

# Daugiasluoksnis perceptronas veidams atpažinti

Ričardas Toliušis, Olga Kurasova

*Lietuvos edukologijos universitetas, Gamtos, matematikos ir technologijų fakultetas*

Studentų g. 39, LT-08106, Vilnius

E. paštas: ricardas.toliusis@gmail.com, olga.kurasova@leu.lt

**Santrauka.** Šiame straipsnyje aprašomas algoritmas, kuris naudojamas būdingus veido bruožų taškus, apskaičiuoja normalizuotus Euklidinius atstumus tarp įvairių veido sričių ir daugiasluoksnio perceptrono dėka atlieka veidų atpažinimą. Ieškant efektyviausio modelio, eksperimento būdu buvo keičiami įvairūs neuroninio tinklo parametrai, tokie kaip paslėptų sluoksnių bei neuronų skaičius juose, gradientinio nusileidimo optimizavimo algoritmai, paklaidos, aktyvacijos funkcijos bei pateikiami skirtingi atstumų rinkiniai.

**Raktiniai žodžiai:** dirbtiniai neuroniniai tinklai, daugiasluoksnis perceptronas, veidų atpažinimas.

## 1 Įvadas

Veidų atpažinimas yra didelės perspektyvas saugos, verslo, pramogų sferose turinti biometrinė identifikavimo priemonė, jau daugelį metų sulaukianti nemažo mokslininkų susidomėjimo. Nors algoritmai yra gerokai pažengę į priekį, tačiau iki šiol susiduriama su įvairiomis atpažinimo problemomis, tokiomis kaip žmonių senėjimas, pasisukimas profiliu, prasta nuotraukų kokybė, apšvietimas, veido išraiškos ir kt.

Vienas svarbiausių veidų atpažinimo algoritmo etapų yra požymių išgavimas. Egzistuoja įvairios požymių išgavimo metodikos, kai kurios jų yra paremtos geometrinų atstumų matavimu.

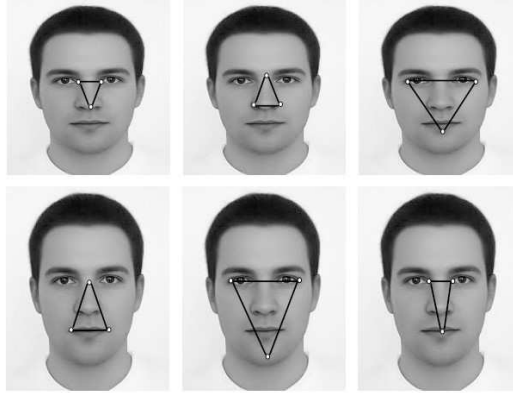
Straipsnyje siūlomas daugiasluoksniu perceptronu pagrįstas algoritmas, kuris veido požymiams identifikuoti naudoja Euklidinius atstumus. Šio algoritmo atpažinimo rezultatus menkai įtakoja apšvietimas ir veido išraiškos.

## 2 Susiję darbai

Dirbtiniai neuroniniai tinklai (DNT) yra vienas efektyviausių klasifikavimo metodų. Veidų atpažinimo algoritmai, grįsti dirbtiniais neuroniniais tinklais, pasiekia vienus geriausių atpažinimo efektyvumo rezultatų.

Sung ir Poggio [5] išgautų požymių klasifikavimui daugiasluoksnį perceptroną pirmieji panaudojo 1994. Jų darbas įkvėpė daugelį kitų mokslininkų veidų atpažinimui naudoti neuroninius tinklus.

Daugelis algoritmų naudoja tik keletą atstumų, pavyzdžiui atstumą tarp akių, nosies plotį, akiduobių gylį, skruostikaulių formą, žandikaulio linijos ilgį. Osman ir kt. [4] veido taškus panaudojo sudaryti 6 trikampiams (1 pav.). Jų sukurtas algoritmas veidus atpažįsta nepriklausomai nuo amžiaus. Naudodami FG-NET veidų bazę,



1 pav. Veidų atpažinimo algoritmas veido taškams sudaryti naudojantis 6 trikampių [4].

kurioje kiekvieno unikalaus žmogaus nuotraukos yra įvairaus amžiaus, pasiekė 94,66% atpažinimo tikslumą.

