių ląstelių analogų linijos daugelyje molekulinis biologinis tyrimų. HeLa ląstelių linijos gali egzistuoti neapibrėžtai ilgai ir jos nemiršta praėjus mirties laikotarpiui, kuris užprogramuotas ląstelėse. Jos ypatingos ir tuo, kad turi žymių genetinių skirtumų, lyginant jas su protėvinėmis *Homo sapiens* ląstelėmis (Van Valen, Maiorana 1991). Taip Henriettos Lacks gydytojas savo atradimu pasidalino su dauguma pasaulio laboratorijų, o šių ląstelių tyrimai itin prisidėjo prie medicinos mokslo pažangos (Rahbari *et al.* 2009).

HeLa vėžinių ląstelių pavyzdžiai pateikiami 1 pav. (VU BChI 2014). Pastebima, kad skirtinguose paveiksluose skiriasi vaizdo ryškumas, ląstelių organoidų švytėjimas, ląstelių kontūrai, formos, dydžiai bei pačių vaizdų kokybė, įtakota biocheminių procesų ruošiant ląsteles *in vitro* tyrimams.



1 pav. Pavyzdiniai vaizdai HeLa vėžinių ląstelių gautų iš: a ir b – epitelio, c – kraujo
 Fig. 1. Image examples of HeLa cancerous cells from: a and b – epithelium, c – blood

Vaizdavimas trimatėje erdvėje tik dar labiau išpopuliarėjo, kai nustatyta, kad sergant specifinėmis ligomis medžiagų kiekis organizme pakinta. Tuomet keičiasi audinio fluorescencijos intensyvumas ir spektras, o juos užregistravus, galima nusakyti ligos pradžią. Fluorescencinio vaizdinimo metu dažnai naudojami specialūs dažikliai, kuriais galima efektyviai vaizdinti net kraujagysles (taikoma fluorescencinėje angiografijoje), pvz., operacijos metu išryškinti širdies raumenį maitinančias vainikines arterijas, siekiant aptikti ir sėkmingai pašalinti kraujagyslių defektus (State 2009).

Tiek daug neapibrėžtumų skatina ieškoti naujų būdų duomenims analizuoti ir klasifikuoti, todėl šioje srityje taip pat siūlomi intelektualieji metodai (Van Valen 1991). Pastaraisiais metais intelektualieji metodai tik dar labiau išpopuliarėjo dėl jų universalumo pritaikant sudėtingose sistemose bei algoritmų galimybių prisiderinti prie duomenų ir greitai bei kokybiškai pateikti rezultatus. Biochemijoje dažnai neįmanoma duomenų apibūdinti dviem būsenomis, todėl dažnai naudojama neraiškioji logika (Zadeh 1965), kuri su negriežtai apibrėžtais kintamaisiais leidžia tiksliau įvertinti ir aprašyti turimus duomenis. Pagal šiuos metodus suklasifikuojami skirtingi ląstelės segmentai net ir realaus laiko sistemose (Wienert *et al.* 2012).

Kuriamas neraiškios logikos klasifikatorius, kitaip tariant, algoritmas, kuris priskiria objektą klasei pagal aprašytas objekto savybes. Intelektualumas slypi klasifikatoriaus galimybeje mokytis prognozuoti klases pagal mokymo algoritmą ir mokymo duomenų rinkinį. Kai tik klasifikatorius yra apmokytas, galima pereiti prie nematytų objektų klasifikavimo naujuose vaizduose. Jei mokymo duomenų rinkinio nėra, klasifikatorius gali būti sukurtas remiantis išankstinėmis žiniomis ir patirtimi, pritaikant *a priori* taisyklių rinkinius, todėl šis metodas ir toliau plinta taikant biocheminius vaizdus. Kitas neraiškiosios logikos klasifikatoriaus taikymo privalumų yra greitaveika, nes įvestis sparčiai apskaičiuojama dėl supaprastintos klasifikatoriaus struktūros.

