



Kauno technologijos universitetas

Elektros ir elektronikos fakultetas

**Odos defektų išskyrimo ir atpažinimo iš nuotraukų tyrimas,
taikant gilaus mokymo modelius**

Baigiamasis magistro projektas

Paulius Kirkus

Projekto autorius

doc. dr. Arūnas Lipnickas

Vadovas

Kaunas, 2021



Kauno technologijos universitetas

Elektros ir elektronikos fakultetas

Odos defektų išskyrimo ir atpažinimo iš nuotraukų tyrimas, taikant gilaus mokymo modelius

Baigiamasis magistro projektas

Valdymo technologijos (6211EX014)

Paulius Kirkus

Projekto autorius

doc. dr. Arūnas Lipnickas

Vadovas

doc. dr. Tomas Tekorius

Recenzentas

Kaunas, 2021



Kauno technologijos universitetas

Elektros ir elektronikos fakultetas

Paulius Kirkus

Odos defektų išskyrimo ir atpažinimo iš nuotraukų tyrimas, taikant gilaus mokymo modelius

Akademinio sąžiningumo deklaracija

Patvirtinu, kad:

1. baigiamąjį projektą parengiau savarankiškai ir sąžiningai, nepažeisdama(s) kitų asmenų autoriaus ar kitų teisių, laikydamasi(s) Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo nuostatų, Kauno technologijos universiteto (toliau – Universitetas) intelektinės nuosavybės valdymo ir perdavimo nuostatų bei Universiteto akademinės etikos kodekse nustatytų etikos reikalavimų;
2. baigiamajame projekte visi pateikti duomenys ir tyrimų rezultatai yra teisingi ir gauti teisėtai, nei viena šio projekto dalis nėra plagijuota nuo jokių spausdintinių ar elektroninių šaltinių, visos baigiamojo projekto tekste pateiktos citatos ir nuorodos yra nurodytos literatūros sąrašė;
3. įstatymų nenumatytų piniginių sumų už baigiamąjį projektą ar jo dalis niekam nesu mokėjęs (-usi);
4. suprantu, kad išaiškėjus nesąžiningumo ar kitų asmenų teisių pažeidimo faktui, man bus taikomos akademinės nuobaudos pagal Universitete galiojančią tvarką ir būsiu pašalinta(s) iš Universiteto, o baigiamasis projektas gali būti pateiktas Akademinės etikos ir procedūrų kontrolieriaus tarnybai nagrinėjant galimą akademinės etikos pažeidimą.

Paulius Kirkus

Patvirtinta elektroniniu būdu

Kirkus Paulius. Odos defektų išskyrimo ir atpažinimo iš nuotraukų tyrimas, taikant gilaus mokymo modelius. Magistro baigiamasis projektas vadovas doc. dr. Arūnas Lipnickas; Kauno technologijos universitetas, Elektros ir Elektronikos fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Elektronikos inžinerija. Inžinerijos mokslai.

Reikšminiai žodžiai: gilieji neuroniniai tinklai, melanoma, vaizdų analizė.

Kaunas, 2021. 50 p.

Santrauka

Šiame tiriamajame darbe atliekamas gilaus mokymo modelių tikslumo tyrimas. Analizuojamas objektas: odos vėžio dariniai – melanoma. Naudojamas vaizdo atpažinimas, segmentacija. Įvesties duomenys, tai medicininės paskirties nuotraukos. Nuotraukos paimtos iš viešai prieinamos, vienos didžiausių odos darinių duomenų bazės „ISIC18“.

Tiriamajame darbe atliekama teorinė vaizdų modelių analizė, apžvelgiami su šia tema susiję anksčiau atlikti darbai bei jų gauti rezultatai. Po teorinės dalies apžvalgos pradedamas „U-net“ modelio tikslumo tyrimas bei tobulinimas. Pasirenkama pradinė „U-net“ modelio architektūra, kuri darbo eigoje yra papildoma tikslumą gerinančiais blokais, keičiama tinklo iškodavimo dalies struktūra.

Naudojamos 3 skirtingos modelio konfigūracijos, su kuriomis atliekamas modelio struktūros tikslumo tyrimas. Mokymo proceso trukmė 50 iteracijų. Naudojami 1815 vaizdų mokymui, testavimo imtį sudaro 520 vaizdų, patvirtinimo – 259. Aukščiausias pasiektas rezultatas pagal „F1“ vertinimo kriterijų yra 89,21 procento. Panaudojus vieną geriausių rezultatų pasiekusią struktūrą išbandoma daugiaklasė segmentacija, kurios metu atliekamas klasifikavimas – iš 1 vaizdo išskiriamos 5 darinių klasės.

Kirkus Paulius. Research of Skin defect Segmentation and Classification from images by deep learning models. Master's Final Degree Project supervisor doc. dr. Arūnas Lipnickas; Faculty of Electrical and Electronics Engineering, Kaunas University of Technology.

Study field and area (study field group): Electronics Engineering. Engineering science.

Keywords: convolutional neural network, melanoma, image analysis.

Kaunas, 2021. 50.

Summary

In this research, the accuracy of deep learning models is investigated. The object of analysis is a derivative of skin cancer - melanoma. Image recognition, segmentation is used. Input data are medical photos. The photos are taken from the one of the largest publicly available database – ISIC18.

In the research work theoretical analysis of deep learning models is performed. Related works on this topic and their results are reviewed. After reviewing the theoretical part, the research and improvement of the accuracy of the U-net model begins. The original architecture of the U-net model is chosen, which is supplemented in the workflow with blocks that improve accuracy, also the structure of the network decoding part is changed.

3 different model configurations are used to know which structure of the model leads to the highest accuracy. Duration of the deep learning process 50 iterations. 1815 images used for training, the test sample consists of 520 images, validation - 259. The highest achieved result according to the F1 evaluation criterion is 89,21 percent. Using one of the best-performing structures, multiclass segmentation is tested, during which classification is performed - 5 classes of derivatives are distinguished from 1 image.

Turinys

Paveikslų sąrašas	8
Santrumpų ir terminų sąrašas	9
Įvadas.....	10
1. Teorinė dalis.....	12
1.1. Dirbtiniai neuroniniai tinklai	12
1.1.1. Konvoliucinis neuroninis tinklas	12
1.1.2. Gilusis mokymasis.....	13
1.1.3. Konvoliucinių tinklų veikimo principas	13
1.2. „U-net“ modelis.....	16
1.2.1. „U-net“ modelio architektūra	17
1.2.2. „U-net“ išėjimo sluoksniai mokomi su pastiprintu mokymu	18
1.3. Požymių pastiprinimo blokas	19
1.3.1. Erdvinio spaudimo būdu požymius pastiprinantis blokas (spaudžiama per geometriją)	20
1.3.2. Erdvinio spaudimo ir kanalų sužadavimo būdu požymius pastiprinantis blokas (bruožų žemėlapių spaudimas per kanalus)	21
1.3.3. Erdvinio / kanalų suspaudimo ir sužadavimo būdu požymius pastiprinantis blokas (tarpkanalinis spaudimas).....	21
1.3.4. Požymių pastiprinimo blokas, kanalų svarbos pabrėžimas („Adaptive Average Pooling 2D“ spaudimas per kanalus).....	22
1.4. Daugiaklasė segmentacija	22
1.5. Nuostolių funkcija	23
1.6. Vertinimo kriterijus	23
2. Panašių darbų apžvalga.....	25
2.1. Geriausių „ISIC18“ kورkurse pasirodžiusių neuroninių tinklų modelių rezultatai	25
2.1.1. „R2U-Net“ modelis	25
2.1.2. „BCDU-net“ modelis.....	26
2.1.3. „MCGU-Net“ modelis.....	27
3. „ISIC18“ duomenų bazė	28
4. Tiriamoji dalis.....	30
4.1. Duombazės paruošimas	30
4.2. Pradinio „U-net“ modelio parametrų parinkimas.....	30
4.3. „U-net“ modelio modernizavimas	31
4.4. „U-net“ išėjimų sustiprinimas	37
4.4.1. Pradinis „U-net“ su sustiprinta išėjimų išvestimi.....	37
4.4.2. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir erdvinio požymių segmentų atmetimo bloku.....	38
4.4.3. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir požymių pastiprinimo blokais	39
4.4.4. Modelių, naudojančių „U-net“ su giliuoju mokymu iškodavimo dalyje palyginimas	42
4.5. Daugiaklasė segmentacija	43
Išvados	46
Literatūros sąrašas	48

Lentelių sąrašas

3.1 lentelė. Geriausių modelių rezultatai, naudota duomenų bazė „ISIC18“	29
4.1 lentelė. „U-net“ modelio rezultatai pagal „F1“ vertimo kriterijų, kai įėjimo vaizdo dydis 128 x 128 x 3	30
4.2 lentelė. „U-net“ modelio rezultatai pagal „F1“ vertimo kriterijų, kai įėjimo vaizdo dydis 256 x 256 x 3	31
4.3 lentelė. Pradinio „U-net“ modelio tikslumo rezultatai	31
4.4 lentelė. „U-net“ skaičiavimų rezultatai įtraukus erdvinius požymių segmentų atmetimo blokus	32
4.5 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE1“, suspaudžiantį broožų žemėlapius	33
4.6 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE2“, spaudimas vyksta per geometriją	34
4.7 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE3“, spaudimą per kanalus	34
4.8 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE4“, spaudimas per kanalus (papildytas)	35
4.9 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE5“ skaičiavimams	35
4.10 lentelė. Modelių palyginimas („128_4_32“ tinklo konfigūracija)	36
4.11 lentelė. Modelių palyginimas („256_5_8“ tinklo konfigūracija)	36
4.12 lentelė. Modelių palyginimas („256_5_32“ tinklo konfigūracija)	37
4.13 lentelė. Pradinio „U-net“ tinklo su daugiasluoksniu išėjimu rezultatai	38
4.14 lentelė. Pradinio „U-net“ tinklo su daugiasluoksniu išėjimu ir „SpatialDropout“ bloku rezultatai	39
4.15 lentelė. Pradinio „U-net“ tinklo su daugiasluoksniu išėjimu ir „SpatialDropout“, požymių pastiprinimo blokais rezultatai	39
4.16 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 1 požymių pastiprinimo bloku „SE1“	40
4.17 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 2 požymių pastiprinimo bloku „SE2“	40
4.18 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 3 požymių pastiprinimo bloku „SE3“	40
4.19 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 4 požymių pastiprinimo bloku „SE4“	41
4.20 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 5 požymių pastiprinimo bloku „SE5“	41
4.21 lentelė. Modelių palyginimas („128_4_32“ tinklo konfigūracija)	42
4.22 lentelė. Modelių palyginimas („256_5_8“ tinklo konfigūracija)	43
4.23 lentelė. Modelių palyginimas („256_5_32“ tinklo konfigūracija)	43

Paveikslų sąrašas

1.1 pav. Dirbtinio neuroninio tinklo struktūra [3].....	12
1.2 pav. Konvoliucinis sluoksnis [6].....	13
1.3 pav. Bruožų žemėlapių sandara [6].....	14
1.4 pav. Bruožų žemėlapis [5].....	14
1.5 pav. Filtras [6]	15
1.6 pav. Sutelkimo sluoksnių pavyzdžiai [7]	15
1.7 pav. Visiškai sujungtas sluoksnis [6]	16
1.8 pav. Viso konvoliucinio neuroninio tinklo architektūra [6].....	16
1.9 pav. „U-net“ modelio architektūra. Kiekviena mėlyna dėžutė atspindi daugiakanalius bruožų žemėlapius. Kanalų skaičius nurodytas virš kiekvienos dėžutės. „X“ ir „Y“ – dydis nurodytas kiekvienos dėžutės kairėje pusėje. Baltos dėžutės atspinti nukopijuotus bruožų žemėlapius. Legendoje matomos rodyklės apibrėžia skirtingas operacijas [9].....	17
1.10 pav. Kodavimo – iškodavimo modelis [7]	18
1.11 pav. „U-net“ tinklas su daugiasluoksniu išėjimu [12].....	19
1.12 pav. Požymių pastiprinimo bloko struktūra [11].....	19
1.13 pav. Požymių pastiprinimo bloko pozicijos [13]	20
1.14 pav. Erdvinio spaudimo būdu požymius pastiprinantis blokas [13]	21
1.15 pav. Erdvinio spaudimo ir kanalų sužadavimo būdu požymius pastiprinantis blokas [13].....	21
1.16 pav. Erdvinio / kanalų suspaudimo ir sužadavimo būdu požymius pastiprinantis blokas [13].	22
1.17 pav. Skirtingų segmentacijos tipų pavyzdys. A – vienos klasės išskyrimas; B – rezultatas gaunamas, kai neišskiriama jokia klasė; C – rezultatas, kai vaizde išskiriama keletas klasių. Žalia spalva reiškia melanocitinę nevusą, raudona – melanomą, o geltona – seborėjinę keratozę [16].	23
2.1 pav. „R2U-net“ architektūra su konvoliuciniais kodavimo / iškodavimo keliais, kuriuose naudojami pasikartojantys konvoliuciniai sluoksniai. Tinklas sukurtas pagal „U-net“ modelį [24].	25
2.2 pav. „BCDU-net“ modelis. Naudojamas dvikryptis „ConvLSTM“ modulis praleidimo jungtyse (angl. <i>skip connections</i>), tankiai sujungtos konvoliucijos [28].	26
2.3 pav. „MCGU-net“ modelis. Modelis naudoja „BCDU-net“ struktūrą, iškodavimo sluoksnyje papildytą požymių pastiprinimo blokais [29].....	27
3.1 pav. „ISIC18“ duombazę, sudarančių nuotraukų pavyzdžiai. Viršuje matoma melanomos segmentacijai skirta duombazės dalis [31], apačioje pateikiamas visų 7 klasių esančių duombazėje vaizdas [32].....	28
4.1 pav. Požymių pastiprinimo blokų įterpimo vieta „U-net“ modelyje.....	33
4.2 pav. „U-net“ išėjimo sustiprinimas	37
4.3 pav. Daugiaklasei segmentacijai naudojami duomenys (mokymui).....	44
4.4 pav. Daugiaklasės segmentacijos rezultatai (1 pvz.).....	44
4.5 pav. Daugiaklasės segmentacijos rezultatai (2 pvz.).....	45
4.6 pav. Daugiaklasės segmentacijos rezultatai (3 pvz.).....	45

Santrumpų ir terminų sąrašas

Santrumpos:

SE – požymių pastiprinimo blokas.

ReLU (angl. *Rectified Linear Unit*) – aktyvavimo funkcija, apibrėžta kaip teigiama jos argumento dalis: kur x yra neurono įvestis.

RGB – (raudona – žalia – mėlyna) spalvų modelis.

CNN (angl. *convolutional neural network*) – konvoliucinis neuroninis tinklas.

F1 – tikslumo matas.

Terminai:

Konvoliucija – tai matematinis operatorius, kuris kaip argumentus paima dvi funkcijas „f“ ir „g“ ir gražina trečią, kuri parodo „f“ ir „g“ persidengimo kiekį.

Įvadas

Atpažinimas bendrąja prasme yra procesas, kurio metu nustatomas informacinis situacijos turinys ir atliekama semantinė klasifikacija. Vykstant šiam procesui iš jutiklių (arba jutimo organų) nustatoma ir įsimenama informacija, bruožai, kurie palyginami su atmintyje saugomais ankstesniais to pačio objekto duomenimis ir jie priskiriami tam tikrai semantinei klasei (kategorijai). Automatizuotas vaizdų atpažinimas, tai tas pats procesas, tik visą darbą atlieka kompiuterinės sistemos. Sparčiai tobulėjant technologijoms, kartu išaugo ir vaizdų atpažinimo panaudojimo galimybės. Automatinės vaizdo apdorojimo sistemos šiuo metu naudojamos buityje, pramonėje, pramogų srityse ar medicinoje.

Išskiriama keletas vaizdų apdorojimo sričių. Pavyzdžiui, vaizdų klasifikacija, siekiant juos suskirsti į tam tikras klases pagal specifinius požymius. Kita plačiai naudojama sritis, tai vaizdų segmentacija, kurios paskirtis išskirti tam tikras struktūras ar fragmentus iš matomo vaizdo ar nuotraukos. Gamybos ir pramonės srityse segmentacijos sistemas galima pamatyti kokybės kontrolės punktuose, kur taikant vaizdo analizės principus ieškoma defektų ar struktūrų pakitimų. Ši technologija leidžia atlikti tą patį darbą šimtus kartų greičiau nei tai padarytų žmogus ekspertas, tikrindamas gaminius rankinius būdu [4].

Ištobulėjus kameroms, sparčiai daugėja vaizdų, nuotraukų, kurios yra aukštos kokybės. Iš tokių vaizdų kuriamos duomenų bazės. Kokybiškų duomenų bazių gausėjimas, tai viena iš priežasčių, kodėl atsiranda vis tikslesnių vaizdo analizės sistemų, kurios nuolatos tobulinamos. Šiuo metu daug dėmesio sulaukianti tyrimų sritis yra medicina. Vaizdų atpažinimo metodų naudojimas medicinoje, tai neinvazinis struktūrų ar jų kontūrų tyrimas naudojant giliojo mokymo modelius. Šių tyrimų tikslas – išskirti iš vaizdų ar nuotraukų matomus pakitimus ar nestandartines struktūras. Nestandartinės odos struktūros dažnu atveju – vėžiniai dariniai. Vėžys, tai viena iš pavojingiausių ligų, kurios ankstyvas diagnozavimas gali padėti išvengti pavojingesnio ligos išplitimo. Viena iš labiausiai paplitusių vėžio formų, tai odos vėžys. Pavojingiausias odos vėžio tipas yra melanoma. Ją nustatyti taikomi vizualiniai diagnostikos būdai pagal tam tikrus kriterijus – dermatoskopija. Melanos išvaizda gali būti įvairi. Bruožai apibūdinantys ligą yra [2]:

- tamsiai rudi arba juodos spalvos odos dariniai;
- asimetrija;
- dantyti, rantuoti, netaisyklingi ir neryškūs kontūrai;
- daugiau nei 6 milimetrų dydžio navikai (besiplečiantys);
- kintantis dydis.

Šiame darbe atliekamas dirbtinio neuroninio tinklo modelio struktūros tyrimas, kurio tikslas yra išgauti kuo paprastesnį modelį, kuris leistų pasiekti aukštą vaizdo segmento (melanos) nustatymo bei klasifikacijos galimybę.

Darbo tikslas: atlikti odos defektų išskyrimo ir atpažinimo iš nuotraukų tyrimą, taikant gilaus mokymo modelius.

Darbo uždaviniai:

1. nustatyti gilaus mokymo modelio tinkamiausius parametrus segmentacijos užduočiai atlikti;
2. atlikti melanos segmentacijos tikslumo tyrimą, į gilaus mokymo modelį įtraukus erdvinį požymių atmetimo modulį bei 5 skirtingus požymių pastiprinimo blokus;

3. atlikti melanomos segmentacijos tikslumo tyrimą, modernizuojant gilaus mokymo modelį, jį papildant blokais;
4. atlikti melanomos vaizdo klasifikaciją, jį išskaidant į skirtingas klases.

