

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
INFORMATIKOS KATEDRA

3D objektų atpažinimas iš 2D nuotraukų

3D object recognition from 2D images

Magistro baigiamasis darbas

Atliko:	Aleksas Vaitulevičius	(parašas)
Darbo vadovas:	prof. habil. dr. Olga Kurasova	(parašas)
Recenzentas:	doc. dr. Vardauskas Pavardauskas	(parašas)

Vilnius – 2020

Santrauka

Glaustai aprašomas darbo turinys, pristatoma nagrinėta problema ir padarytos išvados. Santraukos apimtis ne didesnė nei 0,5 puslapio. Santraukų gale nurodomi darbo raktiniai žodžiai.

Raktiniai žodžiai: raktinis žodis 1, raktinis žodis 2, raktinis žodis 3, raktinis žodis 4, raktinis žodis 5

Summary

Santrauka anglų kalba.

Keywords: keyword 1, keyword 2, keyword 3, keyword 4, keyword 5

Turinys

Įvadas	4
1. Medžiagos darbo tema dėstymo skyriai	6
1.1. Poskyris.....	6
1.1.1. Skirsnis	6
1.1.1.1. Straipsnis.....	6
1.1.2. Skirsnis	6
2. Skyrius	7
2.1. Poskyris.....	7
2.2. Poskyris.....	7
Rezultatai ir išvados	8
Literatūra	9
Santrumpos	10
Priedas Nr.1	
Priedas Nr.2	

Ivadas

Vienas iš fundamentalių kompiuterinės regos uždavinių yra informacijos apie 3 dimensijų (3D) pasaulį išgavimas naudojant 2 dimensijų (2D) nuotraukas. Šio uždavinio tikslas yra atpažinti konkrečius 3D objektus, naudojant jų, skirtingų apžvalgos taškų 2D nuotraukas. Šiam tikslui pasiekti, yra konstruojami objektų atpažinimo algoritmai, kurie klasifikuoja 2D nuotraukas į klases, kurios atstovauja vieną iš 3D objektų modelių.

3D objektų atpažinimas iš 2D nuotraukų yra naudojamas srityse, kuriose turimi 3D objektai turi būti atpažinti iš visų galimų 2D nuotraukų, turint tik poaibį šių nuotraukų. Keli šių sričių pavyzdžiai yra vogtų objektų aptikimas - turint algoritmą, atpažįstantį konkretų automobilį, galima iš viešų erdvių nuotraukų atrasti automobilio poziciją, vietos nustatymas - turint algoritmą, atpažįstantį konkrečius objektus esančius skirtingose vietovėse, ir tų vietovių koordinatės galima nustatyti kurioje vietovėje buvo padaryta nuotrauka. Deja, laiko ir duomenų kaštai yra per dideli, kad pasiekti absoliutų tikslumą sprendžiant šį uždavinį. Todėl taikomi metodai yra euristiniai. Dėl to renkantis metodą, spręsti 3D objektų atpažinimo iš 2D nuotraukų uždaviniui, reikia atsižvelgti į laiko kaštus ir kaip tiksliai tuo metodu pagrįstas algoritmas klasifikuoja 2D nuotraukas, sprendžiamas šį uždavinį. Šiame darbe, bus atliekami tyrimai, skirti nustatyti metodą, kuris būtų pranašesnis sprendžiant 3D objektų atpažinimo iš 2D nuotraukų uždavinį. Pasirinkti metodai yra konvoliucinio ir CapsNet neuroninio tinklo architektūros.

Gana dažnai naudojamas metodas, šiam uždaviniui spręsti, yra dirbtiniai gilieji neuroniniai tinklai. 3D objektų atpažinimo iš 2D nuotraukų uždavinyje, naudojami duomenys - 2D nuotraukos, yra nestruktūrizuoti, jiems sudėtinga vykdyti požymių išgavimą. Todėl daugelis kitų sprendimų nėra tokie patrauklūs kaip dirbtiniai gilieji neuroniniai tinklai, dėl savo sugebėjimo efektyviai vykdyti automatinį požymių išgavimą iš nestruktūrizuotų duomenų. Tačiau, norint pasiekti aukštą klasifikavimo tikslumą, naudojant šį metodą, yra reikalingas didelis kiekis duomenų. Konkrečiai šiam uždaviniui reikia didelio kiekio 3D modelių. Laimei, šiuo metu egzistuoja viešai prieinamos didelės 3D repozitorijos. Tokios kaip 3D Warehouse, TurboSquid, ir Shapeways. Dėl to, šiuo metu daugelis architektūrų jau yra išbandytos sprendžiant 3D objektų atpažinimo iš 2D nuotraukų uždavinį. Pavyzdžiui viena iš architektūrų, kuri buvo išbandyta, yra gilaus pasitikėjimo architektūra. Šiai architektūrai atlikti tyrimai yra aprašyti darbe [WSK⁺15].