Straipsnyje nagrinėjama HeLa ląstelių fluorescencinės mikroskopijos vaizdų charakteringųjų požymių atrankos trimačio vaizdavimo, naudojant neraiškios logikos klasifikatorių, uždaviniui problema. Iš pradžių aptariami esminiai reikalavimai vaizdo požymiams bei vaizdams apdoroti taikomų ir galimų taikyti požymių sandara, po to neraiškios logikos klasifikatorius vaizdams segmentuoti, o straipsnio pabaigoje pateikiamas apibendrinimas ir išvados.

Esminiai reikalavimai vaizdo požymiams ir jų sandara

Naudojami didelės raiškos HeLa vėžinių ląstelių epitelio (1 pav. a, b) ir kraujo (1 pav. c) fluorescencijos mikroskopijos vaizdai. Vaizdai yra trijų kanalų (RGB spalvų modelio), 768×576 vaizdo taškų dydžio.

Vaizdo požymių išskyrimas – procesas, kai norima gauti svarbiausius duomenis iš pirminių vaizdų. Todėl požymių išskyrimas reikalingas rasti parametrų rinkiniams, kurie vienareikšmiškai apibrėžia ieškomą objektą. Pagrindinis tikslas yra išgauti funkcijų rinkinį, kuris maksimizuoja pripažinimo normą su mažiausia elementų suma. Duomenų struktūros atžvilgiu vaizdo požymiai skirstomi į dvi klases (Wang *et al.* 2011):

- globalius požymius;
- lokalius požymius.

Globalūs požymiai apima dydį, spalvą, intensyvumą (minimumą, maksimumą) ir pan. Lokalūs požymiai yra geometrija, forma (briaunos, kampai, atšakos, sritys) ir pan. Pastebima, kad globalūs vaizdo požymiai skirti struktūrai nusakyti, o lokalūs – esamai geometrijai. Šiame eksperimente pasirinkti du perspektyvūs globalūs požymiai, kurie aptariami kitame skyrelyje, pristatančiame taikytą neraiškiosios logikos klasifikatorių.

Naudojami vaizdai yra pakankamai ryškūs ir kontrastingi. Papildomi reikalavimai vaizdams taikomi tada, kai yra itin žymūs vaizdų kokybės iškraipymai dėl biocheminių

procesų paruošiant ląsteles *in vitro*. Tokiu atveju priklausomai nuo iškraipymų gali būti taikomas išankstinis vaizdų apdorojimas.

Neraiškiosios logikos klasifikatorius vaizdams segmentuoti

Atliekant kraujo ląstelių segmentavimą, vienas svarbiausių tikslų yra atskirti sudėtingą vaizdo foną su galimais iškraipymais nuo ląstelės su visais matomais organoidais (Adollah *et al.* 2008). Šiam tikslui intelektualieji metodai itin gerai tinka, nes gali prisitaikyti prie nuolat kintančių duomenų.

Jain ir Abraham (2003) atlikti tyrimai parodė, kad naudojant neraiškiosią logiką su JEI-TADA taisyklių rinkiniais pasiektas net 99,73 % klasifikavimo tikslumas. Taikomas neraiškiosios logikos klasifikatorius leidžia matematiškai apibrėžti dalinį teisingumą ir netikslumą, naudojant lingvistinius kintamuosius. Pasitelkiant neraiškiosią aibių teoriją, galima suprantamai formuoti taisykles ir lengvai užprogramuoti lingvistinę informaciją. Lingvistiniai kintamieji, suformuoti iš kelių persidengiančių neraiškiųjų aibių, fizikine prasme apibūdina kintamąjį. Lingvistinis kintamasis „objekto dydis“ sudarytas iš neraiškiųjų aibių: „mažas“, „vidutinis“ ir „didelis“. Kitas lingvistinis kintamasis „objekto skaistis“ sudarytas iš trijų neraiškiųjų aibių:

„blankus“, „vidutinis“ ir „ryškus“. Lingvistiniai kintamieji sukurti tiesiogiai derinantys prie turimų biocheminių fluorescencinių mikroskopijos vaizdų. Atskirose situacijose neraiškiosios aibės gali turėti skirtingas reikšmių sritis ir skirtingas savybes.

