



Vilniaus universitetas
Duomenų mokslo ir skaitmeninių
technologijų institutas
L I E T U V A



INFORMATIKOS INŽINERIJA (T007)

AKIES DUGNO NUOTRAUKŲ
SEGMENTAVIMAS NAUDOJANT
DIRBTINIUS NEURONINIUS TINKLUS

Ričardas Toliušis

2019 m. spalį

Mokslinė ataskaita DMSTI-DS-T007-19-10

VU Duomenų mokslo ir skaitmeninių technologijų institutas, Akademijos g. 4,

Vilnius LT-08412

www.mii.lt

Santrauka

Akies dugno nuotraukų analizė yra populiari ir perspektyvi dirbtinio intelekto sritis. Ypač svarbus analizės etapas yra kraujagyslių tinklo išskyrimas, kuriam taikomi įvairūs segmentavimo algoritmai. Aptikus kraujagyslių pokyčius galima diagnozuoti įvairias ligas, tokias kaip glaukomą, hipertenziją, diabetinę retinopatiją, makulos degeneraciją ir t. t. Segmentavimui atlikti puikiai tinka konvoliuciniai neuroniniai tinklai, ypač U-Net architektūros. Pastaruoju metu buvo sukurta nemažai U-Net modifikacijų, kurios pasiekia puikius efektyvumo rezultatus.

Reikšminiai žodžiai: gilusis mokymas, dirbtiniai neuroniniai tinklai, U-net, semantinis segmentavimas, akies dugno nuotrauka, akies kraujagyslių išskyrimas

Turinys

1	Įvadas	4
2	Akies vaizdų analizės problematika.....	4
3	Dirbtinių neuroninių tinklų architektūros	5
4	Akies vaizdų semantinis segmentavimas.....	8
5	Išvados	11
6	Literatūra.....	12

1 Įvadas

Automatizuota biomedicininė vaizdų analizė yra aktuali ir perspektyvi dirbtinio intelekto šaka, kuri padeda ypač pagerinti įvairių susirgimų diagnozavimą. Viena medicinos sričių, kurioje tokia analizė gali palengvinti gydytojų darbą, yra akies dugno nuotraukų tyrimas. Iš akies dugno nuotraukų yra nustatomi įvairūs susirgimai, pvz., glaukoma, hipertenzija, diabetinė retinopatija, makulos degeneracija ir kt.

Akies dugno vaizdų analizėje tyrimo objektais gali būti skirtingos akies dalys – kraujagyslės, optinis diskas, regos nervas, makula, lęšiukas. Todėl vaizdų analizėje esminiu momentu tampa semantinis vaizdų segmentavimas, kurio metu kiekvienas vaizdo pikselis yra priskiriamas tam tikrai objekto klasei. Tam yra naudojami įvairūs algoritmai, tačiau pastaruoju metu įsitvirtina dirbtiniai neuroniniai tinklai, o ypač – konvoliuciniai neuroniniai tinklai (angl. *Convolutional Neural Networks*, CNN). Didelį populiarumą įgijo U-Net architektūra, kuri nuo pat pradžių buvo kuriama biomedicininė vaizdų analizei. U-Net pagrindu buvo sukurta nemažai skirtingų modifikacijų, kurių tikslumas viršija 90 proc.

Standartinę akies dugno nuotraukų analizę galima padalinti į keturias fazes: pirminis vaizdo apdorojimas, objektų išskyrimas, požymių išgavimas ir klasifikavimas. Tačiau sudėtingose architektūrose tų fazių gali būti ir daugiau, jos gali būti atliekamos lygiagrečiai.