1. Teorinė dalis

1.1. Dirbtiniai neuroniniai tinklai

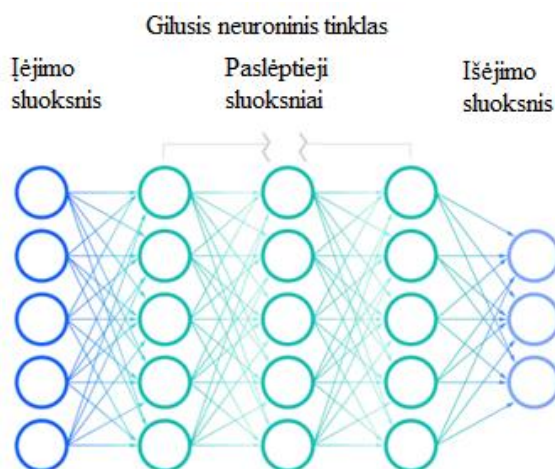
Dirbtinis neuroninis tinklas yra skaičiavimo sistemos dalis sukurta imituoti žmogaus smegenų darbą. Tai dirbtinio intelekto pagrindas, kuris sprendžia problemas, kurios pagal žmogaus standartus pasirodytų neįmanomos ar sunkios. Dirbtinio neuroninio tinklo savarankiško mokymosi galimybės leidžia pasiekti aukštesnių rezultatų, kai naudojamas didelis įvesties duomenų kiekis. Įvesties vienetai, tai įvairios informacijos formos ar struktūros, pagrįstos vidine svorių sistema, kurią neuroninis tinklas mokosi suprasti, kad būtų parengta viena išvesties forma. Išskiriama keletas pagrindinių nuostatų [1]:

- dirbtinis neuroninis tinklas, tai dirbtinio intelekto komponentas, imituojantis smegenų darbą;
- tinkle aprašomi neuronų mazgai sudaro dirbtinį neuroninį tinklą, kurie susideda iš įvesties ir išvesties duomenų;
- klaidos skleidimas atgal (angl. *backpropation*) yra mokymosi taisyklių rinkinys, naudojamas dirbtiniams neuroniniams tinklams apmokyti;
- praktiškas dirbtinio neuroninio tinklo taikymas yra labai platus, apimantis asmeninį gyvenimą, pramonę, švietimą, mediciną;
- dirbtiniai neuroniniai tinklai sukurti atsižvelgiant į žmogaus smegenų struktūrą, neuronų mazgai sujungti kaip tinklas.

1.1.1. Konvoliucinis neuroninis tinklas

Dirbtiniai neuroniniai tinklai yra mašininio mokymosi pogrupis, giliųjų mokymosi algoritmų esmė. Jie sudaryti iš mazgų, kurie sukuria įvesties / išvesties sluoksnius bei vieną ar daugiau paslėptųjų sluoksnių (žr. 1.1 pav.). Kiekvienas mazgas ar dirbtinis neuronas, kuris jungiasi prie kito, turi susijusį svorį bei slenktį. Kai atskiro mazgo išvestis viršija nurodytą ribinę vertę, tas mazgas suaktyvinamas siunčiant duomenis į kitą tinklo sluoksnį [3].

Neuroninių tinklų darbas pagrįstas treniruočių metu gautais duomenimis, kurių metu mažinama rezultatų išvesties paklaida. Tinkamai paruošti algoritmai leidžia greitai ir tiksliai kaupti bei klasifikuoti norimus duomenis. Vienas iš labiausiai žinomų neuroninių tinklų modelių yra „Google“ paieškos algoritmas.



1.1 pav. Dirbtinio neuroninio tinklo struktūra [3]

1.1.2. Gilusis mokymasis

Neuroniniai tinklai sudaryti iš keleto skirtingų sluoksnių. Kiekviename iš jų giliojo mokymosi algoritmai atlieka skaičiavimus ir pakartoja prognozes, palaipsniui gerindami rezultato tikslumą. Mašininiai neuroniniai tinklai dirba su struktūrizuotais, paženklintais duomenimis ir iš jų mokosi. Giliojo mokymosi metu galima apdoroti ir nestructūrizuotus, nesužymėtus duomenis, kitaip sakant, nebūtinai etiketės norint identifikuoti objektus ar informaciją. Gilusis mokymasis naudoja daugiasluoksnį neuroninį tinklą, kad išgautų ypatybes iš duomenų ir vis geriau atpažintų ar klasifikuotų duomenis [4].

Šie modeliai išsiskiria tuo, kad gali nustatyti anksčiau nepastebėtas duomenų savybes ar struktūras, kurios nėra paženklintos etikete, laikantis mažiausio žmogaus priežiūros lygio. Giliojo mokymosi modeliai taip pat geba sustiprinti mokymąsi remdamiesi teigiamais ankstesnių skaičiavimų rezultatais [5].

1.1.3. Konvoliucinių tinklų veikimo principas

Konvoliuciniai neuroniniai tinklai pasižymi geresniu vaizdo, kalbos ir garso signalo įvesčių apdorojimu. Pirmuosiuose sluoksniuose daugiausia dėmesio skiriama paprastoms funkcijoms, tokioms kaip spalvų ar kraštų išskyrimas. Įvesties duomenims keliaujant tinklo struktūros sluoksniuose, pradedami atpažinti tam tikri objekto elementai ar bruožai, galiausiai nustatomas numatytasis rezultatas. Toliau šiame skyriuje plačiau analizuojami trys pagrindiniai sluoksnių tipai [6]:

- konvoliucinis sluoksnis;
- sutelkimo sluoksnis;
- visiškai sujungtas sluoksnis.

Konvoliucinis sluoksnis

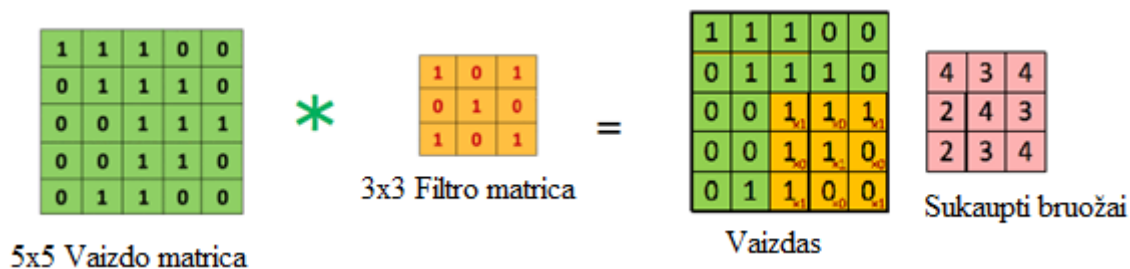
Konvoliucinis (angl. *convolution*) sluoksnis, apdoroja ypatybes iš įvesties vaizdo. Konvoliucija išsaugo ryšius tarp taškų (angl. *pixels*), skenuodama vaizde esančius bruožus. Tai yra matematinė operacija, kuriai atlikti reikalingi du įėjimai, tokie kaip vaizdo matrica, filtras arba branduolys [8]. 1.2 paveiksle matoma iš ko susideda konvoliucinis sluoksnis [6]:

- vaizdo matrica;
- filtras;
- išvestis.



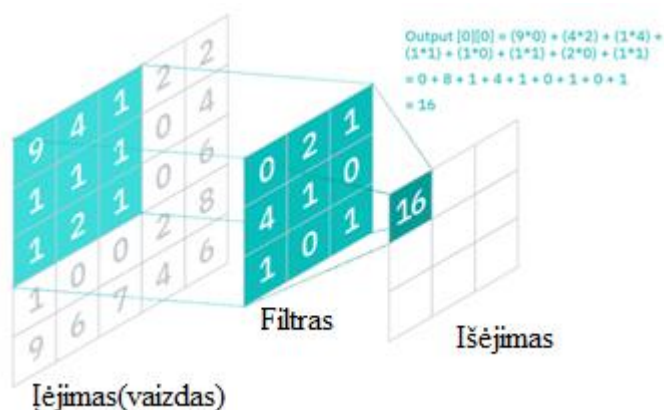
1.2 pav. Konvoliucinis sluoksnis [6]

Sudauginus 5 x 5 dydžio vaizdą, kurio taškų vertės yra 0, 1 ir 3 x 3 dydžio filtro matricą, gaunamas bruožų / ypatybių žemėlapis, kuris parodytas žemiau (žr. 1.3 pav.).



1.3 pav. Bruožų žemėlapis sandara [6]

Jeigu įvestis spalvotas vaizdas, kurį sudaro „3D“ taškų matrica, tai įvestis turės tris matmenis – aukštį, plotį ir gylį, kurie atitinka „RGB“ vaizdą.



1.4 pav. Bruožų žemėlapis [5]

Prieš tai esančiame paveiksle (žr. 1.4 pav.) matoma, kad kiekviena išvesties reikšmė funkcijų žemėlapyje neprisijungia prie kiekvienos įvesties vaizdo taškų vertės. Ji turi prisijungti tik prie imlaus lauko, kuriame naudojamas filtras. Kadangi išvesties masyvo nereikia tiesiogiai susieti su kiekviena įvesties verte, konvoliucinius sluoksnius galima vadinti iš dalies sujungtais sluoksniais. Ši charakteristika literatūroje apibūdinama kaip vietinis ryšys.

Bruožų detektoriaus svoriai lieka fiksuoti, kai jis juda per vaizdą. Šis procesas vadinamas parametru bendrinimu. Kai kurie parametrai, pvz., svorio vertės, apmokymo metu koreguojamos dauginimo proceso ar gradientinio nusileidimo būdu (angl. *descent*). Tačiau yra trys parametrai, turintys įtakos išvesties tūrio dydžiui, kuriuos reikia nustatyti prieš pradedant neuroninio tinklo apmokymą. Tai yra: filtrų skaičius, žingsnis ir užpildymas (angl. *padding*).

Žingsnis

Žingsnis yra taškų skaičius, pasislinkusių per įvesties matricą. Kai žingsnis yra 1, filtrai vienu metu perkeliama per 1 tašką. Toliau pateiktame paveiksle (žr. 1.5 pav.) matoma, kaip atrodo konvoliucija, kurios metu naudojamas 3 x 3 dydžio filtras, veikiantis su žingsniu, kurio dydis 2. Galima paminėti, kad visos reikšmės, patenkančios į 3 x 3 rėmą, yra susumuojamos ir ši operacija atliekama po kiekvieno žingsnio kas du langelius [6].



1.5 pav. Filtras [6]

Užpildymas

Užpildymas naudojamas, kai filtrai neatitinka įvesties vaizdo. Tam nustatomi visi elementai, kurie nepatenka į įvesties matricą sukuriant didesnę arba vienodo dydžio išvestį. Užpildymų tipai [6]:

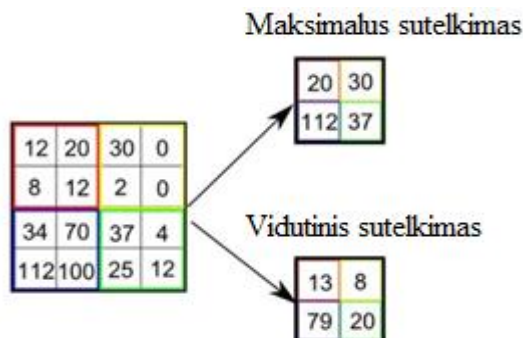
- užpildymas (angl. *valid padding*), kurio metu išlaikoma ta pati struktūra, dar kitaip vadinama be užpildymo. Tokiu atveju paskutinė konvoliucija atmetama, jei matmenys nesutampa;
- toks pats užpildymas (angl. *same padding*). Šio užpildymo metu užtikrinama, kad išvesties sluoksnis būtų tokio paties dydžio kaip įvesties sluoksnis;
- pilnas užpildymas (angl. *full padding*). Tokio užpildymo metu padidinamas išvesties dydis, prie įvesties kraštų pridėdami nulius.

Po kiekvienos konvoliucijos operacijos tinklo funkcijų žemėlapiams taikoma „ReLU“ (angl. *Rectified Linear Unit*) transformacija, kuri įveda modelio pakitimus. Galiausiai, konvoliucinis sluoksnis paverčia vaizdą skaitmeninėmis vertėmis leisdamas neuroniniam tinklui interpretuoti ir išgauti svarbius modelio duomenis.

Sutelkimo sluoksnis

Sluoksnių sutelkimas sumažina parametrų skaičių, kai vaizdai yra per dideli. Erdvinis sutelkimas taip pat vadinamas daliniu, o jo metu sumažinami kiekvieno žemėlapio matmenys, tuo pat metu išsaugant svarbiausią informaciją. Erdvinis sutelkimas gali būti įvairių tipų (žr. 1.6 pav.) [7]:

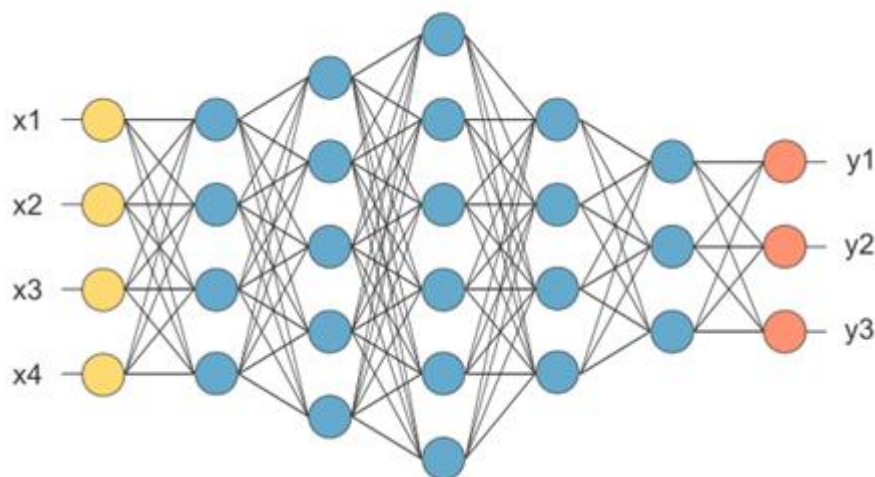
- didžiausio sutelkimo metu paimamas didžiausias elementas iš pataisyto bruožų žemėlapio;
- tokiu pačiu principu veikia ir vidutinis sutelkimas – paimamos vidutinės vertės;
- visų elementų, esančių funkcijų žemėlapio elementuose sumos naudojimas yra vadinamas suminiu sutelkimu.



1.6 pav. Sutelkimo sluoksnių pavyzdžiai [7]

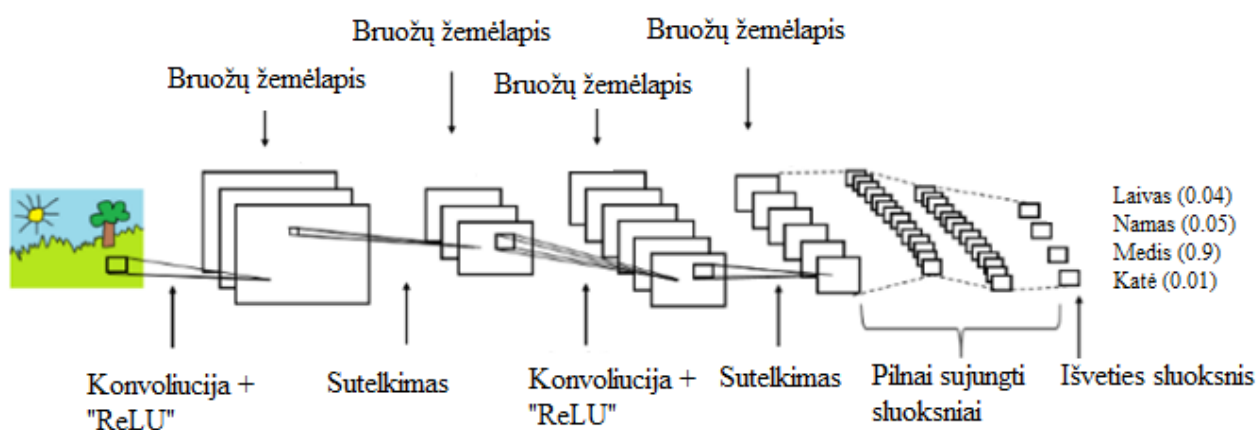
Visiškai sujungtas sluoksnis

Visiškai sujungtu sluoksniu vadinama matrica, kuri yra išskleista į vektorių. Visiškai sujungtame sluoksnyje kiekvieno išvesties sluoksnio mazgas jungiasi tiesiogiai su ankstesniame sluoksnyje esančiu mazgu (žr. 1.7 pav.) [6].



1.7 pav. Visiškai sujungtas sluoksnis [6]

Tokia struktūra leidžia atlikti klasifikavimo užduotį pagal ankstesniuose sluoksniuose išgautas ypatybes ir skirtingus jų filtrus. Po iki galo sujungtų sluoksnių atliekama „softmax“ aktyvavimo transformacija, siekiant užtikrinti tinkamą įvesties klasifikavimą, sukuriant tikimybę nuo 0 iki 1. Galutinis konvoliucinio tinklo vaizdas matomas paveiksle apačioje (žr. 1.8 pav.) [5].

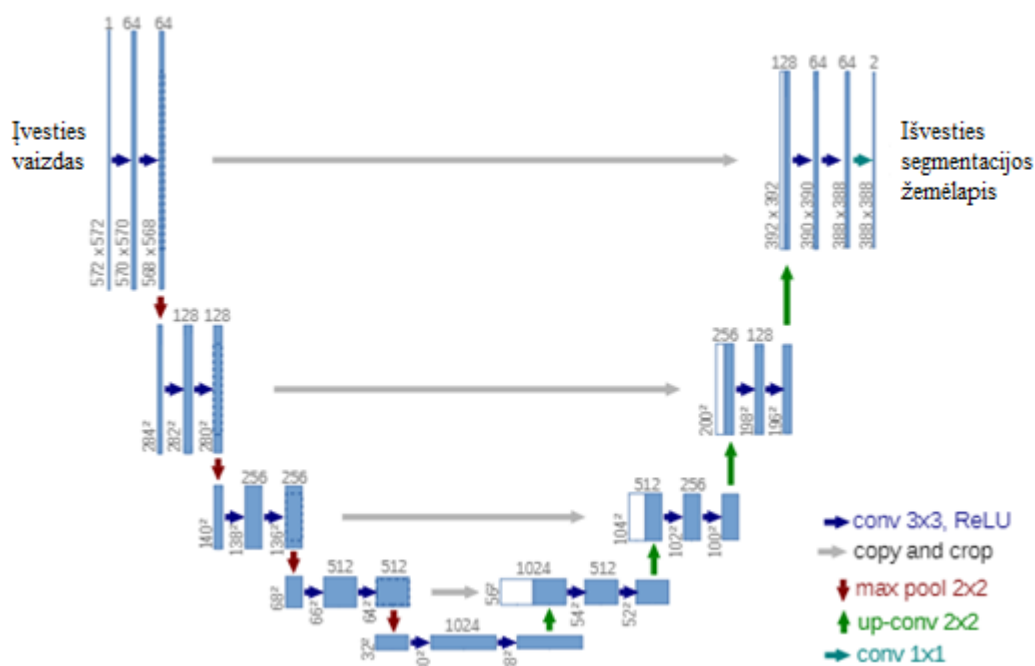


1.8 pav. Viso konvoliucinio neuroninio tinklo architektūra [6]

1.2. „U-net“ modelis

Dažnai susiduriama su problema, kad duomenų bazės yra sudarytos iš mažai duomenų turinčių rinkinių. Mažas duomenų kiekis lemia tai, kad naudojant dirbtinio intelekto modelius gaunami prasti rezultatai, nes jiems apmokyti reikalingi platūs duomenų rinkiniai. „U-net“ modelio struktūra yra ypatinga tuo, kad ji buvo sukurta dirbti su mažai duomenų turinčiomis duomenų bazėmis, tokiomis kaip biomedicininės paskirties vaizdai [9].