1 lentelė. Vaizdo skaidymas į sritis pagal neraiškiosią logiką
Table 1. Image separation to areas based on fuzzy logic

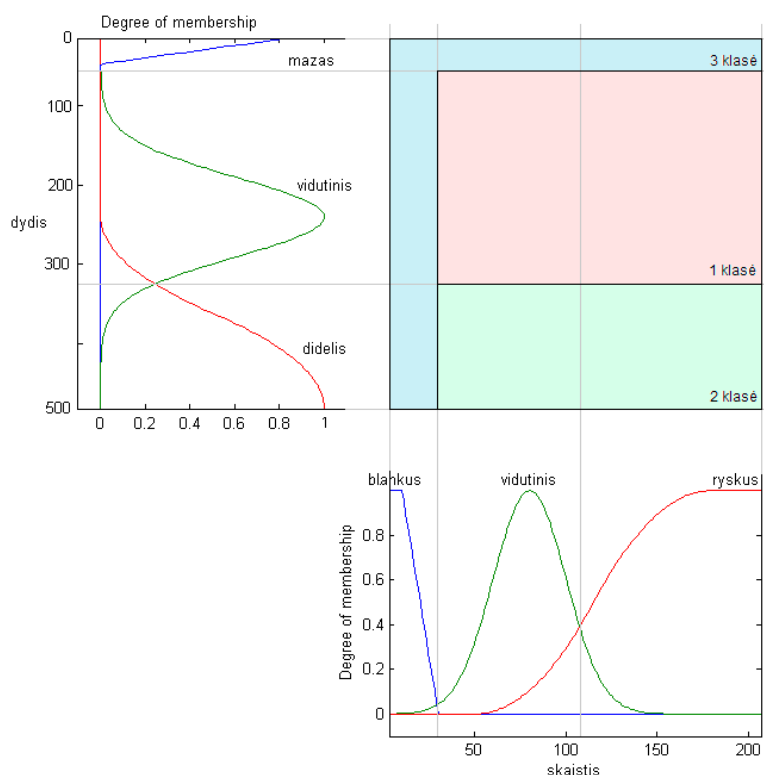
Objekto skaistis	Objekto dydis		
	mažas	vidutinis	didelis
Blankus	3	3	3
Vidutinis	3	1	2
Ryškus	3	1	2

Vaizdo skirstymas (1 lentelė) atliekamas pagal tris skirtingas klases:

- organoidas (1 klasė);
- ląstelė (2 klasė);
- fonas (3 klasė).

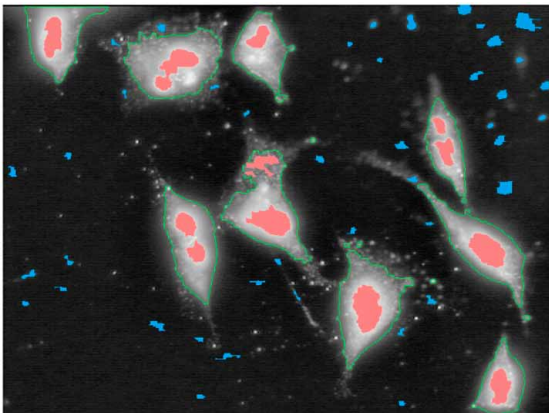
Taigi, sudėjus visas JEI-TADA taisykles, gaunama supaprastinta lentelė. Svarbu, kad neraiškūs klasifikatoriaus pateikia lingvistinės vertės išėjimo skaitines reikšmes, kurios nėra įtrauktos taisyklėse. Taigi kiekviena kalbinė vertė yra perteikiama priklausomybės funkcija (2 pav.).