2 Akies vaizdų analizės problematika

Kalbant apie biomedicininis duomenis, viena pagrindinių problemų yra duomenų priėjimo ribotumas. Realiose situacijose neuroninio tinklo mokymui reikia panaudoti tūkstančius anotuotų vaizdų, o laisvai prieinamose akies dugno nuotraukų duomenų bazėse, pvz., DRIVE ar STARE, jų yra vos po kelias dešimtis. Siekiant padidinti vaizdų aibę yra naudojama duomenų augmentacija. Dažniausias taikomas augmentacijos metodas yra geometrinės vaizdų transformacijos. M. Frid-Adar ir kt. [1] sintetiniams duomenų augmentavimui panaudojo dalinį konvoliucinį neuroninį tinklą GAN (angl. *Generative Adversarial Network*). Naudodamas kompiuterinės tomografijos kepenų vaizdus tinklas generuoja manipuliuodamas įvairiomis augmentavimo technikomis – pasukimai, apvertimai, mastelio keitimas. W. Xianchengas ir kt. [2] DRIVE duomenų bazės padidinimui nuo 20 vaizdų iki 190 000 panaudojo pilkos skalės konvertavimą, riboto kontrasto adaptyvųjį histogramos išlyginimą (CLAHE) ir gamos korekcijas.

Akies dugno nuotraukų prieinamumą ilgą laiką lėmė ir tai, jog jos dažniausiai daromos medicinos įstaigose labai brangia įranga, tačiau jau yra sukurta ir sąlygiškai nebrangių nešiojamųjų kamerų, pavyzdžiui, tokias kameras gamina „Optomed“, „Zeiss“, „Volk Microclear“ ir kt. O „D-Eye“ sukūrė priedą „iPhone“ išmaniesiems telefonams, kuris jį paverčia akies dugno kamera. B. Shen ir Sh. Mukai [3] pristatė mažiau nei 200 dolerių kainuojantį „Rapsbery Pi“ kompiuteriu valdomą kameros prototipą, kuris akį apšviečia infraraudonaisiais spinduliais ir balta šviesa. Naudojant nešiojamuosius fotoaparatus iš dalies galima spręsti vaizdų priėjimo ribotumo problemą, nes atsiranda gerokai daugiau galimybių fotografuoti potencialius ligonius. Tačiau šiuo atveju susiduriama su kita problema – vaizdų kokybe. Nešiojamieji fotoaparatai negali išgauti tokios kokybės nuotraukų, kokia išgaunama naudojant profesionalius stacionarius fotoaparatus. Net su profesionaliais fotoaparatais darytos nuotraukos gali būti netinkamos naudoti. Vaizdų kokybė priklauso nuo fotografavimo procedūros, operatoriaus, apšvietimo, fokusavimo, okliuzijų, plačiai išplitusių pažeidimų, artefaktų. Taip pat sudėtinga nufotografuoti kūdikių ir mažų vaikų akis, jų

nuotraukų kokybė būna prastesnė nei suaugusiųjų [4]. Tad vienas iššūkių yra apdoroti įėjimo aibės vaizdus ir nustatyti, kurie jų yra tinkami semantiniam segmentavimui.

A. Flemingas ir kt. [5] vaizdo ryškumą nustatė identifikuodami kraujagysles, o matymo lauką – centruodami makulos regioną, apimdami visą optinį diską ir palikdami mažiausiai du optinio disko skersmenis nuo vaizdo krašto.

Vienas pagrindinių objektų, kurį reikia identifikuoti ir išskirti, yra akies kraujagyslės, kurių pokyčiai stebimi nustatant ligas. Atliekant kraujagyslių tinklo išskyrimą susiduriama su tokiomis problemomis, kuomet arti esančios kraujagyslės gali būti susiliejusios į vieną, dalis plonesnių kraujagyslių pranyksta, tam tikrose vietose yra per silpnai ar per stipriai apšviestos, matomas šviesos spindulys. Be to, sudėtinga atskirti kraujagysles nuo venų, kas yra būtina atliekant kraujagyslių ir venos santykio matavimą.

3 Dirbtinių neuroninių tinklų architektūros

Biomediciniinių vaizdų segmentavimui pastaruoju metu ypač sėkmingai panaudojami algoritmai, grįsti dirbtiniais neuroniniais tinklais, o konkrečiau – konvoliuciniais neuroniniais tinklais. Nors pirmasis sėkmingai įgyvendintas 7 lygių konvoliucinis neuroninis tinklas LeNet-5 buvo sukurtas dar 1998-aisiais [6], konvoliuciniai neuroniniai tinklai dėl savo dydžio ir poreikio dideliems kompiuterio resursams ilgą laiką buvo šešėlyje ir retai naudojami tyrimuose. Tačiau pastarąjį dešimtmetį ženkliai padidėjus kompiuterių skaičiavimo galiai atsirado daugiau galimybių panaudoti CNN vaizdų tyrimams.