„U-Net“ pavadinimą paėmė iš architektūros, kuri vizualizuota atrodo panaši į raidę „U“, kaip parodyta toliau esančiame paveiksle (žr. 1.9 pav.). Modelio struktūra sudaryta iš suspaudžiamosios ir išplečiamosios dalių. Svarbiausi tinklo aspektai yra išplėtimo (antroji dalis) dalyje. Kiekvieną standartinę konvoliucijos procesą transformuoja „ReLU“ aktyvavimo funkcija [8].



1.9 pav. „U-net“ modelio architektūra. Kiekviena mėlyna dėžutė atspindi daugiakanalius broožų žemėlapius. Kanalų skaičius nurodytas virš kiekvienos dėžutės. „X“ ir „Y“ – dydis nurodytas kiekvienos dėžutės kairėje pusėje. Baltos dėžutės atspindi nukopijuotus broožų žemėlapius. Legendoje matomos rodyklės apibrėžia skirtingas operacijas [9].

1.2.1. „U-net“ modelio architektūra

Tinklo architektūra pavaizduota 1.9 paveiksle. Ją sudaro sutraukimo (angl. *contracting*) kelias (kairė pusė) ir išsiplėtimo kelias (dešinė pusė). Sutraukimo dalis, tai tipinė konvoliucinio tinklo architektūra. Ją sudaro 3 x 3 dydžio pasikartojančios konvoliucijos, po kurių seka ištaisymo transformacijų „ReLU“ algoritmai ir 2 x 2 dydžio maksimalaus sutelkimo operacijos su mažinimo žingsniu, kurio vertė – 2. Kiekviename mažinimo / atrankos etape broožų kanalų skaičius yra padvigubinamas. Kiekvieną plečiamojo kelio žingsnį sudaro broožų žemėlapių praplėtimas, po kurio eina 2 x 2 dydžio konvoliucija. Ši konvoliucija dvigubai padidina broožų kanalų skaičių, prijungdama apkarpytus plečiamojo kelio broožų žemėlapius ir dvi 3 x 3 dydžio konvoliucijas, kurios yra transformuotos „ReLU“ algoritmu. Savybių žemėlapių apkarpymas yra būtinas dėl kraštinių taškų (angl. *pixels*) praradimo kiekvienos atliekamos konvoliucijos metu. Paskutiniame sluoksnyje naudojama 1 x 1 dydžio konvoliucija, kuri skirta 64 komponentų ypatybių vektoriui susieti su norimu klasių skaičiumi. Iš viso tinkle yra 23 konvoliuciniai sluoksniai [9].

Sutelkimo sluoksniai „U-net“ tinkle

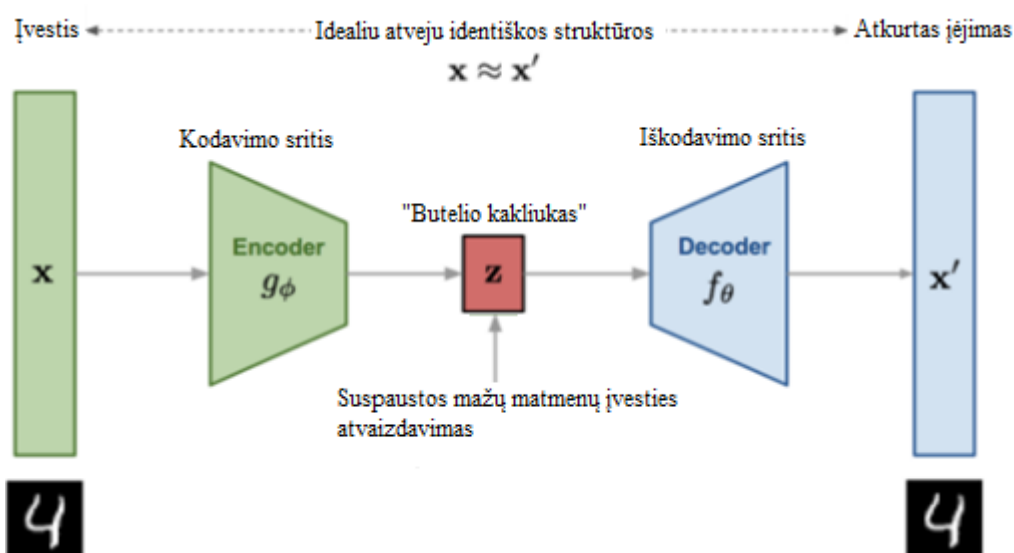
Sutelkimo sluoksnis – tai dimensijų mažinimo / didinimo procesą aukštyje ir plotyje atliekantis sluoksnis. Jis taikomas visame konvoliuciniame neuroniniame tinkle. Sutelkimo sluoksnis sumažina informaciją apie aukštį ir plotį, išlaikydamas pastovų įvesties matricos kanalų skaičių. Tai naudojama

siekiant sumažinti modelio sudėtingumą. Apibendrinant galima sakyti, kad sutelkimo sluoksnis sudaro tašką, kuris atspindi buvusių taškų grupes, jas apjungia.

Lokalizavimo metu išvesties rezultatas, keliaudamas modeliu, yra sujungiamas su didelės raiškos bruožais. Tokiu būdu nuoseklus konvoliucinis sluoksnis sukuria tikslesnę išvestį, pagrįstą prieš tai naudota informacija [7].

„U-net“ tinklo struktūra skiriasi nuo kodavimo – iškodavimo modelio struktūros (žr. 1.10 pav.) tuo, kad klasikinė automatinio kodavimo sistema sumažina įvesties informacijos dydį kartu su sluoksniais. Šis procesas vyksta kol pasibaigia kodavimo dalis ir prasideda iškodavimo architektūros sritis. Šioje srityje toliau naudojamas tiesinis bruožų informacijos perdavimo metodas, kurio metu išvesties dydis palaipsniui didėja. Architektūros pabaigoje išvesties ir įvesties duomenų dydžiai yra lygūs.

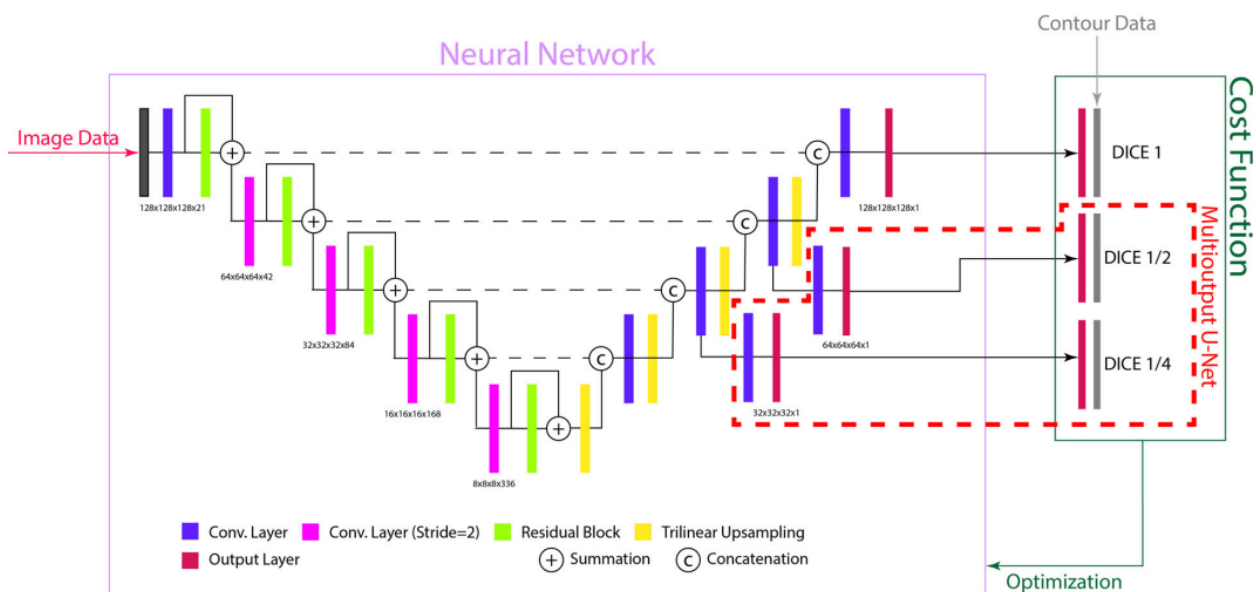
Ši architektūra yra ideali, kai norima išsaugoti išvesties dydį, tačiau atsiranda problema, kad suspaudimą vykdant tiesiškai, centrinėje tinklo dalyje (žr. 1.10 pav. pavaizduota raudonai), prarandama dalis bruožų, kurie nėra atkuriami iškodavimo srityje. „U-net“ modelis skiriasi tuo, kad iškodavimo pusėje atliekama dekonvoliucija, kuri leidžia apeiti ir išvengti prieš tai minėtos problemos [8].



1.10 pav. Kodavimo – iškodavimo modelis [7]

1.2.2. „U-net“ išėjimo sluoksniai mokomi su pastiprintu mokymu

Šiame skyrelyje aprašoma „U-net“ struktūra, kuri skirta sustiprinti tinklo išskleidimo dalies išvestis ir leidžia išsaugoti svarbiausią informaciją iki vidinio paslėptojo sluoksnio. Klasikinio „U-net“ atveju įvesties duomenys apdorojami nuosekliai, išvestis – paskutiniojo išskleidimo sluoksnio išėjimas. Modernizuojant tinklą siekiama sustiprinti mokymosi tikslumą, tai pasiekama pakeičiant „U-net“ modelio išėjimo dalies struktūrą ir priverčiant minimizuoti mokymo klaidą per visus tinklo sluoksnius [12]. 1.11 pav. matomas modernizuotas „U-net“ tinklas, kuris naudoja išskleidimo dalies atskirų sluoksnių išvesčių apjungimą, siekiant išsaugoti svarbiausią informaciją modelio išskleidimo struktūros dalyje. Daugiasluoksnius išėjimus naudojančio „U-net“ tinklo tikslumo rezultatai apžvelgiami tiriamojoje dalyje (žr. 4 skyriuje).

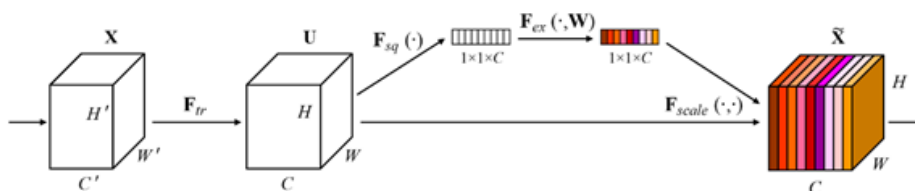


1.11 pav. „U-net“ tinklas su daugiasluoksniu išėjimu [12]

1.3. Požymių pastiprinimo blokas

Požymių pastiprinimo (angl. *Squeeze-and-Excitation*) blokas, tai skaičiavimo vienetas, kurį galima sukurti naudojant bruožų žemėlapių transformacijas. Suspaudimo operacija sukuria kanalų aprašus, kurie vaizduoja kiekvieno kanalo informaciją visuose funkcijų žemėlapuose. Išskleidimo operacija susideda iš visiškai sujungtų sluoksnių ir „ReLU“ transformacijų, kurios adaptyviai sustiprina informacines savybes ir slopina mažiau naudingas [13].

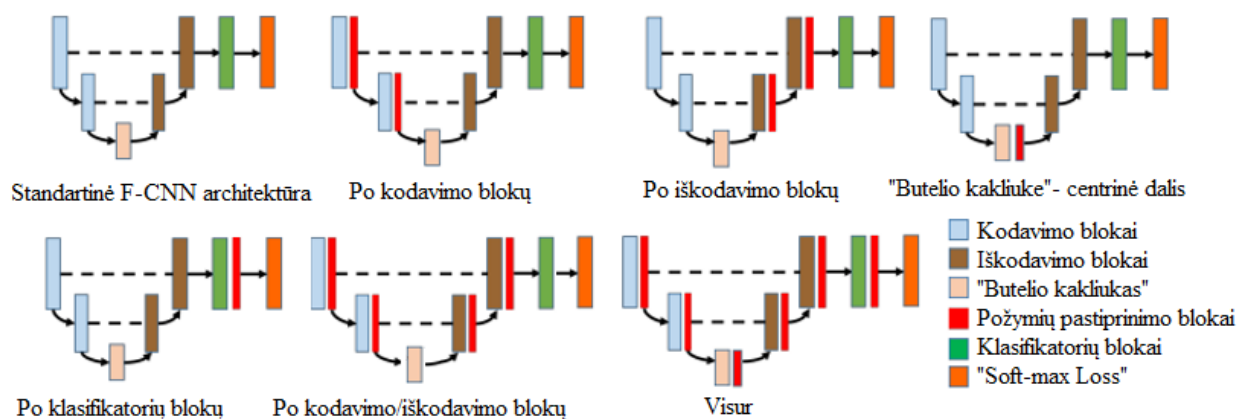
Požymių pastiprinimo bloko struktūra matoma žemiau pateiktame paveiksle (žr. 1.12 pav.) Transformacijai F_{tr} atlikti naudojamas įvesties dėmuo „X“, kuris yra dauginamas su bruožų žemėlapiumi. Iš šio konvoliucijos pavyzdžio galima sukonstruoti atitinkamą požymių pastiprinimo bloką, kurio pagalba atliekamas funkcijų perkaliibravimas. Bruožų žemėlapis (žr. 1.12 pav.), pavaizduotas „U“ raide, pirmiausia keliauja pro suspaudimo operaciją, kuri sukuria visumą apibūdinimų, sujungiant funkcijų žemėlapius su jų erdviniais matmenimis. Ši apibūdinimo funkcija įterpta į visuotinį (angl. *global*) kanalų funkcijų atsakų pasiskirstymą, leidžiantį informaciją iš tinklo imliojo lauko naudoti visiems jo sluoksniams. Po suspaudimo operacijos seka išskleidimo operacija, kuri įgauna savaiminį išėjimų mechanizmą ir sukuria kanalų moduliacijų svorių rinkinius. Šie svoriai pritaikyti bruožų žemėlapiams „U“, leidžia generuoti pastiprinimo bloko išėjimus, kurie gali būti panaudoti tolesniuose tinklo sluoksniuose [11].



1.12 pav. Požymių pastiprinimo bloko struktūra [11].

Vienas svarbiausių kriterijų siekiant išgauti kuo aukštesnį požymių pastiprinimo blokų efektyvumą, tai teisingai parinkta pozicija neuroninio tinklo modelio struktūroje. Galimos 6 pagrindinės pozicijos (žr. 1.13 pav.) [13]:

- po kodavimo blokų;
- po iškodavimo blokų;
- „butelio kakliuke“ centre;
- po klasifikatorių blokų;
- po kodavimo ir iškodavimo blokų;
- po kodavimo, iškodavimo, klasifikatorių blokų bei centre – „butelio kakliuke“.

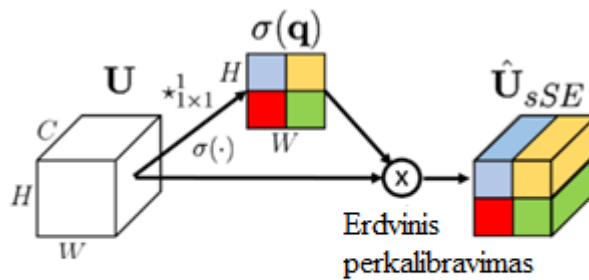


1.13 pav. Požymių pastiprinimo bloko pozicijos [13]

Remiantis literatūroje aprašytais rezultatais, matoma, kad aukščiausi kokybiniai įverčiai gaunami, kai naudojamo požymių pastiprinimo bloko pozicija yra po kodavimo / iškodavimo blokų arba kai išdėstyti pagal paskutinę pateiktą struktūrą paveiksle. Šiame tiriamajame darbe atliekami skaičiavimai, kai požymių pastiprinimo blokai išdėstyti visoje modelio struktūroje (žr. 1.13 pav. „visur“). Toliau 4 skyriuje išbandomi 5 skirtingų tipų požymių pastiprinimo blokai, siekiant išsiaiškinti, kuris yra tinkamiausias, norimai operacijai atlikti, t.y. pagerinti, išskiriamų odos defektų, suradimo tikslumą bei sumažinti rezultatų išsibarstymą.

1.3.1. Erdvinio spaudimo būdu požymius pastiprinantis blokas (spaudžiama per geometriją)

Vienas iš dažnai naudojamų požymių pastiprinimo blokų, tai blokas spaudžiantis bruožų žemėlapius per geometriją. Kanalų spaudimo ir erdvinio sužadavimo blokai, spausdami bruožų žemėlapius per kanalus atlieka erdvinį sužadimą, kuris svarbus smulkių vaizdų segmentavimui. Priklausomai nuo erdvinės padėties apibrėžiamas alternatyvus įvesties tensoriaus skaidymas. Atliekama erdvinio suspaudimo operacija, kurios metu sukuriama tensoriaus projekcija. Kiekviena projekcija tiesiškai priklauso viena nuo kitos ir pritaikoma visiems kanalams pagal jų erdvinę padėtį. Projekcija, praėjusi pro sigmoidinį sluoksnį, pakeičia aktyvacijų vertes į $[0, 1]$ imtį. Šios vertės yra naudojamos perkalibruoti ar sužadinti erdvinius bruožų žemėlapius. Kiekvieną vertę apibūdina erdvinė informacija, esanti bruožų žemėlapyje. Perkalibravimas suteikia, didesnę svarbą erdvinėms reikšmingoms padėtimis ir neatsižvelgia į nereikšmingas. Šios struktūros vaizdą matome paveiksle pateiktame žemiau (žr. 1.14 pav.) [13].