Cox ir kt. [2] veidų atpažinimui panaudojo 30 įvairių atstumų tarp veido dalių ir pasinaudoję nuotraukomis iš įvairių duomenų bazių, pasiekė 95% atpažinimo rodiklį.

### 3 Siūlomas algoritmas

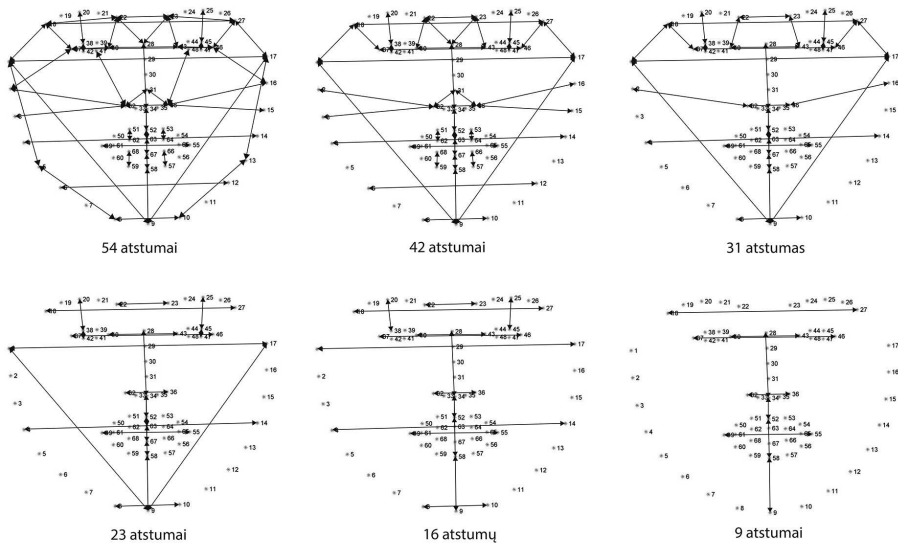
Išanalizavus daugelį įvairių mokslininkų sukurtų algoritmų, buvo nustatyta, jog dauguma jų, prieš išskiriant veido požymius, atlieka nuotraukos normalizavimą. Tam reikalingi papildomi veiksmai – geometrinių iškraipymų šalinimas, išmatavimų mažinimas, dažnai vertimas į nespaltotą formatą. Šiame darbe yra sukurtas algoritmas, kuris veido požymius išgauna iš neapdorotų paveikslėlių, kurie gali būti ir skirtingų išmatavimų.

Kadangi kiekvienas veidas turi unikalias ir nekintančias geometrines formas, vienas paprastesnių, tačiau efektyvių metodų yra matuoti atstumus tarp tam tikrų veido dalių ir juos lyginti. Siekiant juos išmatuoti, reikalinga išskirti būdingus veido taškus (angl. *Landmarks*). Jie dažniausiai naudojami veidams aptikti ir normalizuoti, tačiau yra sukurti algoritmai, kuriuose taškai naudojami ir veido požymių išgavimui.

Veido ir taškų aptikimui buvo panaudotas jau sukurtas Dlib (<http://dlib.net>) bibliotekoje esantis metodas, aptinkantis 68 būdingus veido taškus, kurie apima akių, lūpų, nosies, antakių ir pačio veido kontūrus. Šis metodas sukurtas remiantis Kazemi ir Sullivan aprašytu tyrimu [3], kuriame jie nagrinėjo regresinių medžių panaudojimą veido taškų aptikimui.

Atlikus veido struktūros analizę, buvo nuspręsta sudaryti keletą įvairių atstumų rinkinių tarp įvairių veido dalių – 9, 16, 21, 33, 42 ir 54. 2 paveiksle grafiškai pavaizduoti kiekvieno rinkinio atstumai tarp veido dalių.