Didžiausias proveržis įvyko 2012 m., kuomet A. Krizhevsky ir kt. [7] pristatė AlexNet tinklą, kuris buvo sudarytas iš 5 konvoliucinių sluoksnių ir 3 visiškai sujungtų sluoksnių. ReLU sluoksniu buvo taikomas po kiekvieno konvoliucinio bei visiškai sujungto sluoksnio, o išmetimo (angl. *Dropout*) sluoksniu po pirmojo ir antrojo visiškai sujungto sluoksnio. AlexNet dideliu skirtumu nugalėjo vaizdų atpažinimo konkurse „ImageNet ILSVRC challenge 2012“. Nuo šio momento CNN smarkiai išpopuliarėjo sprendžiant vaizdų atpažinimo ir klasifikavimo uždavinius, o pastaruosius keletą metų itin aktyviai pradėti naudoti ir semantiniam vaizdų segmentavimui.

„Google“ komanda ILSVR 2014 konkursui pristatė tinklą GoogLeNet, dar vadinamą Inception V1 [8], kuris konkurse užėmė pirmą vietą. GoogLeNet kurtas remiantis LeNet architektūra, papildomai įtraukiant naują elementą – Inception modulį, naudojamą normalizavimui, vaizdo distorcijoms ir RMSProp funkcijai. Inception modulis yra tarsi atskiras tinklas, turintis tris skirtingo dydžio konvoliucinius sluoksnius (1x1, 3x3, 5x5) ir vieną sujungimo sluoksnį, todėl išėjime paduodami skirtingo tipo požymiai. Į Inception modulį papildomai buvo įtraukta 1x1 dydžio konvoliucijos, todėl ženkliai sumažėjo dimensijos ir skaičiavimo operacijų skaičius. Šiuo metu yra sukurta 4-oji Inception versija.

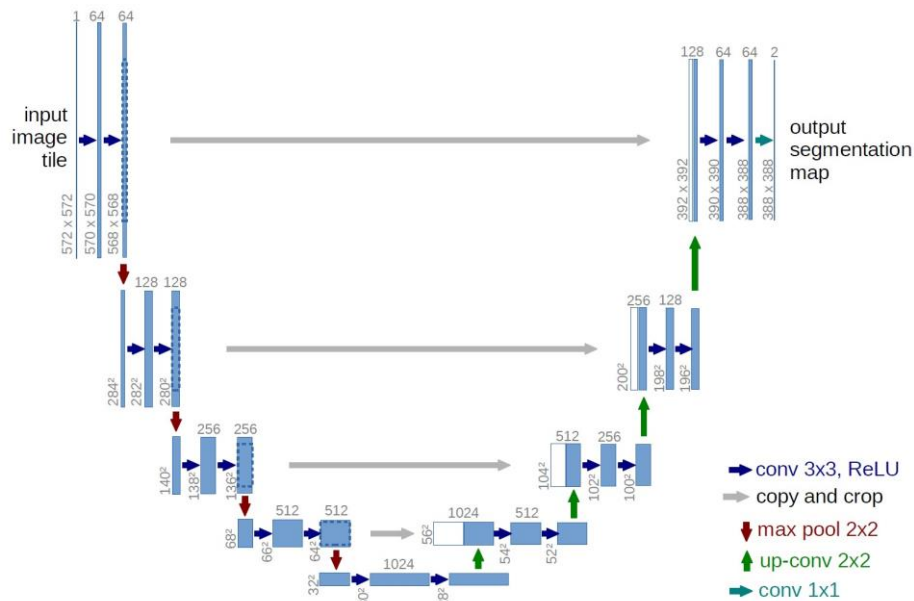
Viena garsiausių šiuo metu architektūrų laikoma VGGNet, kurią sukūrė K. Simonyan ir A. Zissermanas [9]. ILSVRC 2014 konkurse šis tinklas pasižymėjo labai gera sparta ir užėmė antrą vietą. Buvo sukurtos kelios šio tinklo versijos, geriausių tikslumo rezultatus demonstravo 16 sluoksnių turinti versija. Ši architektūra turi apie 140 milijonų parametrų, kurių dauguma yra pirmajame visiškai sujungtame sluoksnyje. Eksperimentų būdu buvo nustatyta, jog visiškai sujungti sluoksniai gali būti pašalinti iš CNN architektūros, taip neprarandant tinklo efektyvumo, tačiau ženkliai sumažinant reikiamų parametrų kiekį. VGGNet yra viena populiariesnių architektūrų, naudojamų kituose moksliniuose tyrimuose, ją modifikuojant ar kombinuojant su kitais metodais.