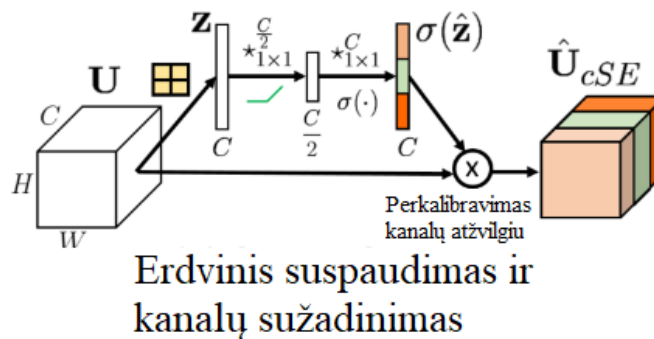


Kanalų spaudimas ir erdvinis išplėtimas

1.14 pav. Erdvinio spaudimo būdu požymius pastiprinantis blokas [13]

1.3.2. Erdvinio spaudimo ir kanalų sužadavimo būdu požymius pastiprinantis blokas (bruožų žemėlapių spaudimas per kanalus)

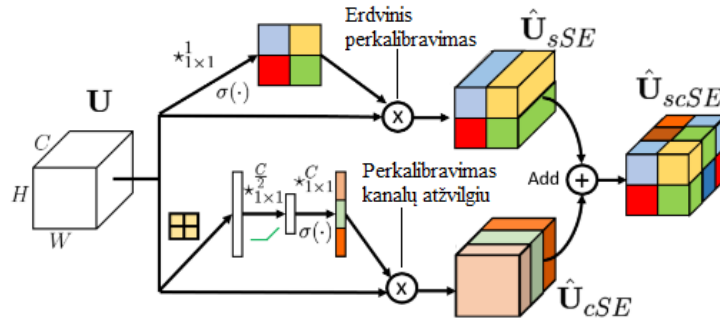
Naudojant erdvinio spaudimo ir kanalų sužadavimo būdu požymius pastiprinantį bloką, įvesties bruožų žemėlapiai „U“ yra apibrėžiami kaip kanalų aibė. Atliekant erdvinį spaudimą, naudojamas bendrasis (angl. *global*) sutelkimo sluoksnis, kuris sukuria „z“ vektorius. Į šį vektorius yra įterpiama bendroji informacija. Po to atliekama transformacija, kurios metu užkoduojamos kanalų priklausomybės. Dinaminis transformuoto vektoriaus „z“ intervalas pereinant per sigmoidinį sluoksnį yra perduodamas į [0, 1] imtį. Gautas rezultatas – vektorius naudojamas bruožų žemėlapio „U“ kalibravimui / sužadimui. Kai tinklas apsimoko, šie aktyvinimai pritaikomi prie nežinomų kanalų, pabrėžiami svarbiausi bruožai. Bloko architektūra matoma 1.15 paveiksle [13].



1.15 pav. Erdvinio spaudimo ir kanalų sužadavimo būdu požymius pastiprinantis blokas [13]

1.3.3. Erdvinio / kanalų suspaudimo ir sužadavimo būdu požymius pastiprinantis blokas (tarpkanalinis spaudimas)

Erdvinio / kanalų suspaudimo ir sužadavimo būdu požymius pastiprinantis blokas, tai anksčiau aprašytų blokų junginys, kuris perkaliavuoja įėjimo žemėlapius geometriškai ir kanaliniame lygmenyje. Tokiu būdu požymių pastiprinimo blokas įgyja papildomų savybių, leidžiančių dar tiksliau išskirti reikiamus bruožus ir at mesti nereikšmingus. Šio apjungto bloko struktūra matoma 1.16 paveiksle pateiktame žemiau (žr. 1.16 pav.) [13].



Tuo pat metu vykstantis kanalų/erdvinis suspaudimas ir kanalų sužadimas

1.16 pav. Erdvinio / kanalų suspaudimo ir sužadimo būdu požymius pastiprinantis blokas [13]

1.3.4. Požymių pastiprinimo blokas, kanalų svarbos pabrėžimas („Adaptive Average Pooling 2D“ spaudimas per kanalus)

Šio požymių pastiprinimo bloko veikimo principas yra apžvelgiamas per skaičiuojamąją dalį. Antroji lygtis apibrėžia konvoliucinio sluoksnio išėjimus \mathbf{X} , kur W_i , H , C yra plotis, aukštis ir kanalų dydis (filtrų kiekis). Kanalų svoriai požymių pastiprinimo blokuose apskaičiuojami kaip (žr. 1 funkciją) [14]:

$$\omega = \sigma \left(f \{ w_1, w_2 \} (g(X)) \right); \quad (1)$$

$$g(X) = \frac{1}{W_{iH}} \sum_{i=1, j=1}^{W_{iH}} X_{ij}; \quad (2)$$

čia, $g(X)$ kanalo atžvilgiu yra bendrasis vidutinis sutelkimas; σ – sigmoidinė funkcija.

Jeigu funkcija $y = g(X)$, $f \{ w_1, w_2 \}$, tai „ReLU“ transformacijos funkcijos išraiška yra (3):

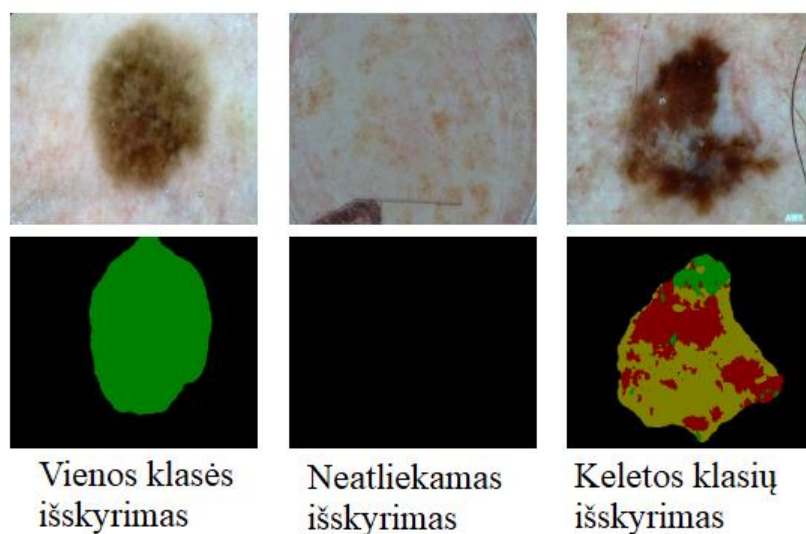
$$f \{ w_1, w_2 \} (y) = W_2 \text{ReLU}(W_1 y); \quad (3)$$

Siekiant išvengti aukšto modelio sudėtingumo, dydžiai W_1 ir W_2 pakeisti į $C * (C/r)$. Pastebima, kad į $f \{ W_1, W_2 \}$ įtraukti parametrai susiję su kanalų svarbos blokais. Nors matmenų mažinimas antroje funkcijoje mažina modelio sudėtingumą, jis taip pat panaikina priklausomybę tarp kanalų ir svorių. Tokiu atveju vienas sluoksnis spėja svorius kiekvienam kanalui, naudodamas tiesines priklausomybes tarp visų kanalų. O antroji funkcija padaro tai, kad pirmiausia kanalų bruožai projektuojami į mažo dydžio erdvę, tada gražinami atgal, taip nutraukiant tiesioginę svorių ir kanalų priklausomybę [14].

1.4. Daugiaklasė segmentacija

Konvoliuciniai tinklai, naudojantys kodavimo – iškodavimo struktūrą, gali atlikti daugiaklasę segmentaciją ir lokalizuoti objektus, atlikdami taškinio lygmens spėjimus. Tai leidžia tinklui išmokti, kiekvieną tašką ir jį priskirti konkrečiai klasei. Šis metodas paplito ir yra plačiai tyrinėjamas medicininės bei kitokios paskirties vaizdams analizuoti [15]. Šiame darbe išbandomi keletas „U-net“ modifikacijų su skirtingais įėjimo vaizdų dydžiais bei filtrų skaičiais, siekiant išsiaiškinti, kuri modifikacija leidžia išgauti aukščiausius išvesties rezultatus. Skirtingų klasių išskyrimo naudojimas tame pačiame vaizde, leidžia atlikti sudėtingesnės analizės operacijas (žr. 1.17 pav. apačioje). Kaip matoma paveiksle, vienos klasės išskyrimo atveju, visa nuotraukoje esanti struktūra priskiriama

vienai klasei. Dešinėje pateiktas vaizdas, kuris gaunamas naudojant daugiaklasę segmentaciją. Galima matyti, kad tame pačiame vaizde yra išskiriama keletas zonų taškiniame lygmenyje (tamsios sritys – melanoma), kas leidžia tiksliau / kokybiškiau įvertinti segmentacijos rezultatus, atlikti sudėtingesnę to pačio vaizdo segmentų analizę, išskaidymą [16].



1.17 pav. Skirtingų segmentacijos tipų pavyzdys. A – vienos klasės išskyrimas; B – rezultatas gaunamas, kai neišskiriama jokia klasė; C – rezultatas, kai vaizde išskiriama keletas klasių. Žalia spalva reiškia melanocitinį nevosą, raudona – melanomą, o geltona – seborėjinę keratozę [16].

1.5. Nuostolių funkcija

Atliekant medicininės paskirties vaizdų analizę susiduriama su viena iš taškinio klasifikavimo problemų – kiekvieno taško priskirimu į pirmąjį planą arba foną. Dažniausiai naudojama nuostolių įvertinimo funkcija, tai kryžminio entropijos praradimo funkcija [17]. Tačiau atsižvelgiant į tai, kad medicininės paskirties vaizduose analizuojami objektai, tokie kaip optinis diskas ar tinklainės kraujagyslės, užimančios mažą paveikslėlio sritį, šis algoritmas neatlieka numatytos funkcijos tinkamai. Todėl šiame darbe yra naudojama „kauliuko“ (angl. *dice*) koeficiento nuostolių funkcija, kuri pakeičia įprastą kryžminio entropijos praradimo funkciją. Koeficientas persidengimo matas naudojamas segmentavimo rezultatams įvertinti, kai yra tenkinama 4 priklausomybė [18]:

$$L_{Dice} = 1 - \frac{\sum_k \omega_k \sum_i p_{(k,i)} g_{(k,i)}}{\sum_i p_{(k,i)}^2 + \sum_i g_{(k,i)}^2}, \quad (4)$$

čia, N – taškų skaičius; $p_{(k,i)}$ priklauso aibei $[0,1]$; $g_{(k,i)}$ priklauso aibei $\{0,1\}$ ir žymi numatomą klasės tikimybės ir pagrindinės tiesos etiketę k . K – klasės numeris; $\sum_k \omega_k = 1$ klasės svoris; $\omega_k = 1/K$. Nuostolių funkcija, leidžia išvengti persimokymo. Jos išraiška pateikiama lygtimi (5):

$$L_{loss} = L_{Dice} + L_{reg}; \quad (5)$$

čia, L_{reg} – reguliavimo praradimas (taip pat vadinamas svorio mažėjimu).

1.6. Vertinimo kriterijus

Neuroninių tinklų rezultatams vertinti sukurta, daug skirtingų vertinimo kriterijų [19], [20], [21]. Vertinimo kriterijaus pasirinkimas priklauso, nuo to kokią funkciją atlieka naudojamas neuroninio

tinklo modelis. Tikslumas, tai vienas lengviausiai suprantamų įvertinimo metrikų. Jis apibrėžiamas kaip teisingai sugrupuotų duomenų skaičius, padalytas iš viso duomenų skaičiaus. Šiame darbe naudojamas „F1“ vertinimo kriterijus, kuris yra tikslesnis rezultatų įvertinimo dydis, nei tikslumo vertinimas, kadangi atsižvelgiama į daugiau faktorių nei tik į teisingai sugrupuotų pavyzdžių skaičių. „F1“ vertinimo kriterijaus formulė yra tikslumo ir pakartojamumo harmoninis vidurkis [22]:

$$F_1 = \frac{2}{\frac{1}{\text{pakartojamumas}} + \frac{1}{\text{tikslumas}}} = 2 * \frac{\text{tikslumas} * \text{atkartojamumas}}{\text{tikslumas} + \text{atkartojamumas}} = \frac{Tp}{Tp + \frac{1}{2}(Fp + Fn)}; \quad (6)$$

čia, *tikslumas* – teisingai teigiamiems priskirtų duomenų dalis tarp duomenų, kuriuos modelis priskyrė teigiamiems. Kitaip tariant, teisingų teigiamų duomenų skaičius, padalytas iš klaidingų teigiamiems priskirtų ir teisingai teigiamiems priskirtų duomenų; *atkartojamumas* – dar vadinamas *jautrumu*, tai trupmena, teisingai teigiamiems duomenims priskirtų duomenų padalijimas iš visų teigiamiems priskirtų duomenų (teisingai ir neteisingai priskirti teigiamiems); *Tp* – teisingai teigiamiems priskirtų duomenų skaičius; *Fn* – klaidingai neigiamiems priskirtų duomenų skaičius; *Fp* – klaidingai teisingiems priskirtų duomenų skaičius.

2. Panašių darbų apžvalga

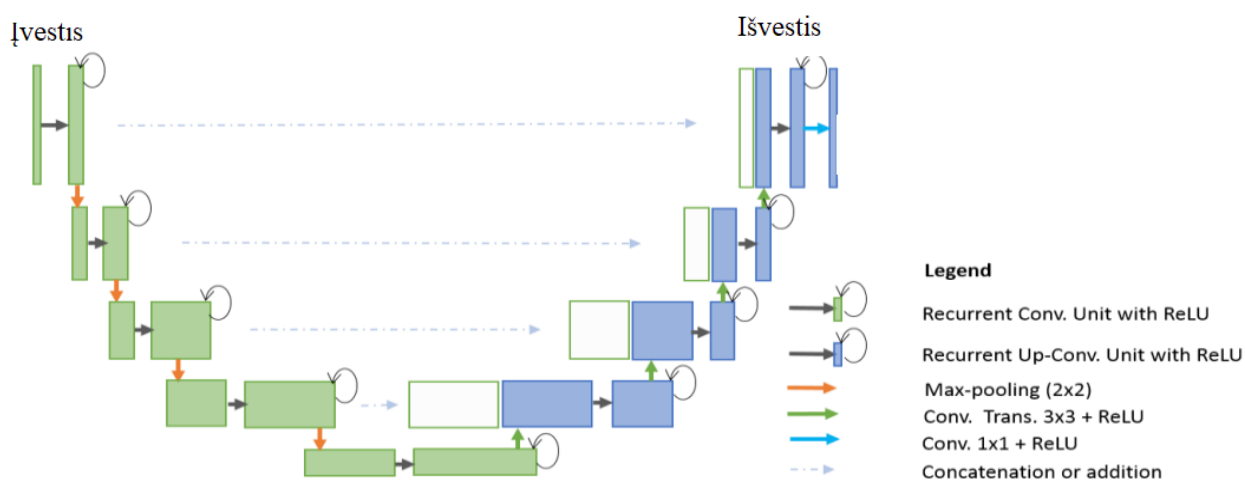
Šiame skyriuje pateikiama darbų apžvalga, kuriuose atliekamas segmentavimas ir klasifikavimas. Didžiausias dėmesys skiriamas „U-net“ struktūrą turintiems modeliams, nes būtent tokios struktūros modelis yra išbandomas eksperimentinėje dalyje. Visi žemiau išvardinti modeliai naudojami tos pačios duomenų bazės analizei [23]. Naudota duombazė „ISIC18“.

2.1. Geriausių „ISIC18“ konkurse pasirodžiusių neuroninių tinklų modelių rezultatai

„R2U-net“, „BCDU-net“ ir „MCGU-net“ – tai trys modeliai („U-net“ tipo), kuriais 2018-2020 metais „ISIC18“ konkurso metu buvo pasiekti aukščiausi rezultatai. Kaip matoma iš pavadinimų, visuose šiuose modeliuose modifikuojamas / papildomas „U-net“ tinklas, kuris tinkamai atlieka mažų segmentų išskyrimo uždavinį.

2.1.1. „R2U-Net“ modelis

„R2U-net“ – tai „U-net“ ir „RRCNN“ modelių junginys [24]. Modelio architektūra matoma, paveiksle pateiktame žemiau (žr. 2.1 pav.). Ši struktūra yra prieš tai minėtų dirbtinių neuroninių tinklų modelių sandūra. Paveikslo legendoje pateikiamos visos naudojamos struktūros / blokai.



2.1 pav. „R2U-net“ architektūra su konvoliuciais kodavimo / iškodavimo keliais, kuriuose naudojami pasikartojantys konvoliuciniai sluoksniai. Tinklas sukurtas pagal „U-net“ modelį [24].

Būtent tokia architektūra pasirinkta dėl keleto privalumų [25]:

- liekamasis vienetas (angl. *residual unit*) yra tinkama architektūra giliojo mokymo modeliams apmokyti;
- funkcijų kaupimo procesas, kurių metu naudojami pasikartojantys liekamieji konvoliuciniai sluoksniai užtikrina geresnį segmentų struktūrų išskyrimą;
- liekamieji konvoliuciniai sluoksniai, leidžia suprojektuoti modelio architektūrą, kuri atlieka kokybišką segmenavimo funkciją per daug neišplečiant „U-net“ modelio struktūros.

Naudojant šį modelį pasiekiami aukšti rezultatai iš kurių geriausiu bandymu pasiekta vertė pagal „F1“ vertinimo kriterijų siekia 69,1 procento.

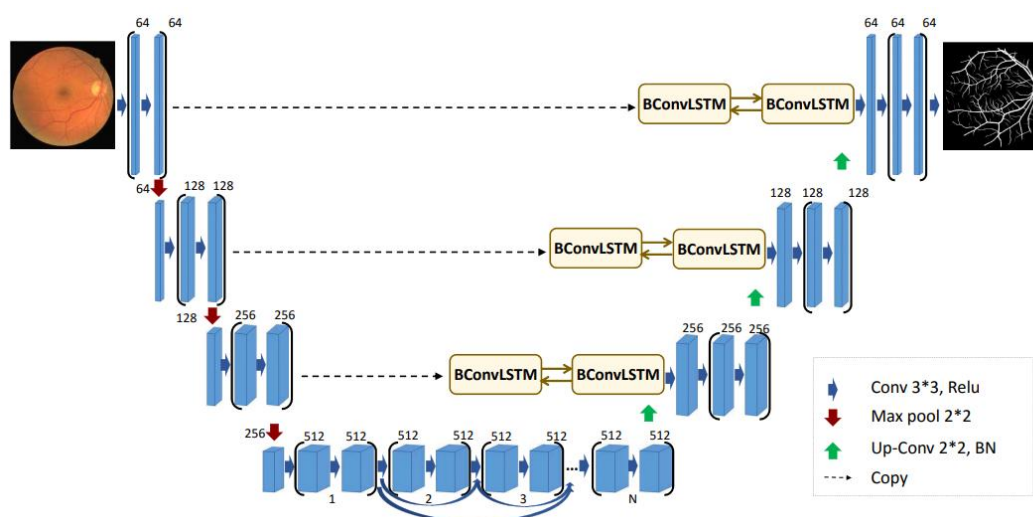
2.1.2. „BCDU-net“ modelis

„BCDU-net“ modelis, tai vieną geriausių rezultatų pasiekianti struktūra „ISIC18“ duombazės duomenims analizuoti. Naudojant šį modelį pasiekiamas iki 84,7 procento „F1“ vertinimo kriterijaus įvertis. Šis modelis turi modifikuotas kodavimo / iškodavimo sritis.