Šiuo metu, šiam uždaviniui spręsti, optimaliausius rezultatus, laiko ir tikslumo atžvilgiu, pasiekusi dirbtinio neuroninio tinklo architektūra yra konvoliuciniai neuroniniai tinklai. Tyrimai, kuriuose ši architektūra buvo išbandyta, yra aprašyta darbuose [Mas18; SMK⁺15]. Darbe [WSK⁺15] atlikto tyrimo rezultatai parodo, kad pateiktas sprendimas, kuriame 3D objektų atpažinimas yra konstruojamas naudojantis tik 2D nuotraukomis, yra tikslesnis 8% (77% → 85%). Todėl šiame darbe bus daromi eksperimentai su dirbtinio neuroninio tinklo architektūrų įgyvendinimais, kurie yra pagrįsti būtent šiuo metodu. Tad šiam darbui, vienas iš pasirinktų metodų yra konvoliuciniai neuroniniai tinklai, kurio įgyvendinimas ir tyrimai yra aprašyti darbe [SMK⁺15]. Mat šio darbo įgyvendinimas naudoja tik 2D nuotraukas, konstruojant 3D objektų atpažinimo algoritmą.

Kita tiriama dirbtinio neuroninio tinklo architektūra yra CapsNet. Lyginant su konvoliuciniu neuroniniu tinklu, tai gana nauja architektūra. Aprašyta 2017 metais darbe [nNH17] CapsNet ar-

chitektūros veikimo principas tiksliau imituoja žmogaus rega, remiantis faktu, kad žmogaus rega ignoruoja nereikšmingas vaizdo detales, naudodama tik seką fokusuotų taškų, taip apdorodama tik dalį vaizdinės informacijos su labai aukšta rezoliucija. Šiame darbe atliktas tyrimas parodo, kad ši architektūra atlieka skaičių klasifikavimo užduotį tiksliau nei konvoliucinis neuroninis tinklas. Tad

1. Medžiagos darbo tema dėstymo skyriai

Medžiagos darbo tema dėstymo skyriuose pateikiamos nagrinėjamos temos detalės: pradiniai duomenys, analizės ir apdorojimo metodai, sprendimų įgyvendinimas, gautų rezultatų apibendrinimas.

1.1. Poskyris

Citavimo pavyzdžiai: cituojamas vienas šaltinis [**PvzStraipsnLt**]; cituojami keli šaltiniai [**PvzStraipsnEn; PvzKonfLt; PvzKonfEn; PvzKnygLt; PvzKnygEn; PvzElPubLt; PvzElPubEn; PvzMagistrLt; PvzPhdEn**].

1.1.1. Skirsnis

1.1.1.1. Straipsnis

1.1.2. Skirsnis

2. Skyrius

2.1. Poskyris

2.2. Poskyris

Rezultatai ir išvados

Rezultatų ir išvadų dalyje išdėstomi pagrindiniai darbo rezultatai (kažkas išanalizuota, kažkas sukurta, kažkas įdiegta), pateikiamos išvados (daromi nagrinėtų problemų sprendimo metodų palyginimai, siūlomos rekomendacijos, akcentuojamos naujovės).

Literatūra

- [Mas18] F. Vitor Suzano Massa. Relating images and 3d models with convolutional neural networks. <https://pastel.archives-ouvertes.fr/tel-01762533/document>, 2018. 7.917 KB, accessed 2020-05-23.
- [nNH17] S. Sabour nad N. Frosst and G. E. Hinton. Dynamic routing between capsules. <https://arxiv.org/pdf/1710.09829.pdf>, 2017. 899 KB, accessed 2020-05-23.
- [SMK⁺15] H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller. Multi-view convolutional neural networks for 3d shape recognition. https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Su_Multi-View_Convolutional_Neural_ICCV_2015_paper.pdf, 2015. 1.439 KB, accessed 2020-05-23.
- [WSK⁺15] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao. 3d shapenets: a deep representation for volumetric shapes. https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Wu_3D_ShapeNets_A_2015_CVPR_paper.pdf, 2015. 2.451 KB, accessed 2020-05-23.

Santrumpos

Sąvokų apibrėžimai ir santrumpų sąrašas sudaromas tada, kai darbo tekste vartojami specialūs terminai, reikalaujantys paaiškinimo, ir rečiau sutinkamos santrumpos.

Priedas Nr. 1

Niauroninio tinklo struktūra

Priedas Nr. 2

Eksperimentinio palyginimo rezultatai

1 lentelė. Lentelės pavyzdys

Algoritmas	\bar{x}	σ^2
Algoritmas A	1.6335	0.5584
Algoritmas B	1.7395	0.5647