Vaizdų atpažinimo srityje pradėjus naudoti giliuosius neuroninius tinklus, jie buvo pritaikyti ir semantiniam segmentavimui. Vienas pirmųjų reikšmingesnių darbų buvo atliktas 2015 metais, kai J. Longas ir kt. [10] aprašė visiškai konvoliucinių neuroninių tinklų (angl. *fully convolutional network*, FCN) architektūros panaudojimą semantiniam segmentavimui. Skirtingai nuo klasikinių konvoliucinių neuroninių tinklų, FCN nenaudoja visiškai sujungtų sluoksnių (angl. *fully connected layers*). Tai leidžia neprisirišti prie absoliučių dimensijų reikšmių, o operuoti santykinėmis, todėl tinklas gali atlikti segmentavimą skirtingo dydžio vaizdams. Taip pat tokios architektūros tinklas pasižymi gerokai didesne sparta nei tipinis konvoliucinis tinklas. Lyginant su AlexNet, FCN segmentavimą atlieka nuo 5 iki 25 kartų greičiau.

V. Badrinarayananas ir kt. [11] semantiniam vaizdų žymėjimui atlikti sukūrė konvoliucinio tinklo variaciją SegNet, kuri neturi visiškai sujungtų sluoksnių. SegNet remiasi enkoderio–dekoderio principu. Enkoderis – tai 13 konvoliucinių sluoksnių ir vieną sujungimo sluoksnį turintis VGG grįstas tinklas, skirtas sumažinti ir požymių žemėlapiui sukurti, o dekoderis atlieka padidrinimo funkciją, sukurdamas išretintą originalaus vaizdo matmenų požymių žemėlapi.

Biomedicininį vaizdų segmentavimui O. Ronnebergeras ir kt. [12] sukūrė tinklą U-Net. Tinklo struktūra primena U raidę ir esminis jos bruožas, kad sudaryta iš dviejų kelių – mažėjančio kelio kairėje pusėje ir besiplečiančio kelio dešinėje. Mažėjantis kelias mažina vaizdo dydį ir kiekvienoje operacijoje dvigubina požymių kanalų skaičių, o besiplečiantis kelias didina savybių žemėlapi mažindamas požymių kanalų skaičių. U-Net tinklas, apdorodamas biomedicininis vaizdus, 2015 metais laimėjo du ISBI (angl. *International Symposium on Biomedical Imaging*) vaizdų segmentavimo konkursus.

U-Net tinklo architektūra pavaizduota 1 pav. Ji susideda iš mažėjančio kelio (kairėje pusėje) ir besiplečiančio kelio (dešinėje pusėje). Mažėjantis kelias yra tipinės architektūros konvoliucinis tinklas. Jis susideda iš pasikartojančių 3x3 konvoliucijų (angl. *unpadded convolutions*), po kurių yra ReLU sluoksnis ir 2x2 dydžio sujungimo (angl. *pooling*) sluoksnis su žingsniu 2, skirtu sumažinti. Kiekviename sumažinime yra dvigubinamas požymių kanalų skaičius. Kiekvienas žingsnis besiplečiančiame kelyje susideda iš požymių žemėlapio padidrinimo naudojant 2x2 konvoliucijas, kurios perpus sumažina požymių kanalų skaičių, susiejant atitinkamai apkirptus požymių žemėlapius iš mažėjančio kelio, ir dvi 3x3 konvoliucijas, po kurių eina ReLU sluoksnis. Šis apkirpimas yra būtinas, kad kiekvienos konvoliucijos metu būtų išvengta pikselių praradimų vaizdo kraštuose.