Spaudžiamasis kelias „BCDU-net“ modelyje sudarytas iš 4 etapų. Kiekvienas iš jų susideda iš dviejų 3×3 dydžio konvoliucijų, po kurių eina 2×2 maksimalaus sutelkimo ir „ReLU“ transformacijų operacijos. Spaudžiamajame kelyje palaipsniui išskiriami vaizdo bruožai, kurių kiekis didėja po kiekvieno sluoksnio. Paskutiniame kodavimo kelio sluoksnyje gaunama aukšto lygio semantinė informacija. Konvoliucinių sluoksnių seka tinkle leidžia modeliui išmokti įvairių rūšių ypatybes. Naudojamos tankiai sujungtos konvoliucijos, kas lemia, kad funkcijų žemėlapiai, išmokti iš visų ankstesnių konvoliucinių sluoksnių yra susieti su funkcijų žemėlapiu iš esamo sluoksnio. Visa ši informacija persiunčiama į sekančią konvoliuciją – tampa jos įvestimi. Tankiai sujungtų konvoliucijų privalumai [26]:

- sukuriama informatyvus funkcijų žemėlapis, nekaupiami nereikalingi bruožai;
- ši struktūra užtikrina informacijos tėkmę tinkle, užtikrinamas pakartotinis jau buvusių bruožų naudojimas;
- tankiai sujungtas sluoksnis leidžia naudoti prieš tai buvusią informaciją ir taip sumažina bruožų pamiršimo galimybę;
- atgaliniu keliu gradientai greičiau persiunčiami į reikiamus tinklo taškus.

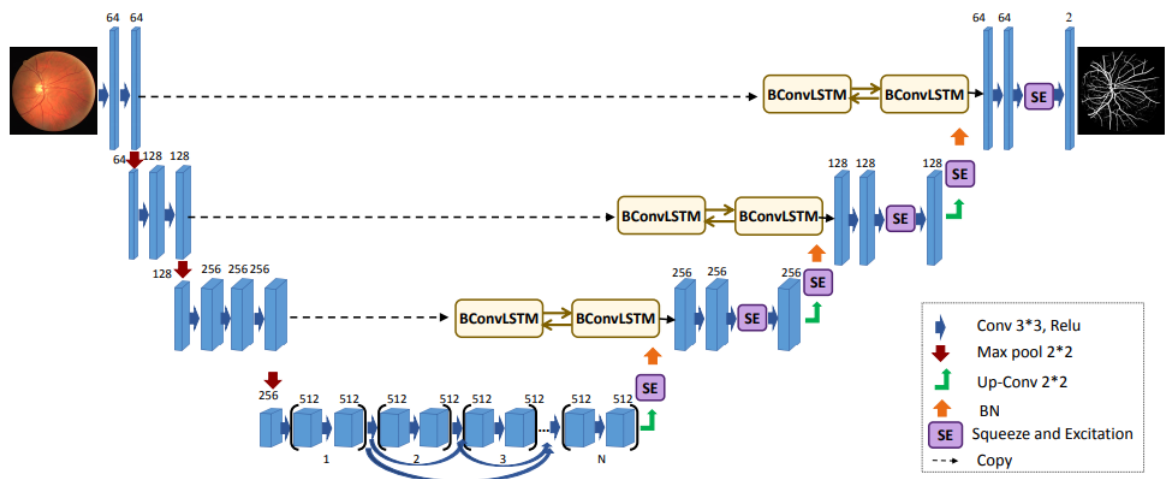
Perėjus į iškodavimo struktūrą, kiekvienas žingsnis prasideda nuo atrankos funkcijos, kuri atliekama naudojant ankstesnio sluoksnio išvestį. Standartiniame „U-net“ bruožų žemėlapiu esantys suspaudimo dalyje apkarpmi ir kopijuojami į iškodavimo struktūros dalį. Abiejų kelių žemėlapiu apjungiami panaudojant atrankos funkciją (angl. *upsampling*). „BCDU-net“ struktūros atveju naudojamas „BConvLSTM“ algoritmas [27], apdorojantis funkcijų žemėlapius sudėtingesniu būdu, taip pateikiama tikslesnė išvestis. Kitaip tariant, išplečiamajame kelyje yra išplečiamas funkcijų žemėlapiu dydis [26]. Visa šio tinklo struktūra matoma 2.2 paveiksle.



2.2 pav. „BCDU-net“ modelis. Naudojamas dvikryptis „ConvLSTM“ modulis praleidimo jungtyse (angl. *skip connections*), tankiai sujungtos konvoliucijos [28].

2.1.3. „MCGU-Net“ modelis

Trečiasis aprašomas modelis yra „MCGU-net“. Ši struktūra, tai geriausią rezultatą šiuo metu gaunantis modelis. Nuo prieš tai aprašytos „BCDU-net“ tinklo struktūros skiriasi tuo, kad ši struktūra papildyta požymių pastiprinimo blokais (žr. 2.3. pav.). Paveiksle matoma, kad požymių pastiprinimo blokai naudojami tik iškodavimo dalyje (plačiau apie pastiprinimo blokų pozicijas žr. 1.3 poskyryje) [29].

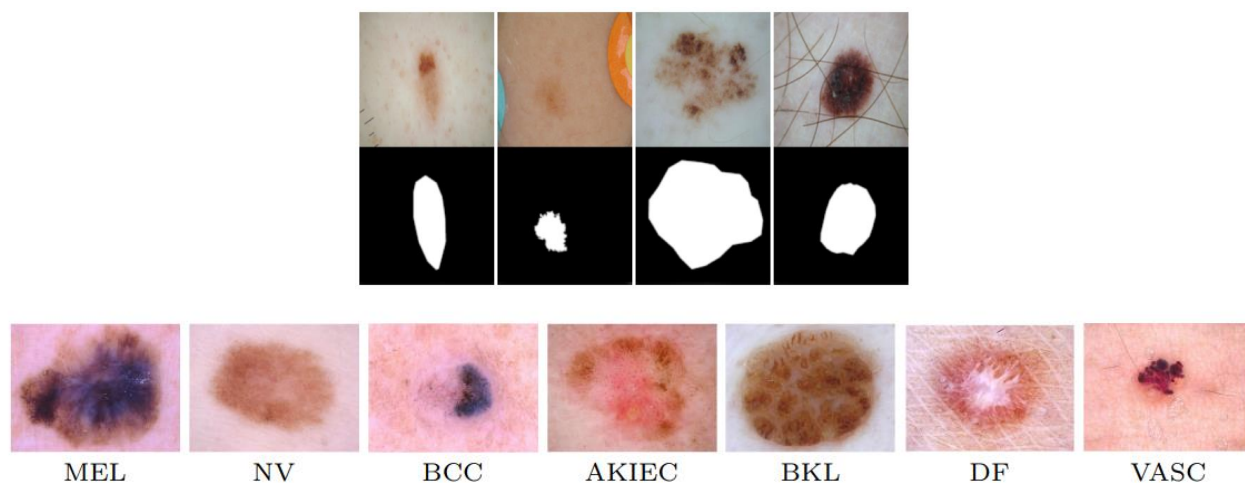


2.3 pav. „MCGU-net“ modelis. Modelis naudoja „BCDU-net“ struktūrą, iškodavimo sluoksnyje papildytą požymių pastiprinimo blokais [29].

Šios architektūros modelis pasiekia iki 89.5 procentų vertę „F1“ vertinimo skalėje.

3. „ISIC18“ duomenų bazė

„ISIC18“ – tai duomenų bazė sudaryta iš nuotraukų. Vaizduose matomi odos pažeidimai, kuriuos sukelia skirtingos odos ligos, tokios kaip melanoma. „ISIC18“, tai didžiausia viešai prieinama duombazė, naudota „ISIC18“ konkurse. Šio konkurso tikslas atlikti kuo tikslesnį vaizduose matomų odos pažeidimų segmentavimą, defekto išskyrimą bei klasifikavimą pagal tam tikrus požymius [30]. Nuotraukos skiriasi dydžiu bei jose matomi dariniai yra skirtingų spalvų ir formų.



3.1 pav. „ISIC18“ duombazę, sudarančių nuotraukų pavyzdžiai. Viršuje matoma melanomos segmentacijai skirta duombazės dalis [31], apačioje pateikiamas visų 7 klasių esančių duombazėje vaizdas [32].

Duombazė sudaryta iš 7 skirtingų kategorijų ligų grupių žr. 3.1 pav. [33]:

- melanoma („MEL“);
- apgamai („NV“);
- bazalinių ląstelių karcinoma („BCC“);
- aktininė keratozė („AKIEC“);
- gerybinė keratozė („BKL“);
- dermatofibroma („DF“);
- kraujagyslių pažeidimai („VASC“).

Ši duomenų bazė naudota konkurse, kuriame reikėjo atlikti tris užduotis. Šiame darbe aprašomas tyrimas atitinka 1 ir 2 užduotis, tai yra melanomos darinių segmentacija bei melanomos vaizdų atskirų segmentų iš viso 5 klasių klasifikacija. Kiekvienai užduočiai skirtų duomenų kiekis skiriasi. Segmentacijos užduotims atlikti pateikiamos 2594 skirtingų dydžių (daugiausia 4 / 3 santykio) nuotraukos, kuriose matomi vėžiniai dariniai bei 2594 vaizdų kaukės. Ligos klasifikavimo uždavinys sprendžiamas naudojant 10015 nuotraukas, iš kurių 1113 – melanoma, 6705 – apgamai, 514 – bazalinių ląstelių karcinoma, 327 – aktininė keratozė, 1099 – gerybinė keratozė, 115 – dermatofibroma, 142 – kraujagyslių pažeidimai.

Šiuos duomenis segmentuoti ir klasifikuoti bandoma įvairių struktūrų modeliais. Iš atliktos literatūros apžvalgos matoma, kad aukščiausi rezultatai pasiekiami naudojant „U-net“ pagrindu sudarytus modelius (žr. 3.1 lentelę) [34]. Taip yra todėl, nes duomenų / nuotraukų kiekis nėra didelis. Apžvelgus senesnės odos darinių duombazės „ISIC17“ segmentacijos rezultatus [35], matoma, panaši statistika – vieni aukščiausių rezultatų pasiekiami naudojant „U-net“ struktūrų modelius, tačiau kadangi

„ISIC17“ duombazė yra senesnė galima rasti ir kitų tinklų pagrindu sudarytų modelių, kurie nežymiai viršija „U-net“ gaunamus rezultatus, pvz. „LinkNet152“ [36], „SegNet“ [37], „II-FCN“ [38].

3.1 lentelė. Geriausių modelių rezultatai, naudota duomenų bazė „ISIC18“

Modelio pavadinis	„F1“ vertė
MCGU-Net	89,5
ResNet50	85
BCDU-net	84,7
MobileNet	82
DenseNet169	82
Inception V3	80
Inception ResNet v2	79

Šių medicininių duomenų apdorojimo rezultatai neviršija 90 proc. Taip yra todėl, nes:

- apmokymui skirtų duomenų kiekis yra per mažas;
- duombazėje pateikiamų vaizdų / kaukių kokybė nėra ideali;
- vykdomos segmentacijos objektas yra sąlyginai mažas lyginant su fono vaizdo kiekio dydžiu;
- siekiant išgauti aukštus rezultatus modelių struktūra sudėtingėja, įgauna daug parametrų, dėl ko reikalingi galingi kompiuteriniai resursai.

4 skyriuje aprašomiems eksperimentams atlikti naudojami šios duomenų bazės vaizdai, kurie yra apdorojami, naudojantis papildytos / modifikuotos „U-net“ struktūros modeliu. Prieš pateikiant modifikuotos „U-net“ struktūros rezultatus apžvelgiami keleto panašių darbų rezultatai (žr. 2 skyriuje). Tyrimo metu duombazės duomenys atsitiktinai priskiriami 80 proc. mokymui ir 20 proc. patikrinimui. Rezultatų vertinimo kriterijus – „F1“ koeficiento vertė.

4. Tiriamoji dalis

Šiame skyriuje atliekama odos defektų išskyrimo iš nuotraukose matomų vaizdų užduotis. Tikslui pasiekti naudojamos 128 x 128 taškų bei 256 x 256 taškų dydžio pilno vaizdo nuotraukos. Skaičiavimai atlikti naudojant standartinį bei modernizuotą „U-net“ modelius. Gaunami rezultatai pateikiami lentelėse, vertinimo kriterijus – „F1“ koeficientas.

4.1. Duombazės paruošimas

Vienas svarbiausių tinkamai paruoštų duomenų kriterijų yra duomenų atsitiktinumas. Tai svarbu, kad vykdant mokymo, testavimo ar patvirtinimo operaciją gaunamiems rezultatams nebūtų daroma įtaka, kuri galėtų nulemti netikslumus / klaidingus skaičiavimus, duomenys nesikartotų. Šiam tikslui pasiekti, atliekant duomenų paruošimą, naudojamas sumaišymo algoritmas. Algoritmas užtikrina duomenų atsitiktinumą. Visų pirmą visiems duomenims (nuotraukoms) priskiriama atsitiktinė pozicija, tai lemia, kad kiekvieną kartą atlikus skaičiavimus ta pati nuotrauka [0:2594] imtyje turės kitą poziciją. Toliau vykdomas duomenų suskirstymas į 3 skirtingas imtis. O kadangi prieš tai duomenų bazėje esančių nuotraukų vieta priskiriama atsitiktinai, tai kiekvieną kartą prieš atliekant skaičiavimus testavimo, mokymo ir patikrinimo imtyse naudojami skirtingi duomenys [39]. Taigi, tokio algoritmo naudojimas paruošiant duomenis skaičiavimams leidžia užtikrinti atsitiktinių duomenų panaudojimo galimybę, taip išvengiama pasikartojimų, kurie gali lemti rezultatų netikslumus.

Testavimo imtį sudaro 520 vaizdų, patvirtinimo – 259, o mokymo daugiausia – 1815. Visų pirma bandymai atliekami naudojant nuotraukas, kurios suspaustos / sudarytos iš 128 x 128 taškų, po to pereinama į sudėtingesnę vaizdų išskaidymo lygmenį, naudojami vaizdai sudaryti iš 256 x 256 taškų.

4.2. Pradinio „U-net“ modelio parametrų parinkimas

Siekiant atlikti odos defektų segmentavimo uždavinį reikia surasti kokios struktūros modelis geriausiai atlieka šią funkciją. Todėl visų pirma atliktas bandymas, kurio tikslas nustatyti kaip kinta „U-net“ modelio tikslumas vertinant pagal „F1“ kriterijų, kai keičiamas filtrų bei sluoksnių kiekis naudojamo modelio struktūroje. Bandymo sąlygos: įvesties vaizdai yra sudaryti iš 128 x 128 taškų arba iš 256 x 256 taškų kiekio. Filtrų kiekis kinta nuo 8 iki 64, o sluoksnių kiekis atitinkamai nuo 3 iki 5. Gauti rezultatai pateikiami 4.1 ir 4.2 lentelėse.

4.1 lentelė. „U-net“ modelio rezultatai pagal „F1“ vertimo kriterijų, kai įėjimo vaizdo dydis 128 x 128 x 3

Įėjimo vaizdo dydis 128 x 128 x 3				
Gylis	Filtrų sk.			
	8	16	32	64
3	0,8384	0,8527	0,8695	0,8638
4	0,8441	0,8668	0,8724	0,8679
5	0,8724	0,8459	0,8559	0,8722

4.2 lentelė. „U-net“ modelio rezultatai pagal „F1“ vertimo kriterijų, kai įėjimo vaizdo dydis 256 x 256 x 3

Įėjimo vaizdo dydis 256 x 256 x 3				
Gylis	Filtrų sk.			
	8	16	32	64
3	0,8385	0,8241	0,8292	0,8383
4	0,8523	0,8650	0,8601	0,8569
5	0,8683	0,8565	0,8673	0,8626

Iš gautų rezultatų pastebima, kad dvi aukščiausios „F1“ vertinimo kriterijaus vertės pasiekia modelio struktūros, kuriuos sudarytos iš 5 sluoksnių, kai naudojamas filtrų kiekis 8 ir 32, o įėjimo vaizdo dydis 256 x 256 x 3 taškų. Trečią aukščiausią įvertį pasiekia 4 sluoksnių modelio architektūra su 32 filtrais, kai įėjimo vaizdo dydis yra 128 x 128 x 3 taškų. Būtent šios trys modelio konfigūracijos pasirinktos tolimesniems skaičiavimais atlikti tobulinant modelį.

4.3 lentelė. Pradinio „U-net“ modelio tikslumo rezultatai

Konfigūracija (keičiama imtis modelis U_net)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8588	0,7622	0,8121
2	0,8446	0,7622	0,8433
3	0,8511	0,8324	0,8639
4	0,9101	0,8763	0,9053
5	0,8978	0,8974	0,9094
Vidurkis:	0,8725	0,8261	0,8668
MIN/MAX	0,8446 / 0,9101	0,7622 / 0,8975	0,8121 / 0,9094
Standartinis nuokrypis:	0,0263	0,0562	0,0370

Prieš pradėdant „U-net“ modelio modernizavimą atliekami skaičiavimai, kurių metu gaunami duomenys, pagal kuriuos yra vertinamas modernizuojamos struktūros pagerėjimas ar pablogėjimas. Gauti rezultatai pateikiami lentelėje (žr. 4.3 lentelę). Šiuo atveju keitėsi tik imtys – jose naudojami atsitiktinai parinkti vaizdai.

Taigi, kaip matoma iš 4.3 lentelėje gautų rezultatų vidutinė tikslumo vertė svyruoja nuo 82,61 iki 87,25 procentų vertės, priklausomai nuo modelio konfigūracijos. Standartinis nuokrypis taip pat yra ganėtinai didelis, kinta nuo 2,6 iki 5,62 procento. Šios gautos vertės naudojamos kaip atliekamo darbo tobulinimo kriterijus.

4.3. „U-net“ modelio modernizavimas

Šiame poskyryje aprašomi rezultatai, gaunami „U-net“ modelio architektūrą papildant blokais. Pirmasis išbandytas modernizavimui skirtas blokas yra erdvinis požymių segmentų atmetimo blokas (angl. *SpatialDropout2D*). Toliau bandymai vyksta naudojant požymių ištraukimo / pastiprinimo blokų (angl. *Squeeze-extract*) elementus, išbandomos 5 skirtingos struktūros (žr. 1.3 poskyryje).