Neraiškiosios logikos klasifikatoriaus veikimas remiasi dviem požymiais, kai kiekvienas iš jų vertinamas trimis



2 pav. Neraiškiosios priklausomybės funkcijos
Fig. 2. Fuzzy membership functions

neraiškiosiomis aibėmis. Svarbiausia yra atskirti ląstelę nuo „fono“ ir „triukšmo“, todėl vienas pagrindinių požymių yra objekto vaizde skaistis. Objektai esantys vaizde išskiriami į tris minėtas klases. Kiekvienas jų nusakomas požymio lingvistine verte. Klasė „fonas“ apima dvi dalis, t. y. vaizdo foną, dažniausiai tamsiausią vaizdo sritį, bei „triukšmo“ – tai negyvos ląstelės, bei objektai, kurie nėra ląstelės dalis, tačiau švyti dėl juose esančios fluorescencijos. „Triukšmo“ švytėjimas dažnai prilygsta ir ląstelės organoidų skaičiui, tačiau yra daug mažesnis. Jei skaistis „blankus“, galime iš 1 lentelės matyti, kad šio požymio pastarosios neraiškiosios aibės dėka gali būti atskiriamas „fonas“ nuo „ląstelės“. Todėl „fonas“, arba kitaip, dar 3 klasės dalis objektų randama pagal minimalų skaistį, o kita 3 klasės dalis yra „triukšmas“, kuris randamos esant „mažam“ dydžiui ir skirtingo skaisčio taškams. Kadangi organoidai turi tilpti ląstelės rėmuose ir yra ryškiausi dėl intensyvaus fluorescencijos pasisavinimo, išskirtos dvi požymių sritys: „ryškaus“ skaisčio bei „vidutinio“ dydžio. Tokiu būdu atskiriami fluorescencijos turintys ir visai neturintys organoidai.



3 pav. HeLa vėžinių ląstelių vaizdo segmentavimas
Fig. 3. HeLa cancerous cells image segmentation

Organoidų apjungimo ląstelės skeletas yra „didelio“ dydžio ir vidutinio skaisčio, dėl dalinio organoidų fluorescencijos įsisavinimo. Taip sukuriama lingvistinė taisyklė apibrėžiančios sąlygas neraiškiai logikai.

Platus neraiškaus klasifikatoriaus apibrėžimas leidžia sukurti ne vieną tinkamą klasifikavimo modelį. Šiuo atveju naudojamas taisyklių rinkiniu paremtas modelis požymių priklausomybėms nusakyti, pvz., net turint porą skaitinių požymių x_1 (objekto dydis) ir x_2 (objekto skaistis) taisyklėse nusakomos lingvistinės vertės:

$$\tau_1(X) = \mu_{\text{mažas}}^{(1)}(x_1) \wedge \mu_{\text{blankus}}^{(2)}(x_2), \quad (1)$$

čia „JEI“ x_1 {1 požymis} yra „mažas“ {1 požymio lingvistinė vertė} „IR“ x_2 {2 požymis} yra „blankus“ {2 požymio lingvistinė vertė} „TADA“ klasė 1 {atliekamas

klasifikavimas}. Loginė operacija „IR“ paprastai įgyvendinama kaip minimumas, bet kita t -norma gali būti naudojama. Čia $\tau_1(X)$ yra pasirinkimo stiprumo koeficientas.

Norint gauti klasifikatoriaus atsaką, visų taisyklių stiprumo koeficientai sudedami. Tarp daugybės įvairių metodų, kurie čia gali būti taikomi, parinktas maksimalios agregacijos metodas. Kiekvienos klasės vertė x_i patenka į intervalą nuo 0 iki 1:

$$g_k(X) \in [0, 1], \quad k = 1, \dots, c, \quad (2)$$

čia k yra klasių skaičius, $c = 3$.

Tegul $i \rightarrow k$ reiškia tą i taisyklę, kuri pasirenka k klasę, tada:

$$g_k(X) = \max_{i \rightarrow k} \tau_i(X). \quad (3)$$

Sukurto neraiškiosios logikos klasifikatoriaus rezultatai matomi 3 pav., čia mėlynai išryškintas „triukšmas“, žaliai paryškintas ląstelės kontūras, o rausvai išryškinti ląstelėje esantys organoidai.