1 pav. U-Net tinklo architektūra (Ronneberger et al., 2015).

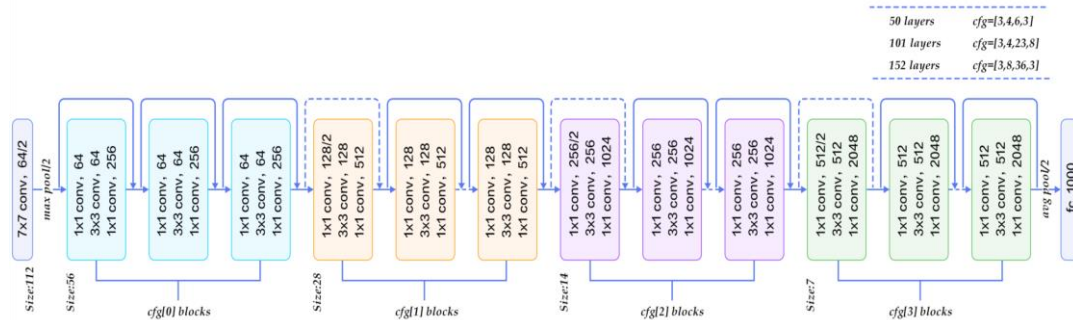
Apkirptų požymių žemėlapių sudarymas iš aukštos raiškos vaizdų leidžia tiksliai ir greitai išgauti požymius ir pasiekti gerus rezultatus su mažomis mokymo aibėmis. Tai labai svarbu dirbant su biomedicininiais vaizdais, prie kurių prieiga yra labai ribota.

Paskutinis sluoksnis yra 1x1 konvoliucija, kuri naudojama žemėlapio iš kiekvieno 64 komponentų požymių vektoriaus sudarymui, priskiriant jį siekiamam klasių skaičiui. Iš viso tinklas turi 23 konvoliucinius sluoksnius.

Kadangi U-Net neturi visiškai sujungtų sluoksnių, tai leidžia jam dirbti su skirtingo dydžio vaizdais.

Biomedicininį 3D vaizdų segmentavimui F. Milletari ir kt. [13] pristatė VNet tinklo architektūrą, kuri struktūriškai primena V raidę ir U-Net tinklo architektūrą. Kairėje pusėje vyksta vaizdų sumažinimas, dešinėje – padidinimas iki originalaus dydžio. Tinklas yra padalintas į po tris konvoliucinius sluoksnius turinčius etapus, kuriuose vyksta darbas su skirtingomis vaizdų raiškomis. Kiekvieno etapo pabaigoje yra naudojama likutinė funkcija, kuri, skirtingai nuo U-Net architektūros, užtikrina konvergenciją. Išėjime yra sugeneruojami du požymių žemėlapiai, kuriuos skaičiuoja greitas konvoliucinis sluoksnis, turintis 1x1x1 dydžio branduolį.

ILSVRC 2015 konkurse pirmą vietą užėmė K. He ir kt. [14] sukurtas ResNet tinklas (angl. *Residual Neural Network*). Tinklo architektūros idėja yra liekamieji blokai (angl. *residual block*), dėl kurių kiekvienas sluoksnis naudoja ne tik ankstesnio sluoksnio išėjimus, bet ir originalius nepalietus duomenis požymių išgavimui. ResNet naudoja trumpesnio kelio jungtis, kad būtų peršokama per 2–3 sluoksnius. Tinklo architektūra pavaizduota 2 pav.



2 pav. ResNet tinklo architektūra (medium.com)

S. Sabour ir kt. [15] pristatė kapsulinį neuroninį tinklą CapsNet. Tinklo pagrindinis elementas – kapsulė [16], kuri pati yra neuronų sluoksnių rinkinys. Kapsulė išėjimo duomenis pateikia vektoriuje, kuriame pateikiama informacija apie objekto pozą. Tai leidžia išsaugoti objektų pozos informaciją (orientacija, dydis ir pan.) ir, lyginant su įprastais konvoliuciniais tinklais, CapsNet geba kur kas efektyviau atpažinti objektus, kurių poza yra pakitusi, pavyzdžiui, pasuktas dideliu kampiu.