Erdvinis požymių segmentų atmetimo blokas

Šis blokas atlieka panašią funkciją kaip požymių segmentų atmetimo (angl. *Dropout*) blokas [40]. Erdvinis požymių segmentų atmetimo blokas nuo segmentų atmetimo bloko skiriasi tuo, kad vietoje atskirų elementų išmeta „2D“ funkcijų žemėlapius, dėl ko turėtų būti gaunami geresni išvesties rezultatai, jei gretimi elementų žemėlapių taškai yra glaudžiai susiję (kaip paprastai būna ankstyvosios konvoliucijos sluoksniuose). Šiuo atveju erdvinis požymių segmentų atmetimo blokas turėtų padėti skatinti žemėlapių nepriklausomumą [41]. Šios struktūros blokai įtraukti į kiekvieną modelio tinklo sluoksnį.

4.4 lentelė. „U-net“ skaičiavimų rezultatai įtraukus erdvinis požymių segmentų atmetimo blokus

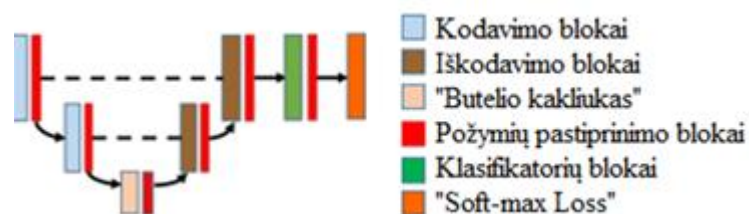
Konfigūracija („U_net_n2“ su „SpatialDropout“)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8305	0,8283	0,8390
2	0,8087	0,8205	0,8158
3	0,8505	0,8030	0,8522
4	0,8952	0,8846	0,8860
5	0,9051	0,8694	0,8940
Vidurkis:	0,8580	0,8412	0,8574
MIN/MAX	0,808 / 0,905	0,803 / 0,869	0,816 / 0,894
Standartinis nuokrypis:	0,0370	0,0308	0,0292

Atliekamiems skaičiavimams kiekvienu atveju naudota 50 iteracijų. Gauti rezultatai matomi 4.4 lentelėje. Lentelės duomenys parodo, kad erdviniai požymių segmentų atmetimo blokai 2 iš 3 modelio konfigūracijų atvejų neleido pasiekti aukštesnių rezultatų. Matomas 1,5-2 proc. tikslumo sumažėjimas. Naudojant „256_5_8“ modelio konfigūraciją gaunamas 2 proc. pagerėjimas. Rezultatų išsibarstymas 2 iš 3 atvejų yra mažesnis lyginant su pradinio „U-net“ modelio rezultatais ir siekė apie 3–3,7 proc.

Požymių pastiprinimo blokai

Požymių pastiprinimo (angl. *Squeeze-extract*) blokas, tai struktūra, kurios tikslas sudaryti požymių tikimybinę kaukę ir naudojamąjį kintamąjį „x“ sudauginti su ja. Tokiu būdu išryškinami svarbiausi nuotraukoje esantys, defekto suradimui būtini, bruožai [10]. Šiais blokais papildoma „U-net“ struktūra taip pat yra papildyta anksčiau aprašytais erdvinio segmentų atmetimo blokais.

Teorinėje dalyje (žr. 1.3 poskyryje) pabrėžiama, kad reikia pasirinkti tinkamą požymių pastiprinimo blokų vietą „U-net“ modelio sandaroje. Išbandomos architektūros atveju požymių pastiprinimo blokai (pavaizduoti raudonai) naudojami visoje „U-net“ struktūroje. Blokai įtraukti į kodavimo, iškodavimo bei centrinę dalis (žr. 4.1 pav.).



4.1 pav. Požymių pastiprinimo blokų įterpimo vieta „U-net“ modelyje

Išbandomi 5 tipų, skirtingi požymių pastiprinimo blokai, siekiant išsiaiškinti, kuris yra tinkamiausias norimai operacijai atlikti, t.y. pagerinti odos defektų išskyrimo tikslumą bei sumažinti duomenų išsibarstymą, gaunamą naudojant pradinį „U-net“ tinklą.

Toliau pateikiamos 5 lentelės, kuriose matomi atliktų segmentacijos užduočių rezultatai. Skaičiavimų sąlygos yra identiškos. Ta pati struktūra naudojama 5 kartus su vis kitomis testavimo, patikrinimo bei mokymui skirtomis imtimis, skaičiavimams atliekama po 50 iteracijų.

Požymių pastiprinimo blokai

Papildyto „U-net“ tinklo eksperimentų rezultatai matomi 4.5; 4.6; 4.7; 4.8 ir 4.9 lentelėse. Kiekvienoje iš jų pateikiami skaičiavimai su 3 modelio konfigūracijomis, naudojant 1 iš 5 tipų požymių pastiprinimo blokų. Naudoti pastiprinimų blokai:

- nedaug parametrų turintis požymių pastiprinimo blokas (rezultatai matomi 4.5 lentelėje);
- erdvinio spaudimo ir kanalų sužadavimo būdu požymius pastiprinantis blokas (rezultatai matomi 4.6 lentelėje);
- erdvinio spaudimo būdu požymius pastiprinantis blokas (rezultatai matomi 4.7 lentelėje);
- erdvinio / kanalų suspaudimo ir sužadavimo būdu požymius pastiprinantis blokas (rezultatai matomi 4.8 lentelėje);
- požymių pastiprinimo blokas, kanalų svarbos pabrėžimas (rezultatai matomi 4.9 lentelėje).

4.5 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE1“, suspaudžiantį bruožų žemėlapius

Konfigūracija (U_net_n3 su SP+SE1)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8655	0,8596	0,8648
2	0,8757	0,8596	0,8852
3	0,8809	0,8666	0,8808
4	0,8842	0,8672	0,8541
5	0,8587	0,8134	0,7874
Vidurkis:	0,8730	0,8533	0,8545
MIN/MAX	0,8587 / 0,8842	0,8134 / 0,8672	0,7874 / 0,8852
Standartinis nuokrypis:	0,009548612	0,020206375	0,035335059

4.6 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE2“, spaudimas vyksta per geometriją

Konfigūracija (U_net_n3 su SP+SE2)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8702	0,8399	0,8593
2	0,8672	0,8058	0,8470
3	0,8575	0,8483	0,8677
4	0,8729	0,8238	0,8600
5	0,8430	0,7827	0,8329
Vidurkis:	0,86216	0,8201	0,8533
MIN/MAX	0,843 / 0,8729	0,7827 / 0,8483	0,8329 / 0,8677
Standartinis nuokrypis:	0,010901486	0,023682145	0,012198918

4.7 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE3“, spaudimą per kanalus

Konfigūracija (U_net_n3 su SP+SE3)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,868	0,8539	0,8612
2	0,8777	0,8571	0,8647
3	0,8872	0,8567	0,8763
4	0,8818	0,8584	0,8721
5	0,8518	0,7982	0,8318
Vidurkis:	0,8733	0,84486	0,86122
MIN/MAX	0,8518 / 0,8872	0,7982 / 0,8584	0,8318 / 0,8763
Standartinis nuokrypis:	0,012451185	0,023376107	0,015642429

Apžvelgus 4.5; 4.6; 4.7, 4.8 ir 4.9 lentelių duomenis matoma, kad į pradinę „U-net“ struktūrą įtraukus erdvinį segmentų atmetimo ir požymių pastiprinimo blokus, priklausomai nuo naudoto bloko, gaunamas tikslumo pagerėjimas lyginant su pradine „U-net“ struktūra. „128_4_32“ modelio konfigūracijos atveju, naudojant, bet kurią iš 5 skirtingų požymių pastiprinimo bloką, išskyrus „SE2“, gaunamas tikslumas siekia ~87 procentus. Nors tai tokia pati vertė kokią pasiekia pradinė struktūra (žr. 4.3 lentelę), standartinis nuokrypis, įtraukus požymių pastiprinimo blokus, sumažinamas ~1,5 procento, kas rodo kad gaunamų rezultatų patikimumas, atkartojamumas yra tikslesnis.

„256_5_8“ konfigūracijos skaičiavimai rodo, kad struktūros, su požymių pastiprinimo blokais, gaunamas rezultatas išlaiko pradinio „U-net“ modelio vertes arba jas pagerina iki 3 procentų. Taip pat standartinis nuokrypis mažėja 3 procentais.

Atlikus skaičiavimus naudojant „256_5_32“ modelio konfigūraciją išgaunamas tikslumas svyruoja tarp 85-86 procentų. Lyginant su pradine struktūra pastebima, kad išvestys išlieka stabilios arba pablogėja 1 procentu, priklausomai nuo to, kuris požymių pastiprinimo blokas yra naudojamas. Rezultatų išsibarstymą 3 iš 5 išbandytų atvejų galima sumažinti iki 2 procentų.

4.8 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE4“, spaudimas per kanalus (papildytas)

Konfigūracija (U_net_n3 su SP+SE4)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8706	0,8136	0,8614
2	0,8699	0,8207	0,8678
3	0,885	0,8495	0,8800
4	0,8716	0,8606	0,8745
5	0,8558	0,8134	0,8439
Vidurkis:	0,87058	0,83156	0,8655
MIN/MAX	0,8558 / 0,885	0,8134 / 0,8606	0,8439 / 0,8800
Standartinis nuokrypis:	0,0092	0,0196	0,0124

4.9 lentelė. Rezultatai naudojant požymių pastiprinimo bloką „SE5“ skaičiavimams

Konfigūracija (U_net_n3 su SP+SE5)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8727	0,8519	0,8636
2	0,877	0,8497	0,8645
3	0,8791	0,7859	0,8751
4	0,8816	0,8195	0,862
5	0,8521	0,7971	0,813
Vidurkis:	0,8727	0,82082	0,85564
MIN/MAX	0,8134 / 0,8672	0,7827 / 0,8483	0,7982 / 0,8584
Standartinis nuokrypis:	0,0106	0,0267	0,0218

Taigi, priklausomai nuo modelio konfigūracijos, visos penkios požymių pastiprinimo bloką struktūros gali pagerinti prieš tai buvusio „U-net“ modelio, su erdvinio požymių segmentų atmetimo bloku, rezultatų atkartojamumą ir mažina jų sklaidą (rezultatai yra atkartojami ir nuo mokymo imties nepriklauso). Kai naudojamas įvesties vaizdas yra 128 x 128 taškų geriausius rezultatus pasiekia struktūra su erdvinio spaudimo būdu požymius pastiprinančiu bloku. Ji leidžia pasiekti iki 87,33 proc. tikslumą ir užtikrina iki 1,2 procento siekiantį rezultatų išsibarstymą. Kai naudojamas iš 256 x 256 taškų sudarytas vaizdas, geriausias rezultatas gaunamas, kai naudojamas erdvinio / kanalų suspaudimo ir sužadavimo būdu požymius pastiprinantis blokas. Pasiekiamas iki 86,55 procento tikslumas pagal „F1“ vertinimo kriterijų bei užtikrinamas iki 1,2 procento standartinis nuokrypis.

Modelių palyginimas

Šiame skyrelyje apžvelgiami rezultatai, kurie leidžia daryti išvadas, ar atlikti pradinio „U-net“ modelio patobulinai statistiškai naudingi. Lentelėse pateikti „F1“ vertinimo kriterijaus verčių vidurkiai taip pat standartiniai nuokrypiai. Šios dvi vertės leidžia įvertinti ar pakeista modelio struktūra pasiekia aukštesnius kokybinius parametrus. Pateikiamas vertinimas pagal Studento kriterijų. Žemiau lentelėse 4.10; 4.11; 4.12 pateikiama informacija / rezultatai. 4.10 lentelėje matomos „128_4_32“ konfigūracijos gaunamos vertės, kurios parodo, kad tik naudojant antrąjį požymių pastiprinimo bloką „SE2“, gaunamas rezultato kritimas žemiau 87 proc. ribos. Visais kitais atvejais gaunama panaši tikslumo vertė kaip pradinės „U-net“ struktūros t.y. ~87 proc. Lyginant pradinį modelį ir „U-net“ su erdvinio požymių segmentu atmetimo bloku („128_5_32“ konfigūracijos atveju) matomas ženklus galutinių verčių pablogėjimas tikslumo klasėje ir standartinio nuokrypio atžvilgiu.

Geriausią pagerėjimo rezultatą gaunami, kai naudojama „258_5_8“ modelio konfigūracija (žr. 4.11 lentelę). Šių skaičiavimų vertės rodo, kad erdvinio požymių segmentų atmetimo ir požymių pastiprinimo blokų įtraukimas duoda teigiamą naudą tikslumui (pagerėjimas iki 3 proc.) arba išlaikoma buvusi vertė. Blokų pridėjimas neblogina struktūros darbo. Įtraukus požymių pastiprinimo blokus standartinis nuokrypis mažėja nuo 2,6 proc iki ~1 proc.

4.10 lentelė. Modelių palyginimas („128_4_32“ tinklo konfigūracija)

Konfigūracija: 128_4_32	Pradinis "U-net" modelis	Su erdvinio požymių segmentų atmetimo bloku "SP"	Su požymių pastiprinimo bloku				
			"SE1"	"SE2"	"SE3"	"SE4"	"SE5"
Tikslumas, pagal F1 koeficientą	0,8725	0,8580	0,8730	0,8622	0,8733	0,8706	0,8725
Standartinis nuokrypis	0,0264	0,0370	0,0095	0,0109	0,0125	0,0093	0,0106
Statinis pagerėjimas (Taip/Ne), lyginant su baziniu modeliu		Ne	Taip	Ne	Taip	Taip	Taip

4.11 lentelė. Modelių palyginimas („256_5_8“ tinklo konfigūracija)

Konfigūracija: 256_5_8	Pradinis "U-net" modelis	Su erdvinio požymių segmentų atmetimo bloku "SP"	Su požymių pastiprinimo bloku				
			"SE1"	"SE2"	"SE3"	"SE4"	"SE5"
Tikslumas, pagal F1 koeficientą	0,8262	0,8412	0,8533	0,8201	0,8449	0,8316	0,8208
Standartinis nuokrypis	0,0563	0,0308	0,0202	0,0237	0,0234	0,0197	0,0268
Statinis pagerėjimas (Taip/Ne), lyginant su baziniu modeliu		Taip	Taip	Ne	Taip	Taip	Ne

Trečiosios konfigūracijos „256_5_32“ skaičiavimų rezultatai pateikiami 4.12 lentelėje. Matoma, kad kai modelio struktūra yra sudėtingiausia, naudojamas 5 gylio ir 32 filtrų modelis, gaunamos vertės yra panašios į „128_4_32“ modelio gaunamus išvesties dydžius – išlaikomas tas pats pradinio

modelio tikslumas arba jis krenta 1 procentu. Erdvinis segmentų atmetimo blokas neduoda norimų pakitimų kokybės gerinime, tikslumas krenta 1,1 procento.

4.12 lentelė. Modelių palyginimas („256_5_32“ tinklo konfigūracija)

Konfigūracija: 256_5_32	Bazinis "U-net" modelis	Su erdvinio požymių segmentų atmetimo bloku "SP"	Su požymių pastiprinimo bloku				
			"SE1"	"SE2"	"SE3"	"SE4"	"SE5"
Tikslumas, pagal F1 koeficientą	0,8669	0,8574	0,8545	0,8534	0,8612	0,8655	0,8556
Standartinis nuokrypis	0,0370	0,0292	0,0353	0,0122	0,0156	0,0125	0,0218
Statinis pagerėjimas (Taip/Ne), lyginant su baziniu modeliu		Ne	Ne	Ne	Ne	Ne	Ne

Taigi, atlikus skirtingų konfigūracijų modelių rezultatų apibendrinimą galima teigti, kad tinkamiausias aprašytos struktūros modelis naudoja 128 x 128 taškų dydžio įėjimo vaizdą. Kai „U-net“ struktūra yra papildyta erdvinio spaudimo būdu požymius pastiprinančiu bloku, kuris leidžia truputį pagerinti tikslumą bei dvigubai sumažina rezultatų išsibarstymą.

4.4. „U-net“ išėjimų sustiprinimas

Šiame poskyryje išbandoma „U-net“ struktūra, kuri skirta sustiprinti tinklo išskleidimo dalies išvestis ir leidžia išsaugoti svarbiausią informaciją iki vidinio paslėptojo sluoksnio [19], [42]. Modifikuotos struktūros vaizdas matomas 4.2 paveiksle.

```

out10 = average([out6, out7, out8, out9])
#out10 = Conv2D(1, (1, 1), activation='sigmoid', name='side_10')(out10)

model = Model(inputs=[inputs], outputs=[out6, out7, out8, out9, out10])
model.compile(optimizer=RMSprop(lr=1e-3), loss=bce_dice_loss, metrics=[dice_loss], loss_weights=[0.1, 0.1, 0.1, 0.1, 0.6])

```

4.2 pav. „U-net“ išėjimo sustiprinimas

Paveiksle pateiktame kode, modelio išvesties rezultatas yra „U-net“ iškodavimo dalies atskirų sluoksnių vidurkis. Klasikinėje struktūroje išvesties vertė, tai paskutiniojo sluoksnio išvestis. Būtent toks pakeitimas leidžia sustiprinti svarbiausios informacijos išsaugojimą. Toliau pateikiami įverčiai, gauti naudojant „U-net“ su sustiprinta išėjimų struktūra. Bandymai atliekami identiškais sąlygomis kaip prieš tai aprašyti 4.2 ir 4.3 poskyriuose.

4.4.1. Pradinis „U-net“ su sustiprinta išėjimų išvestimi

Šiame skyrelyje pateikiami rezultatai (žr. 4.13 lentelę), kurie gaunami naudojant pradinę „U-net“ struktūrą (ta pati kaip 4.2 poskyryje) su sustiprinta išėjimų išvestimi. Bandymai atlikti su 3 modelio konfigūracijomis, kai įvesties vaizdų dydis 256 x 256 taškų ir 128 x 128 taškų dydžio, naudojamas gylis yra 4 arba 5 su 8, 32 filtrais. Pastebima, kad šios struktūros modelio rezultatai yra aukštesni lyginant su pradinės „U-net“ struktūros pasiekiamais tikslumo įverčiais, pateikiamais 4.3 lentelėje. „128_4_32“ konfigūracijos skaičiavimų atveju pasiektas tikslumas yra 88,4 procento, tai ~1,15 proc. geresnis rezultatas lyginant su pirminiu modeliu. Bandymas rodo, kad naudojant „256_5_8“

konfigūracijos struktūrą išvesties tikslumas pagerėja nuo 82,6 iki 87 procentų. Toks pats rezultatas matomas atlikus „256_5_32“ konfigūracijos tyrimą – tikslumas didėja, pasiekiami 88,2 procentinė vertė. Taigi, iš 4.13 lentelėje pateiktų duomenų galima teigti, kad „U-net“ modelio iškodavimo dalies pakeitimai duoda teigiamas savybes, su kuriomis pasiekiamos vertės pradinės struktūros (su sustiprintai išėjimais) atveju viršija prieš tai naudotų modelių pasiektus aukščiausius įverčius.