Tarp pasirinktų taškų buvo apskaičiuoti Euklidiniai atstumai. Kad atstumų reikšmės būtų ne absoliučios, o santykinės, jas reikia normalizuoti, t. y. pervesti į intervalą  $[0, 1]$ . Santykinės reikšmės suteikia galimybę naudoti neapdorotas skirtingo dydžio nuotraukas.



2 pav. Atstumų rinkiniai tarp įvairių veido dalių.

Iš vieno veido gaunamas vektorius, kuris perduodamas dirbtiniam neuroniniam tinklui, kuris turi būti išmokomas atpažinti (klasifikuoti) veidus. Kaip klasifikatorius buvo pasirinktas daugiasluksnis perceptronas, kuris apmokomas klaidos skleidimo atgal algoritmu. Šis DNT tipas pasirinktas dėl gero efektyvumo ir gana nesudėtingos struktūros.

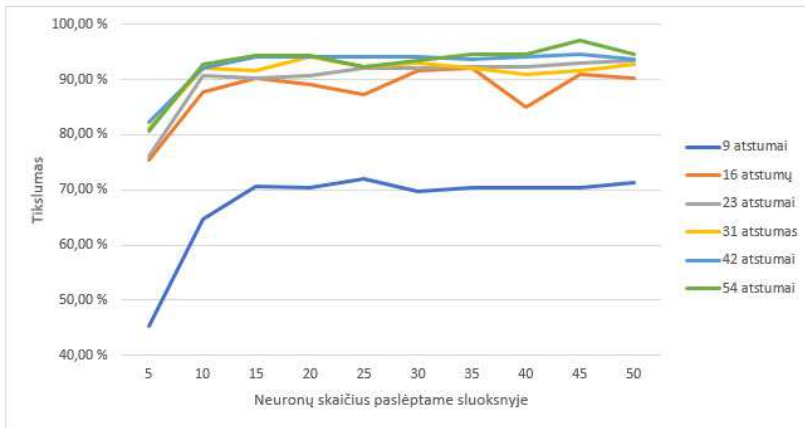
## 4 Eksperimento eiga ir rezultatai

Algoritmo testavimas vyko naudojant tris veidų duomenų bazes – Faces96 (<http://cswww.essex.ac.uk/mv/allfaces/faces96.html>), Grimace (<http://cswww.essex.ac.uk/mv/allfaces/grimace.html>) ir Caltech (<http://www.vision.caltech.edu/html-files/archive.html>). Su pirmąja buvo atliekama pagrindinė eksperimento dalis. Faces96 pasirinkta dėl to, kad joje esantys veidai išgauti iš vaizdo įrašo, kuris darytas prie prasto apšvietimo ir gana nekokybiški, tuo tarpu Caltech – kokybiškų nuotraukų rinkinys. Grimace veidai sukurti kaip ir Faces96 – iš vaizdo kameros, naudojant dar prastesnį apšvietimą, be to, žmonės juose užfiksuoti keičiantys veido išraiškas, tad atpažinimui tai rimtas iššūkis.

Eksperimente buvo siekiama nustatyti, kiek veido atstumų reikalinga efektyviam veidų atpažinimui ir kokie dirbtinio neuroninio tinklo parametrai yra optimalūs.

Pradiniai dirbtinio neuroninio tinklo parametrai buvo parinkti remiantis rekomendacijomis kituose tyrimuose bei atlikus preliminarinius bandymus. Buvo naudojamas stochastinio gradientinio nusileidimo optimizavimo algoritmas su *Momentum* konstanta (toliau – *Momentum* algoritmas), kuri yra pridėjama prieš kiekvieną svorių keitimą.

Paklaidos skaičiavimą atliko kryžminės entropijos paklaidos funkcija binarinei klasifikacijai (angl. *Binary crossentropy*). Paslėptame sluoksnyje naudota sigmoidinė aktyvacijos funkcija.