4 Akies vaizdų semantinis segmentavimas

Iš akies dugno nuotraukų galima nustatyti įvairias akies, širdies ir kraujotakos ligas. Analizuojant akies dugno nuotraukas, esminis momentas yra išskirti akies dalį, pagal kurią bus vykdomas susirgimų diagnozavimas. Tam naudojami semantinio segmentavimo algoritmai. Vienas esminių objektų yra akies kraujagyslių tinklas, iš kurio pakitimų yra nustatomos tokios ligos kaip diabetinė retinopatija, glaukoma, makulos degeneracija, choroidinė neovaskuliarizacija, hipertenzija, insultas.

Vieni pirmųjų dirbtinius neuroninius tinklus kraujagyslių klasifikacijai iš rentgeno nuotraukų panaudojo R. Nekovei ir S. Ying [17]. Jų sukurtas algoritmas remiasi 3 sluoksnių klaidos skleidimo atgal neuroniniu tinklu. Algoritmas aptinka pilkus vaizdų pikselius, kiekvieną jų apdoroja, sukurdamas aplink langą, kuriame aptinka kitus pilkus pikselius. Tokiu būdu slinkdamas per visą nuotrauką algoritmas išskiria visus pilkus rentgeno nuotraukos pikselius, kurie pateikiami neuroniniam tinklui kaip įėjimo duomenys. Mokymo duomenų aibė sudaryta iš rankiniu būdu pasirinktų lopinėlių, kuriuose kraujagyslių ir fono pikselių yra maždaug po lygiai. Šis metodas atpažįstant kraujagysles iš rentgeno nuotraukų pasiekė 92 proc. tikslumą.

Kraujagyslių atskyrimui nuo fono taikomi įvairūs metodai, kuriuos galima suskirstyti į dvi grupes – taisyklėmis grįsti metodai ir mašininio mokymosi metodai [18]. Prie taisyklėmis grįstų metodų priskiriami dvimatis suderintas filtras (angl. *two-dimensional matched filter response* (2D MFR)), morfologinis metodas, kraujagyslių sekimo metodas. Prie mašininio mokymosi metodų priskiriami k-NN, SVM, Bayesian decision rule, Fuzzy C-means, K-means [19] ir dirbtiniai neuroniniai tinklai.

Algoritmų efektyvumui įvertinti yra taikomi keli metodai. Vienas populiariausių yra ROC (angl. Receiver operating characteristic) kreivė, parodanti, kaip modelis geba atskirti klases. AUC (angl. *Area Under Curve*) laipsnis parodo ribą žemiau šios kreivės. Idealus skaičius yra 1, o jei jis žemiau 0,5, vadinasi, modelis yra nepatikimas. Kitas taikomas metodas tikslumui nusakyti yra DICE koeficientas, dar vadinamas F-Score.

D. Maji ir kt. [20] kraujagyslių atskyrimui nuo fono panaudojo dvylikos konvoliucinių neuroninių tinklų grupę. Kiekvienas konvoliucinis neuroninis tinklas sudarytas iš trijų sluoksnių, kurių kiekvienas buvo apmokytas atskirai naudojant 60 000

atsitiktinai parinktų 31x31 dydžio lopinėlių iš 20-ies spalvotų akies dugno vaizdų iš DRIVE duomenų bazės. AUC laipsnis – 0,947.

Dar geresnius tikslumo rezultatus pasiekė P. Liskowski ir kt. [21]. Jų neuroninis tinklas turėjo tris konvoliucinius sluoksnius, vieną sujungimo ir du visiškai sujungtus sluoksnius. Konvoliuciniai sluoksniai naudojo žingsnį 1, o sujungimo – 2. Šio algoritmo AUC siekė 0,972 su DRIVE baze ir 0,9605 su STARE, o klasifikavimo tikslumas viršijo 0,9495 (DRIVE) ir 0,9416 (STARE).