4.13 lentelė. Pradinio „U-net“ tinklo su daugiasluoksniu išėjimu rezultatai

Konfigūracija (U_net)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8817	0,8748	0,8805
2	0,8879	0,8747	0,8807
3	0,8894	0,8880	0,8936
4	0,8876	0,8766	0,8878
5	0,8730	0,8380	0,8690
Vidurkis:	0,8839	0,8704	0,8823
MIN/MAX	0,8730 / 0,8894	0,8381 / 0,8881	0,8691 / 0,8936
Standartinis nuokrypis:	0,006	0,0017	0,008
Parametrų sk. mln.	2,1444	0,5387	8,5831

4.4.2. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir erdvinio požymių segmentų atmetimo bloku

Šiame skyrelyje pateikiami rezultatai, kurie gaunami, kai prie pradinės struktūros „U-net“ modelio su daugiasluoksniu išėjimu pridedami erdviniai požymių segmentų atmetimo blokai. Tikslumo vertės matomos 4.14 lentelėje. Galima pastebėti, kad gautos išvestys kinta taip pat kaip 4.4 lentelėje atlikto bandymo metu – prastėja. Skaičiavimų metu, naudojamos visos trys skirtingos konfigūracijos ir gaunama ta pati tendencija, rezultatų prastėjimas ~3 procentai. Atliktas šios struktūros papildymas požymių pastiprinimo blokais. Gautos vertės pateikiamos 4.15 lentelėje. Matomas rezultatų pagerėjimas lyginant su struktūra, kurioje buvo naudoti tik erdviniai požymių segmentų atmetimo blokai. Tačiau lentelėje pateikiami duomenys rodo, kad gautos tikslumo vertės yra mažesnės lyginant su 4.13 lentelėje gautomis išvesčių vertėmis, kai naudojama klasikinė „U-net“ struktūra su modifikuotais išėjimais. Todėl, atsižvelgus į rezultatų pablogėjimą, atsisakyta naudoti erdvinius požymių segmentų atmetimo blokus. 4.4.3 skyrelyje pateikiami duomenys, gauti atlikus eksperimentus, struktūrą papildant tik požymių pastiprinimo blokais.

4.14 lentelė. Pradinio „U-net“ tinklo su daugiasluoksniu išėjimu ir „SpatialDropout“ bloku rezultatai

Konfigūracija (U_net+ DP)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8633	0,8472	0,8685
2	0,8699	0,8268	0,8645
3	0,8699	0,8605	0,8780
4	0,8722	0,8404	0,8752
5	0,8529	0,8400	0,8493
Vidurkis:	0,8656	0,8430	0,8671
MIN/MAX	0,8529 / 0,8722	0,8268 / 0,8605	0,8493 / 0,8780
Standartinis nuokrypis:	0,0070	0,01108	0,01
Parametrų sk. mln.	2,1444	0,5387	8,5831

4.15 lentelė. Pradinio „U-net“ tinklo su daugiasluoksniu išėjimu ir „SpatialDropout“, požymių pastiprinimo blokais rezultatai

Konfigūracija (U_net+ DP+SE5)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8774	0,8512	0,8949
2	0,8804	0,8542	0,8721
3	0,8835	0,8417	0,8851
4	0,8800	0,8436	0,8847
5	0,8403	0,8206	0,8521
Vidurkis:	0,8723	0,8423	0,8778
MIN/MAX	0,8403 / 0,8835	0,8206 / 0,8542	0,8521 / 0,8949
Standartinis nuokrypis:	0,0159	0,011	0,012
Parametrų sk. mln.	2,5259	0,6350	10,113

4.4.3. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir požymių pastiprinimo blokais

4.16; 4.17; 4.18; 4.19; 4.20 lentelėse pateikiami bandymų rezultatai, kurių metu naudojamas pradinis „U-net“ tinklas su pakeista iškodavimo dalies struktūra prie jos pridendant 5 skirtingus požymių pastiprinimo blokus. Naudoti požymių pastiprinimo blokai yra analogiški 4.2 poskyryje naudotiems. Lentelėse matomi rezultatai:

- nedaug parametrų turintis požymių pastiprinimo blokas (rezultatai matomi 4.16 lentelėje);
- erdvinio spaudimo ir kanalų sužadavimo būdu požymius pastiprinantis blokas (rezultatai matomi 4.17 lentelėje);
- erdvinio spaudimo būdu požymius pastiprinantis blokas (rezultatai matomi 4.18 lentelėje);
- erdvinio/kanalų suspaudimo ir sužadavimo būdu požymius pastiprinantis blokas (rezultatai matomi 4.19 lentelėje);
- požymių pastiprinimo blokas, kanalų svarbos pabrėžimas (rezultatai matomi 4.20 lentelėje).

4.16 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 1 požymių pastiprinimo bloku „SE1“

Konfigūracija (U_net_+SE1)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8995	0,8822	0,8949
2	0,8919	0,8809	0,8919
3	0,9010	0,8878	0,8987
4	0,8930	0,8886	0,8948
5	0,8749	0,8864	0,8697
Vidurkis:	0,8921	0,8852	0,8900
MIN/MAX	0,8749 / 0,9010	0,8809 / 0,8886	0,8697 / 0,8988
Standartinis nuokrypis:	0,0085	0,0027	0,0101
Parametrų sk. mln.	2,1996	0,5528	8,8031

4.17 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 2 požymių pastiprinimo bloku „SE2“

Konfigūracija (U_net_+SE2)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8817	0,8727	0,8826
2	0,8836	0,8732	0,8813
3	0,8923	0,8872	0,8880
4	0,8834	0,8753	0,8837
5	0,8607	0,8659	0,8600
Vidurkis:	0,8803	0,8749	0,8791
MIN/MAX	0,8607 / 0,8923	0,8659 / 0,8872	0,8600 / 0,8881
Standartinis nuokrypis:	0,0105	0,0069	0,0096
Parametrų sk. mln.	2,1451	0,5391	8,5846

4.18 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 3 požymių pastiprinimo bloku „SE3“

Konfigūracija (U_net_+SE3)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8875	0,8893	0,8927
2	0,8921	0,8812	0,8949
3	0,8965	0,8838	0,8996
4	0,8961	0,8875	0,8954
5	0,8725	0,8455	0,8740
Vidurkis:	0,8890	0,8775	0,8914
MIN/MAX	0,8725 / 0,8965	0,8455 / 0,8893	0,8741 / 0,8997
Standartinis nuokrypis:	0,0088	0,0152	0,0089
Parametrų sk. mln.	2,4701	0,6205	9,8918

4.19 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 4 požymių pastiprinimo bloku „SE4“

Konfigūracija (U_net_+SE4)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8963	0,8841	0,8931
2	0,8894	0,8827	0,8903
3	0,8954	0,8937	0,8983
4	0,8901	0,8897	0,8916
5	0,8803	0,8515	0,8700
Vidurkis:	0,8903	0,8803	0,8887
MIN/MAX	0,8803 / 0,8963	0,8515 / 0,8937	0,8701 / 0,8983
Standartinis nuokrypis:	0,0049	0,0148	0,0094
Parametrų sk. mln.	2,4708	0,6209	9,8933

4.20 lentelė. Pradinis „U-net“ tinklas su daugiasluoksniu išėjimu ir 5 požymių pastiprinimo bloku „SE5“

Konfigūracija (U_net_+SE5)	128_4_32	256_5_8	256_5_32
Bandymo nr.	F1 verte	F1 verte	F1 verte
1	0,8850	0,8845	0,89110
2	0,8885	0,8807	0,89126
3	0,8928	0,8880	0,90037
4	0,8915	0,8831	0,89009
5	0,8698	0,8556	0,86005
Vidurkis:	0,8855	0,8784	0,88658
MIN/MAX	0,8698 / 0,8928	0,8556 / 0,8880	0,8601 / 0,9004
Standartinis nuokrypis:	0,0083	0,0122	0,0136
Parametrų sk. mln.	2,5259	0,6350	10,113

Iš gautų duomenų pastebima, kad į struktūrą įtraukus, bet kurį iš požymių pastiprinimo blokų išskyrus antrąjį „SE2“ gaunami geresni tikslumo rezultatai lyginant su pradinės struktūros modeliu (su pakeista iškleidimo dalimi) be požymių pastiprinimo blokų (žr. 4.13 lentelę). „128_4_32“ ir „256_5_8“ konfigūracijos atveju geriausi įverčiai matomi, kai naudojamas 1 požymių pastiprinimo blokas (žr. 4.16 lentelėje). Gaunamas tikslumas siekia 89,21 ir 88,5 procentus atitinkamai. Naudojant „256_5_32“ modelio konfigūraciją pastebima, kad geriausią tikslumo vertę pagal „F1“ vertinimo kriterijų pasiekia modelis, kuriame naudojamas 3 požymių pastiprinimo blokas – erdvinio spaudimo būdu požymius pastiprinantis blokas. Tikslumas siekia 89,13 procento. Visais trimis modelio konfigūracijų skaičiavimo atvejais rezultatų išsibarstymas neviršiją 1 procento ribos, tai reiškia, kad modelio darbas yra tikslus ir prognozuojamas – nėra atsitiktinio rezultato.

4.4.4. Modelių, naudojančių „U-net“ su giliuoju mokymu iškodavimo dalyje palyginimas

Šiame skyrelyje apžvelgiami duomenys, kurie leidžia daryti išvadas, ar atlikti pradinio „U-net“ modelio (su modifikuota išskleidimo architektūros dalimi) patobulinai užtikrina kokybinį modelio darbo pagerėjimą. Lentelėse pateikti „F1“ vertinimo kriterijaus verčių vidurkiai taip pat standartiniai nuokrypiai, parametrų skaičius. Taip pat panaudojant Stjudento kriterijų įvertinama ar modelis įgauna statistinį pagerėjimą lyginant su nemodifikuota struktūra. Žemiau lentelėse 4.21; 4.22; 4.23 pateikiama informacija / rezultatai. 4.21 lentelėje matomi „128_4_32“ konfigūracijos skaičiavimų įverčiai, kurie parodo, kad naudojant 4 iš 5 požymių pastiprinimo blokų gaunamas rezultatų pagerėjimas. Pasiiekiamas iki 89,21 procentų tikslumas, standartinis nuokrypis neviršija 1 procento vertės ribos. Įtraukti požymių pastiprinimo blokai modelį išplečia nuo 2-18 procentų.

Atlikti skaičiavimai su „256_5_8“ konfigūracija parodo (žr. 4.22 lentelėje), kad „U-net“ struktūrą papildžius, bet kuriuo iš 5 požymių pastiprinimo blokų gaunamas rezultatų pagerėjimas lyginant su modelių be požymių pastiprinimo blokų darbu. Standartinio nuokrypio verčių išsibarstymas svyruoja nuo 0,2-1,52 procento. Ši struktūra turi mažiausiai parametrų, todėl ir skaičiavimų laikui yra mažiausiai imli.

„256_5_32“ modelio konfigūracijos skaičiavimų rezultatai matomi 4.23 lentelėje. Šios kaip ir „128_4_32“ konfigūracijos atveju 4 iš 5 požymių pastiprinimo blokų leidžia pagerinti modelio gaunamus rezultatus. Geriausias gautas įvertis yra 89,13 proc. tikslumo pagal „F1“ vertinimo kriterijų. Standartinis nuokrypis neviršija 1 proc. vertės ribos. Kaip matoma iš lentelėje pateiktų duomenų nors modelis turi beveik 20 kartų daugiau parametrų lyginant su „256_5_8“ modelio konfigūracijos parametrų skaičiumi, gaunami rezultatai yra panašūs.

Visų trijų konfigūracijų atveju į struktūrą įtraukus erdvinius požymių segmentų atmetimo blokus, rezultatai prastėjo, todėl galutiniuose skaičiavimuose požymių atmetimo blokų atsisakyta, siekiant gauti kuo aukštesnius galutinius įverčius. Lyginant patobulintos struktūros geriausią rezultatą, kai naudojama „128_4_32“ konfigūracija t.y. 89,21 procento tikslumą su 4.2 poskyryje pateiktu pradiniu (žr. 4.3 lentelėje) modeliu, gaunamas ~2 procentų tikslumo pagerėjimas (pirminio „U-net“ modelio geriausias rezultatas 87,2 proc).

4.21 lentelė. Modelių palyginimas („128_4_32“ tinklo konfigūracija)

Konfigūracija: 128_4_32	"U-net" modelis	Su erdvinio požymių segmentų atmetimo bloku "SP"	Su požymių pastiprinimo bloku				
			"SE1"	"SE2"	"SE3"	"SE4"	"SE5"
Tikslumas, pagal F1 koeficientą	0,8839	0,8656	0,8921	0,8803	0,8890	0,8903	0,8855
Standartinis nuokrypis	0,0059	0,0070	0,0085	0,0105	0,0088	0,0049	0,0083
Parametrų sk. mln.	2,1444	2,1444	2,1996	2,1451	2,4701	2,4708	2,5259
Statinis pagerėjimas (Taip/Ne), lyginant su baziniu modeliu		Ne	Taip	Ne	Taip	Taip	Taip

4.22 lentelė. Modelių palyginimas („256_5_8“ tinklo konfigūracija)

Konfigūracija: 256_5_8	"U-net" modelis	Su erdvinio požymių segmentų atmetimo bloku "SP"	Su požymių pastiprinimo bloku				
			"SE1"	"SE2"	"SE3"	"SE4"	"SE5"
Tikslumas, pagal F1 koeficientą	0,8704	0,8430	0,8852	0,8749	0,8775	0,8803	0,8784
Standartinis nuokrypis	0,0167	0,0108	0,0027	0,0069	0,0152	0,0148	0,0112
Parametrų sk. mln.	0,5387	0,5388	0,5528	0,5391	0,6205	0,6209	0,6350
Statinis pagerėjimas (Taip/Ne), lyginant su baziniu modeliu		Ne	Taip	Taip	Taip	Taip	Taip

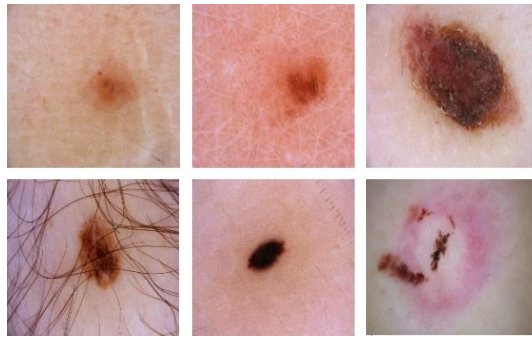
4.23 lentelė. Modelių palyginimas („256_5_32“ tinklo konfigūracija)

Konfigūracija: 256_5_32	"U-net" modelis	Su erdvinio požymių segmentų atmetimo bloku "SP"	Su požymių pastiprinimo bloku				
			"SE1"	"SE2"	"SE3"	"SE4"	"SE5"
Tikslumas, pagal F1 koeficientą	0,8823	0,8671	0,8900	0,8791	0,8913	0,8887	0,8865
Standartinis nuokrypis	0,0082	0,0100	0,0101	0,0096	0,0089	0,0094	0,0135
Parametrų sk. mln.	8,5831	8,5831	8,8031	8,5846	9,8918	9,8933	10,1131
Statinis pagerėjimas (Taip/Ne), lyginant su baziniu modeliu		Ne	Taip	Ne	Taip	Taip	Taip

4.5. Daugiaklasė segmentacija

Šiame poskyryje atliekama daugiaklasė segmentacija. Segmentacija atliekama naudojant vieną geriausių rezultatų parodžiusią modernizuoto „U-net“ modelio struktūrą. Šiuo atveju pasirinktas naudoti „256_5_32“ konfigūracijos modelis su 1 požymių pastiprinimo bloku „SE1“. Jis pasiekė iki 89,13 procento tikslumo rezultatą bei yra viena daugiausiai parametrų, turinti struktūra. Didelis parametrų skaičius lemia lėtesnius skaičiavimus. Atliekant daugiaklasę segmentaciją tikimasi gauti 5 skirtingas klases iš pateikto vaizdo. Naudojami tie patys 2594 vaizdai (duomenų pvz. matomi 4.3 pav.), kurie yra skirti melonomos darinių segmentavimui atlikti. Išskiriamos 5 klasės:

- neigiama (1 klasė);
- rutuliai (įskaitant ir taškus) (2 klasė);
- cistos (3 klasė);
- pigmento tinklas (4 klasė);
- dryželiai (5 klasė).



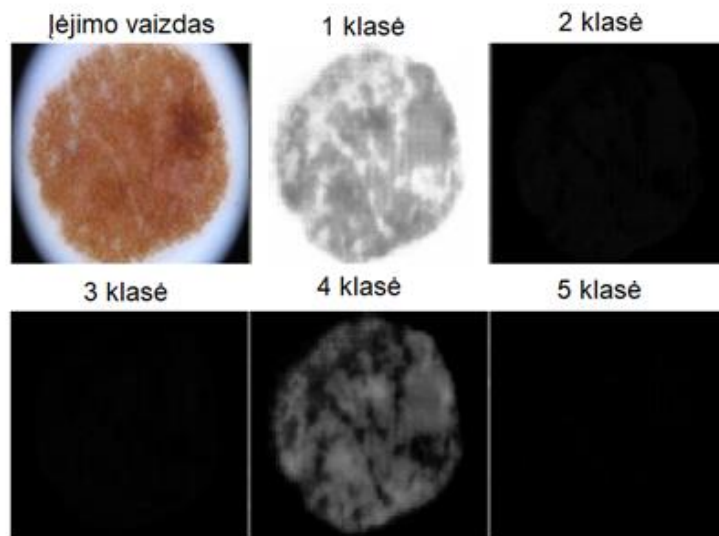
4.3 pav. Daugiaklasei segmentacijai naudojami duomenys (mokymui)

Atliekant segmentaciją visa nuotrauka suskirstoma į spalvų sritis nuo 0 iki 200. Kiekviena spalvų sritis priskirta tam tikrai klasei:

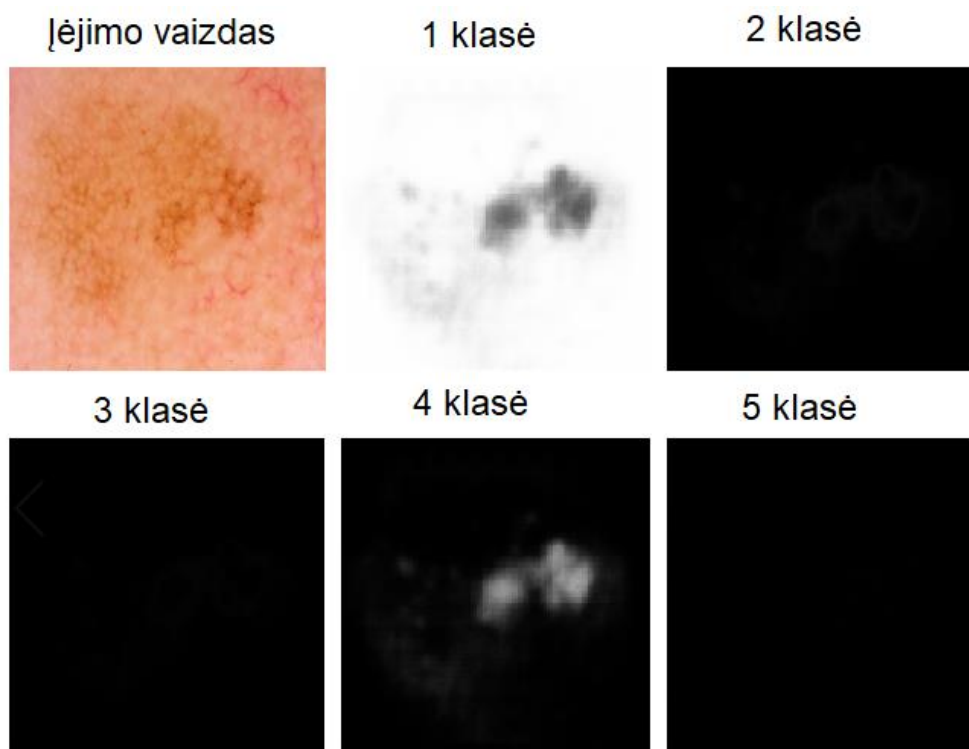
- 1 klasė: 0;
- 2 klasė: 50;
- 3 klasė: 100;
- 4 klasė: 150;
- 5 klasė: 200.