3 pav. Skirtingų atstumų rinkinių atpažinimo efektyvumo palyginimas pagal neuronų skaičių.

1 lentelė. Aukščiausią tikslumą parodę rezultatai, pasiekti su skirtingais atstumų rinkiniais.

Atpažinimo tikslumas, %	Neuronai	Trukmė	Epochos	Atstumai
97,00	45	00:01:44	2723	54
94,67	45	00:02:11	3519	42
94,67	35	00:01:34	2585	54
94,67	40	00:01:49	2881	54
94,67	50	00:01:47	2699	54
94,33	15	00:01:54	3529	54
94,33	20	00:01:38	2857	54
94,00	20	00:01:41	2948	42
94,00	25	00:01:50	3103	42
94,00	30	00:01:45	2959	42

Kadangi su kiekvienu atstumų rinkiniu dirbtiniam neuroniniam tinklui nustatomas skirtingas neuronų skaičius įėjimo sluoksnyje, nuo to priklauso ir neuronų skaičius paslėptame sluoksnyje, tad lyginant atstumų rinkinius buvo keičiami ir jie. 3 paveiksle pateikiami palyginimo rezultatai, gauti su Face96 veidų rinkiniu.

Kaip matome, 9 atstumai yra per mažas skaičius efektyviam veidų atpažinimui, tuo tarpu atpažinimo rezultatai su kitais atstumų rinkiniais išsidėstė kelių procentų skirtumu. Geriausias maksimalus atpažinimo rezultatas buvo pasiektas naudojant 54 atstumų rinkinį ir 45 neuronus paslėptame sluoksnyje – 97,00%. 1 lentelėje pateikiami aukščiausią tikslumą parodę rezultatai, pasiekti su skirtingais atstumų rinkiniais. Lentelėje taip pat pateikiamas epochų skaičius, parodantis kiek kartų buvo pakeisti svoriai visiems įėjimo neuronams.

Pearsono koreliacija tarp atstumų skaičiaus ir atpažinimo tikslumo pasireiškė gana stipri – 0,72. Dar ženklesnė koreliacija tarp atstumų skaičiaus ir greitaveikos, siekianti 0,78. Tai rodo, jog esant didesniai atstumų skaičiui, dirbtinis neuroninis tinklas greičiau apsimoko ir pasiekia didesnę atpažinimo tikslumą.

Toliau keičiant įvairius dirbtinio neuroninio tinklo parametrus buvo ieškoma maksimalaus tikslumo. Jis pasiektas naudojant gradientinio nusileidimo optimizavimo

**2 lentelė.** Veidų atpažinimo algoritmų palyginimas naudojant Faces96 veidų duomenų bazę [1].

Algoritmas	Atpažinimo tikslumas, %
CT-FBM + KNN	97,55
PCA + DCT	94,00
PCA + SVM	94,00
Gabor transformacija	92,35
LDA + SVM + RBF kernel function	88,24
LBP	82,94

algoritmą *Adadeltra*, RMSLE (angl. *Root Mean Squared Logarithmic Error*, šaknis iš vidutinės kvadratinės logaritminės paklaidos) funkciją bei aktyvacijos funkciją *Hard-Sigmoid* – 97,33%. Tačiau tam DNT sugaišo net 22 min 3 s. Tokius didelius laiko skirtumus labiausiai įtakoja gradientinio nusileidimo optimizavimo algoritmai.

Paskutinėje eksperimento dalyje buvo atliktas veidų atpažinimo palyginimas naudojant Grimace ir Caltech veidų duomenų bases. Su kokybiškų veidų rinkiniu Caltech buvo pasiektas maksimalus atpažinimo tikslumas – 98,00%. Lyginant su Faces96, tai mažiau nei vienas procentas. Toks rezultatas parodo sukurto algoritmo atsparumą nuotraukų kokybei bei apšvietimui. Pakartojus eksperimentą su skirtingomis išraiškomis nufotografuotų veidų rinkiniu Grimace, buvo pasiektas pakankamai aukštas rezultatas – 94,44%, kuris parodo, jog veido išraiškos gana silpnai įtakoja efektyvumą.