A. Dasgupta ir S. Singhas [22] kraujagyslių išskyrimui pasiūlė linijinę 8 sluoksnių konvoliucinio tinklo architektūrą. Pirmieji du sluoksniai konvoliuciniai, trečiasis sujungimo, po to vėl du konvoliuciniai, šeštasis – padidinimo (angl. upsampling), skirtas erdvinėms išėjimo dimensijoms padidinti, ir paskutiniai du – konvoliuciniai. Panaudojus daugiaklasę klasifikaciją tinklas buvo mokomas atspėti ne vieną reikšmę, o reikšmių vektorius. Tinklo tikslumas naudojant DRIVE duomenų bazę buvo 95,33 proc., o AUC siekė – 0,974.

Diabetinės retinopatijos diagnozei J. Brownas ir kt. [23] panaudojo dvi skirtingas neuroninių tinklų architektūras: kraujagyslių segmentavimui U-Net, o klasifikavimui GoogLeNet. Šis algoritmas buvo išbandytas prieš aštuonis žmones ekspertus, turinčius daugiau nei 10 metų patirties medicinoje. Algoritmas pasiekė geresnius rezultatus net prieš 6 iš 8 ekspertų. Geriausias pasiektas AUC laipsnis buvo 0,98.

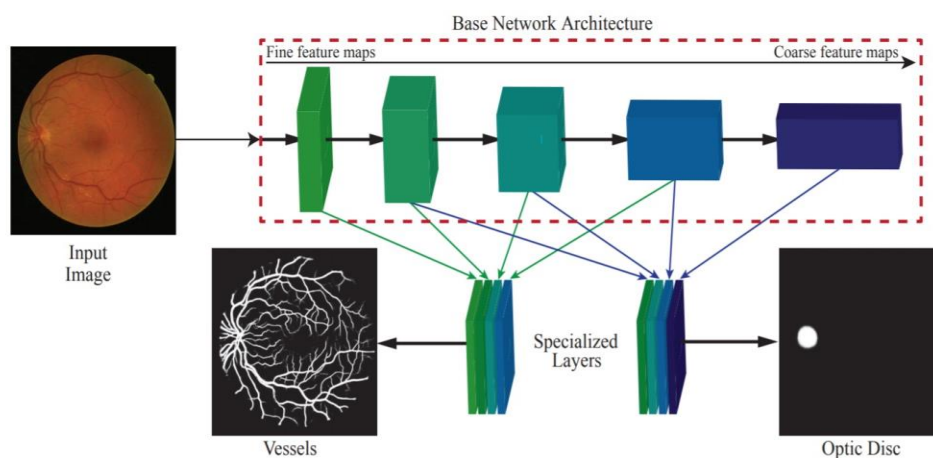
T. Laibacheris ir kt. [24] U-Net pagrindu sukūrė naują architektūrą akies kraujagyslių segmentavimui – M2U-Net. Į enkoderį buvo įtraukti apmokyti MobileNetV2 [25] komponentai, o dekoderyje panaudoti mažėjantys butelio kakliuko blokai (angl. *contracting bottleneck blocks*). Tai leido sumažinti parametrų skaičių nuo 31 milijono (originalus U-Net) iki 0,55 milijono. Ši architektūra pasižymi itin didele sparta bei gerais tikslumo rezultatais pagal AUC – 0,9714 (DRIVE). Kadangi ši architektūra nereikalauja didelių kompiuterio resursų, ji turi dideles panaudojimo perspektyvas įterptinėse sistemose.

Kita svarbi akies dalis yra optinis diskas. Pagal jo pokyčius yra diagnozuojama glaukoma ar diabetinė retinopatija. H. Alghamdi ir kt. [26] panaudodami konvoliucinį tinklą sukūrė automatinę optinio disko pakitimų aptikimo sistemą. Matuodama disko įdubos ir disko skersmens santykį ji identifikuoja glaukomą, o identifikuo-dama tam tikrus pažeidimus, pvz., kraujavimą, nustato kitas anomalijas. Klasifikatoriaus mokymui ir savybių išgavimui jie panaudojo kaskadinius klasifikatorius (AdaBoost, Haar-like). Sistema pasižymi didele greitaveika ir geru tikslumu – su STARE duomenų baze pasiektas 86,7 proc., o su DRIVE – net 100 proc. tikslumas.