Apibendrinimas. Išvados

Neraiškiosios logikos klasifikatoriaus parinktas vaizdo segmentavimui atlikti, nes biocheminiai vaizdai pasižymi sudėtingomis savybėmis.

Atlikus biocheminiams vaizdams apdoroti taikomų ir galimų taikyti požymių analizę, požymiai suklasifikuoti išskiriant labiausiai perspektyvius ir tenkinančius suformuluotus reikalavimus – taip pasiūlyta skaisčio ir dydžio vaizdų charakteringų požymių sandara.

Naudojant epitelio ir kraujo vėžinių ląstelių vaizdus, apskaičiuoti vaizdų charakteringieji požymiai ir jais remiantis atliktas vaizduose rastų objektų segmentavimas į tris sritys (ląstelę, ląstelėje esančius branduolius, triukšmą bei foną).

Padėka

Tyrimą finansuoja Lietuvos mokslo taryba (sutarties Nr. MIP-083/2015).

Dėkoju Vilniaus Universiteto Biochemijos instituto, Ląstelės molekulinės biologijos skyriui už originalių epitelio ir kraujo vėžinių HeLa ląstelių fluorescencinės mikroskopijos vaizdus.

Literatūra

- Adollah, R.; Mashor, M. Y.; Mohd Nasir, N. F.; Rosline, H.; Mahsin, H.; Adilah, H. 2008. Blood cell image segmentation: a review, *Biomed* 141–144.
- Jain, R.; Abraham, A. 2003. A comparative study of fuzzy classification methods on breast cancer data, *Australasian*

Physics & Engineering Sciences in Medicine 27(4): 213–218.
<http://dx.doi.org/10.1007/BF03178651>

Masters, J. R. 2002. HeLa cells 50 years on: the good, the bad and the ugly, *Nature Reviews Cancer* 2: 315–319.
<http://dx.doi.org/10.1038/nrc775>

Rahbari, R.; Sheahan, T.; Modes, V.; Collier, P.; Macfarlane, C.; Badge, R. M. 2009. A novel L1 retrotransposon marker for HeLa cell line identification, *BioTechniques* 46(4): 277–84.

State, P. 2009. Naturally fluorescent molecules may serve as cancer biomarker, *Science Daily* [interaktyvus], [žiūrėta 2015 m. vasario 20 d.]. Prieiga per internetą: <http://www.sciencedaily.com/releases/2009/04/090402104734.htm>.

Van Valen, L.; Maiorana, V. C. 1991. HeLa, a new microbial species, *Evolutionary Theory* 10(2): 71–74.

Wang, L.; Zhang, J.; Zang F. 2011. An efficient feature extraction method, global between maximum and local within minimum, and its applications, *Mathematical Problems in Engineering* 1–15. <http://dx.doi.org/10.1155/2011/176058>

Wienert, S.; Heim, D.; Saeger, K.; Stenzinger, A.; Beil, M.; Hufnagl, P.; Dietel, M.; Denkert, C.; Klauschen, F. 2012. Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach, *Scientific reports* 2: 503. <http://dx.doi.org/10.1038/srep00503>

Zadeh, L. A. 1965. Fuzzy sets, *Information and Control* 353: 338–353. [http://dx.doi.org/10.1016/S0019-9958\(65\)90241-X](http://dx.doi.org/10.1016/S0019-9958(65)90241-X)

THE STRUCTURE OF SPECIFIC FEATURES OF BIOCHEMICAL IMAGES

D. Kurpytė

Abstract

Nowadays three-dimensional imaging technology is used in biochemistry very intensively. It helps to distinguish image features faster with less resources, which are necessary to describe cell characteristics and to explain the origin of the biochemical processes. The article on the strength of Fuzzy logic classification solves fluorescent microscopy of HeLa cells image three-dimensional visualisation characteristic features problem. It is shown that object “luminance“ and “size“ characteristic features are useful to apply for three-dimensional biochemical images analysis.

Keywords: biochemical images, three-dimensional visualization, image features, feature set taxonomy, fuzzy logic classifier.