Žvelgiant į neuroninio tinklo mokymo greitaveikos rezultatus, pastebėta tendencija – kuo kokybiškesnės nuotraukos, tuo greičiau tinklas apsimoko.

Lyginant su kitų mokslininkų atliktais tyrimais, šis metodas demonstruoja puikius rezultatus. Al-Waisy ir kt. [1] sukūrė metodą, kuris su Faces96 pasiekė 97,55% tikslumą bei sudarė palyginamąją lentelę (2 lentelė). Iš jos matome, jog šiame straipsnyje aprašytas algoritmas lenkia daugelį kitų ir nenusileidžia geriausiems atpažinimo rezultatus parodžiusiems veidų atpažinimo metodams.

## 5 Išvados

Tyrimo metu sukurtas ir realizuotas veidų atpažinimo algoritmas, kuris iš būdingų veido bruožų taškų išgauna tam tikrus atstumus tarp įvairių veido sričių, o daugiasluoksnis perceptronas atlieka klasifikavimą. Atlikus eksperimentinį tyrimą nustatyta:

1. Daugiasluoksnis perceptronas naudojamas 54 atstumų rinkinį su veidų duomenų baze Faces96 pasiekė aukščiausią atpažinimo rezultatą – 97,33%, tam sugaišdamas 22 min 3 s. Rezultatai buvo pasiekti su vienu paslėptu sluoksniu bei 45 neuronais jame, naudojant stochastinio gradientinio nusileidimo optimizavimo algoritmą *Adadeltra*, RMSLE paklaidos funkciją ir aktyvacijos funkciją *HardSigmoid*.
2. Efektyviausiai klasifikavimo užduotį DNT atliko su *Momentum* algoritmu, kryžminės entropijos paklaidos funkcija daugiaklasei klasifikacijai bei sigmoidine funkcija – atpažinimo tikslumas 97% ir trukmė 1 min 44 s.
3. Veidų atpažinimo algoritmas, naudojantis santykinius atstumus tarp įvairių veido sričių, gali pasiekti atpažinimo rezultatus, kurie prilygsta ar net pralenkia kitus metodus naudojančius algoritmus.

## Literatūra

- [1] A.S. Al-Waisy, R. Qahwaji, S. Ipson and S. Al-Fahdawi. A robust face recognition system based on curvelet and fractal dimension transforms. In *Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM)*, pp. 548–555. IEEE, 2015.
- [2] I.J. Cox, J. Ghosn and P.N. Yianilos. Feature-based face recognition using mixture-distance. In *Proceedings of the 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1996.*, pp. 209–216. IEEE, 1996.
- [3] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1867–1874, 2014.
- [4] Amal Seralkhatem Osman A., Vijanth Sagayan Asirvadam, Aamir Saeed Malik, Mohamed Meselhy Eltoukhy and Azrina Aziz. Age-invariant face recognition using triangle geometric features. *Int. J. Patt. Rec. Art. Int.*, **29**(05):1556006, 2015.
- [5] K.-K. Sung and T. Poggio. Example-based learning for view-based human face detection. *Proc. Image Understanding Workshop*, **2**:843–850, 1994.

### SUMMARY

#### **Multilayer perceptron for face recognition**

*R. Toliušis, O.Kurasova*

In this paper, an algorithm is proposed which uses facial landmarks to calculate normalized Euclidean distances between different facial parts and performs faces recognition by using Multilayer Perceptron. In order to determine the most effective model, different neural network parameters have been changed in an experimental way, such as hidden layers and the number of neurons, gradient descent optimization algorithms, error and activation functions, and different sets of distances.

*Keywords:* artificial neural networks, multilayer perceptron, face recognition.