Atliekant mokymo procesą naudojamos melanomos vaizdų nuotraukos ir kaukės, kurios yra įvairių dydžių. Visų pirma prieš pradedant mokymo procesą visi turimi duomenys yra suvienodinami, paruošiami mokymo procesui. Visos 5 klasės priskiriamos spalvų sritims. Atliekamas 50 epochų mokymo procesas. Galiausiai atliekamas spėjimo veiksmas su patvirtinimui skirta vaizdų imtimi.

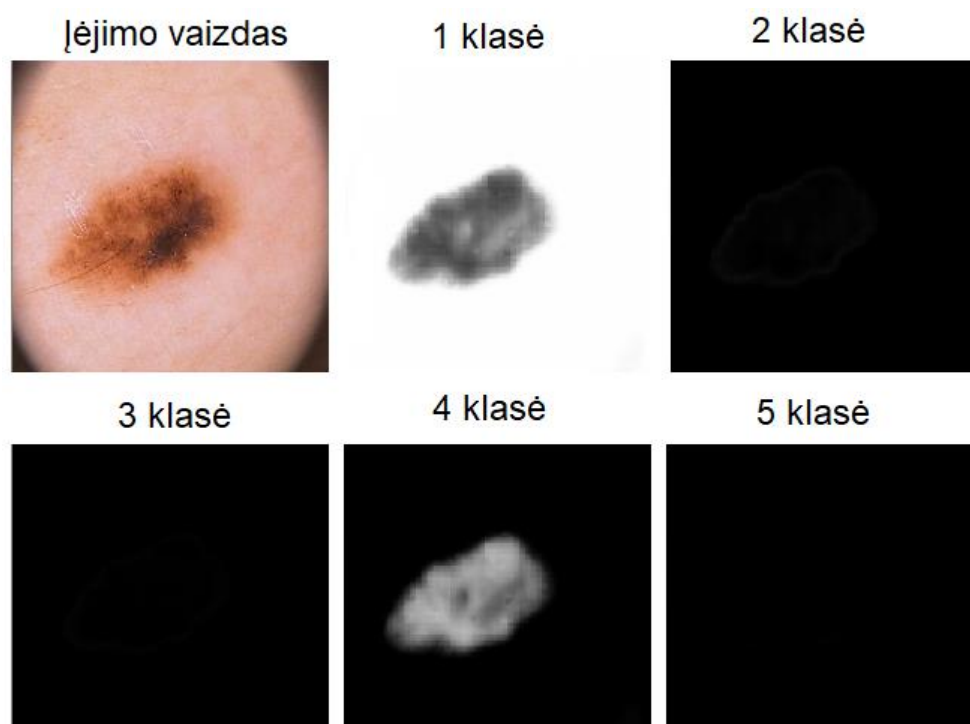
Toliau pateikiamos nuotraukos (žr. 4.4 pav., 4.5 pav., 4.6 pav.), kuriose vizualiai matomi gaunami rezultatai – vaizdo segmentų priskyrimas tam tikrai klasei pagal požymius. Gauta išvestis parodo, kad naudojant šios struktūros modelį galima atlikti daugiaklasę segmentaciją. Paveiksle matomos visos 5 klasės. 3 ir 5 klasių struktūrų vaizdai yra labai neryškūs, nes segmentų plotas yra labai mažas. 1, 2 ir 4 klasės išskyrimas yra pakankamai kokybiškai matomas, o tai leidžia daryti išvadą, kad naudojama modifikuota „U-net“ struktūra leidžia atlikti skirtingų klasių klasifikavimą. Šie rezultatai užtikrina galimybę atlikti sudėtingesnę vaizdo analizę.



4.4 pav. Daugiaklasės segmentacijos rezultatai (1 pvz.)



4.5 pav. Daugiaklasēs segmentācijas rezultāti (2 pvz.)



4.6 pav. Daugiaklasēs segmentācijas rezultāti (3 pvz.)

Išvados

1. Atliktas pradinio „U-net“ modelio parametrų parinkimas melanomos segmentacijos užduočiai atlikti. Išskiriamos 3 aukščiausius įverčius pagal „F1“ vertinimo kriterijų pasiekiančios struktūros. Pirmoji, tai 4 sluoksnių modelio architektūra su 32 filtrais, kai įėjimo vaizdo dydis yra 128 x 128 x 3 taškų, gaunamas tikslumas iki 87,25 proc. Kitos dvi aukščiausios „F1“ vertinimo kriterijaus vertės pasiekiančios modelio struktūros sudarytos iš 5 sluoksnių, kai naudojamas filtrų dydis 8 ir 32, o įėjimo vaizdo dydis 256 x 256 x 3 taškų. Pasiekiamos „F1“ vertės 82,61 ir 86,68 procentai;
2. Atliktas melanomos segmentacijos tikslumo tyrimas, į giliojo mokymo modelį įtraukus erdvinio segmentų atmetimo ir požymių pastiprinimo blokus. Gaunami rezultatai skiriasi priklausomai, nuo naudotos modelio konfigūracijos. „128_4_32“ konfigūracijos skaičiavimų įverčiai rodo, kad tik naudojant antrąjį požymių pastiprinimo bloką „SE2“, gaunamas rezultato kritimas žemiau 87 proc. ribos. Visais kitais atvejais gaunama panaši tikslumo vertė kaip pradinės „U-net“ struktūros t.y. ~87 proc. Lyginant pradinį modelį ir „U-net“ su erdvinio požymių segmentų atmetimo bloku („128_5_32“ konfigūracijos atveju) matomas ženklus rezultatų pablogėjimas tikslumo klasėje ir standartinio nuokrypio atžvilgiu – didėjimas.

„258_5_8“ konfigūracijos rezultatai rodo, kad erdvinio požymių segmentų atmetimo ir požymių pastiprinimo blokų įtraukimas duoda teigiamą poveikį tikslumui (pagerėjimas iki 3 proc.) arba yra išlaikoma pradinė „U-net“ tinklo vertė. Įtraukus požymių pastiprinimo blokus standartinis nuokrypis mažėja nuo 2,6 proc iki ~1 proc.

Iš trečiosios konfigūracijos „256_5_32“ skaičiavimo rezultatų matoma, kad kai modelio struktūra yra sudėtingiausia, naudojamas 5 gylio ir 32 filtrų modelis, gaunami rezultatai yra panašūs į „128_4_32“ modelio gaunamus išvesties dydžius – išlaikomas tas pats standartinio modelio tikslumas arba jis krenta 1 procentu. Maksimali tikslumo vertė 86,55 proc. Erdvinis požymių segmentų atmetimo blokas neduoda norimų rezultatų kokybės gerinimui, tikslumas krenta 1,1 procento lyginant su pradinės struktūros modeliu;

3. Atliktas melanomos segmentacijos tikslumo tyrimas, kai naudojamas gilus mokymo struktūros „U-net“ modelis, jį papildant blokais. „128_4_32“ konfigūracijos skaičiavimų įverčiai, parodo, kad naudojant 4 iš 5 požymių pastiprinimo blokų gaunamas rezultatų pagerėjimas. Pasiekiamas iki 89,21 procentų tikslumas, standartinis nuokrypis neviršija 1 procento vertės ribos.

Atlikti skaičiavimai su „256_5_8“ konfigūracija rodo, kad į „U-net“ struktūrą įtraukus, bet kurią iš 5 požymių pastiprinimo blokų gaunamas rezultatų pagerėjimas lyginant su modelių be požymių pastiprinimo blokų darbu. Rezultatų išsibarstymas svyruoja nuo 0,2-1,52 procento.

Iš „256_5_32“ modelio konfigūracijos skaičiavimų rezultatų matoma, kad šios bei „128_4_32“ konfigūracijų atveju 4 iš 5 požymių pastiprinimo blokų leidžia pagerinti modelio gaunamus rezultatus. Geriausias gautas įvertis yra 89,13 proc. tikslumo. Standartinis nuokrypis svyruoja iki 1 proc. ribos vertės. Šis modelis turi beveik 20 kartų daugiau parametrų lyginant su „256_5_8“ modelio konfigūracijos parametrų skaičiumi, bet jo gaunami rezultatai yra panašūs.

Visų trijų konfigūracijų atvejų į struktūrą įtraukus erdvinio požymių segmentų atmetimo blokus, rezultatai prastėjo, todėl galutiniuose skaičiavimuose jų atsisakyta, siekiant gauti kuo aukštesnius galutinius įverčius. Lyginant šios patobulintos struktūros geriausią rezultatą, kai naudojama „128_4_32“ konfigūracija t.y. 89,21 procento tikslumą su 4.2 poskyryje pateiktu pirminiu modeliu, matoma kad galutinį rezultatą pavyksta pagerinti ~2 proc. Visų trijų konfigūracijų atveju požymių pastiprinimo blokai giliojo mokymo modelį išplečia ~2-18 procentais, priklausomai nuo pasirinktos struktūros;

4. Atlikta daugiaklasė segmentacija naudojant daugiausiai parametrų turintį „256_5_32“ konfigūracijos modelį. Gautas rezultatas yra išskirtos 5 skirtingos klasės iš 1 vaizdo, kuriame matomas melanomos darinys. 2 iš 5 klasių yra sunkiau matomos, nes išskiriamas darinys yra labai mažas. Pagrindinis išskirtas vaizdas – melanomos struktūra yra geros kokybės. Gauti rezultatai yra tinkami sudėtingai vaizdo analizei atlikti vykdant atskirų segmentų klasifikavimą.

Literatūros sąrašas

1. AGATONOVIC-KUSTRIN, S.; BERESFORD, Rosemary. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of pharmaceutical and biomedical analysis*, 2000, 22.5: 717-727. [Žiūrėta 2021.01.09];
2. Melanoma. [Žiūrėta 2021.01.09]. Prieiga per internetą: <https://www.iveikvezi.lt/melanoma/>;
3. Dirbtinis intelektas. [Interaktyvus]. [Žiūrėta 2021.01.09]. Prieiga per internetą: <https://www.ibm.com/cloud/learn/neural-networks>;
4. Gilusis mokymasis. [Interaktyvus]. [Žiūrėta 2021.01.09]. Prieiga per internetą: <https://builtin.com/artificial-intelligence>;
5. Gilusis mokymasis. [Interaktyvus]. [Žiūrėta 2021.01.09]. Prieiga per internetą: <https://www.ibm.com/cloud/learn/deep-learning>;
6. PRABHU, Raghav. Understanding of convolutional neural network (CNN)—deep learning. A Medium Corporation, US, 2018. [Žiūrėta 2021.01.09];
7. Užpildymų tipai. [Interaktyvus]. [Žiūrėta 2021.01.20]. Prieiga per internetą: <https://heartbeat.fritz.ai/deep-learning-for-image-segmentation-u-net-architecture-ff17f6e4c1cf>;
8. BALDI, Pierre. Autoencoders, unsupervised learning, and deep architectures. In: *Proceedings of ICML workshop on unsupervised and transfer learning. JMLR Workshop and Conference Proceedings*, 2012. p. 37-49. [Žiūrėta 2021.01.20];
9. RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015. p. 234-241. [Žiūrėta 2021.01.20];
10. HU, Jie; SHEN, Li; SUN, Gang. Squeeze-and-excitation networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018. p. 7132-7141. [Žiūrėta 2021.01.20];
11. ROY, Abhijit Guha; NAVAB, Nassir; WACHINGER, Christian. Recalibrating fully convolutional networks with spatial and channel “squeeze and excitation” blocks. *IEEE transactions on medical imaging*, 2018, 38.2: 540-549. [Žiūrėta 2021.02.07];
12. BOUSABARAH, Khaled, et al. Deep convolutional neural networks for automated segmentation of brain metastases trained on clinical data. *Radiation Oncology*, 2020, 15: 1-90. [Žiūrėta 2021.04.26];
13. ROY, Abhijit Guha; NAVAB, Nassir; WACHINGER, Christian. Concurrent spatial and channel ‘squeeze & excitation’ in fully convolutional networks. In: *International conference on medical image computing and computer-assisted intervention*. Springer, Cham, 2018. p. 421-429. [Žiūrėta 2021.02.07];
14. QILONG, Wang, et al. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. 2020. [Žiūrėta 2021.03.16];
15. GOULD, Stephen, et al. Multi-class segmentation with relative location prior. *International Journal of Computer Vision*, 2008, 80.3: 300-316. [Žiūrėta 2021.03.16];
16. GOYAL, Manu; YAP, Moi Hoon; HASSANPOUR, Saeed. Multi-class semantic segmentation of skin lesions via fully convolutional networks. *arXiv preprint arXiv:1711.10449*, 2017. [Žiūrėta 2021.03.16];
17. MARTINEZ, Manuel; STIEFELHAGEN, Rainer. Taming the cross entropy loss. In: *German Conference on Pattern Recognition*. Springer, Cham, 2018. p. 628-637. [Žiūrėta 2021.03.16];
18. GU, Zaiwang, et al. Ce-net: Context encoder network for 2d medical image segmentation. *IEEE transactions on medical imaging*, 2019, 38.10: 2281-2292. [Žiūrėta 2021.03.16];

19. SHI, Ran; NGAN, King Ngi; LI, Songnan. Jaccard index compensation for object segmentation evaluation. In: 2014 IEEE international conference on image processing (ICIP). IEEE, 2014. p. 4457-4461. [Žiūrėta 2021.03.23];
20. BOOS, Dennis D.; STEFANSKI, Leonard A. P-value precision and reproducibility. *The American Statistician*, 2011, 65.4: 213-221. [Žiūrėta 2021.03.23];
21. NARKHEDE, Sarang. Understanding auc-roc curve. *Towards Data Science*, 2018, 26: 220-227. [Žiūrėta 2021.03.23];
22. WHITE, Brandyn; SHAH, Mubarak. Automatically tuning background subtraction parameters using particle swarm optimization. In: 2007 IEEE international conference on multimedia and expo. IEEE, 2007. p. 1826-1829. [Žiūrėta 2021.03.23];
23. Geriausiai rezultatus pasiekientys modeliai, naudojantys „ISIC18“ duombazę. [Žiūrėta: 2021.03.23]. Prieiga per internetą: <https://github.com/rezazad68/BCDU-Net>;
24. GIRSHICK, Ross, et al. Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 2015, 38.1: 142-158. [Žiūrėta 2021.03.23];
25. ALOM, Md Zahangir, et al. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*, 2018. [Žiūrėta 2021.03.25];
26. AZAD, Reza, et al. Bi-directional convlstm u-net with densley connected convolutions. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. 2019. p. 0-0. [Žiūrėta 2021.03.25];
27. KIM, Seongchan, et al. Deeprain: ConvLstm network for precipitation prediction using multichannel radar data. *arXiv preprint arXiv:1711.02316*, 2017. [Žiūrėta 2021.03.25];
28. LI, Jing, et al. Study on Strategy of CT Image Sequence Segmentation for Liver and Tumor Based on U-Net and Bi-ConvLSTM. *Expert Systems with Applications*, 2021, 115008. [Žiūrėta 2021.03.25];
29. ASADI-AGHBOLAGHI, Maryam, et al. Multi-level context gating of embedded collective knowledge for medical image segmentation. *arXiv preprint arXiv:2003.05056*, 2020. [Žiūrėta 2021.03.25];
30. „ISIC18“ duomenų bazė. [Interaktyvus]. [Žiūrėta: 2021.04.25]. Prieiga per: <https://challenge2018.isic-archive.com/>;
31. Melanomos vaizdas. [Interaktyvus]. [Žiūrėta: 2021.04.25]. Prieiga per: <https://challenge2018.isic-archive.com/task1/>;
32. WEI, Dong, et al. Learning and Exploiting Interclass Visual Correlations for Medical Image Classification. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2020. p. 106-115. [Žiūrėta 2021.04.26];
33. RAZZAK, Imran; NAZ, Saeeda. Unit-wise: Deep Shallow Unit-Vise Residual Neural Networks with Transition Layer For Expert Level Skin Cancer Classification. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2020. [Žiūrėta 2021.04.26];
34. REZVANTALAB, Amirreza; SAFIGHOLI, Habib; KARIMIJESHNI, Somayeh. Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms. *arXiv preprint arXiv:1810.10348*, 2018. [Žiūrėta 2021.05.03];
35. ALOM, Md Zahangir, et al. Recurrent residual U-Net for medical image segmentation. *Journal of Medical Imaging*, 2019, 6.1: 014006. [Žiūrėta 2021.03.25];

36. TSCHANDL, Philipp; SINZ, Christoph; KITTLER, Harald. Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation. *Computers in biology and medicine*, 2019, 104: 111-116. [Žiūrėta 2021.04.25];
37. AL-MASNI, Mohammed A., et al. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Computer methods and programs in biomedicine*, 2018, 162: 221-231. [Žiūrėta 2021.04.25];
38. KAYMAK, Ruya; KAYMAK, Cagri; UCAR, Aysegul. Skin lesion segmentation using fully convolutional networks: A comparative experimental study. *Expert Systems with Applications*, 2020, 161: 113742. [Žiūrėta 2021.04.25];
39. RODRIGUEZ, Juan D.; PEREZ, Aritz; LOZANO, Jose A. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE transactions on pattern analysis and machine intelligence*, 2009, 32.3: 569-575. [Žiūrėta 2021.04.26];
40. MAI, Xiaochun, et al. A Density Map Estimation Model with DropBlock Regularization for Clustered-Fruit Counting. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019. p. 5512-5519. [Žiūrėta 2021.04.26];
41. LEE, Sanghun; LEE, Chulhee. Revisiting spatial dropout for regularizing convolutional neural networks. *Multimedia Tools and Applications*, 2020, 79.45: 34195-34207. [Žiūrėta 2021.04.26];
42. KHANH, Trinh Le Ba, et al. Enhancing U-Net with Spatial-Channel Attention Gate for Abnormal Tissue Segmentation in Medical Imaging. *Applied Sciences*, 2020, 10.17: 5729. [Žiūrėta 2021.04.28].