A. Mitra ir kt. [27] sukūrė algoritmą optinio disko segmentavimui. Panaudoję 24 sluoksnių konvoliucinį neuroninį tinklą su MESSIDOR vaizdų baze pasiekė 99,05 proc. atpažinimo tikslumą, o su DRIVE – net 99,41 proc. Jų sukurtas tinklas kontekstinę informaciją išgauna tiesiai iš akies nuotraukos, praleisdamas požymių išgavimo fazę, ir be jokio žmogaus įsikišimo.

Efektyvų automatinio segmentavimo algoritmą DRIU (angl. *Deep Retinal Image Understanding*) (3 pav.), naudojant giliuosius neuroninius tinklus, pasiūlė K. Maninis ir kt. [28]. DRIU atliko ir kraujagyslių, ir optinio disko segmentavimą. Pagrindinis tinklas buvo grįstas VGG architektūra su pašalintais visiškai sujungtais sluoksniais. Galutiniame vaizdų apdorojimo etape buvo pridėti specializuoti sluoksniai, kurių vienas segmentavo kraujagysles, o kitas optinį diską. Tokia architektūra pademonstravo puikius spartos rezultatus, vieną vaizdą segmentuodama per 65–110 ms (greitis priklausė nuo pasirinktos duomenų bazės ir vaizdo dydžio). Tai pranoko daugelio kitų algoritmų pasiektus rezultatus. Kraujagyslių nustatymo tikslumas siekė

0,822 (DRIVE) ir 0,831 (STARE), o optinio disko – 0,971 (DRIONS-DB) ir 0,959 (RIM-ONE). Palyginkime: žmogaus eksperto pasiekti rezultatai nustatant optinio disko pakitimus siekė 0,967 (DRIONS-DB) ir 0,952 (RIM-ONE).



3 pav. DRIU architektūra leidžia atlikti akies kraujagyslių ir optinio disko segmentavimą (Maninis et al., 2016).

A. Sevastopolsky ir kt. [29] optinio disko ir jo įdubos segmentavimui panaudojo modifikuotą U-Net architektūrą Stack-U-Net, kurios esmė yra tipišku U-Net arba Res-U-Net (U-Net su liekamosiomis (angl. *residual*) jungtimis) jungimas į blokus. Didžiausias tikslumas buvo pasiektas panaudojus 15-a Res-U-Net bloką. Pagal DICE koeficientą (F-Score) optinio disko segmentavimo rezultatai buvo 0,96 (DRIONS-DB), 0,95 (RIM-ONE) ir 0,97 (DRISHTI-GS), o disko įdubos segmentavimo – 0,84 (RIM-ONE) ir 0,89 (DRISHTI-GS).

A. Palas ir kt. [30] glaukoms aptikti sukūrė daugiamodelinį tinklą G-EyeNet. Jis sudarytas iš enkoderio, dekoderio bei klasifikatoriaus. Dekoderis atlikdamas rekonstrukcijos funkciją padeda klasifikatoriui sumažinti praradimus ir pagerinti klasifikavimą. Glaukoma identifikuojama pagal optinio nervo ir kraujagyslių pakitimus. G-EyeNet akies dugno nuotraukose pirmiausia segmentuoja optinį diską, po to išskiria dominantį regioną (angl. *Interest of Region*, ROI). Identifikuojant glaukomą AUC pasiekė 0,923.

H. Fu ir kt. [31] pasiūlė daugiakontekstį gilųjį tinklą (angl. Multi-Context Deep Network, MCDN) uždarąjo kampo glaukoms aptikti. MCDN architektūroje požymių žemėlapiui sudaryti lygiagrečiai naudojami du VGG-16 tinklai – vienas globaliam vaizdui (visa nuotrauka), kitas lokaliai (iškirptas lopinėlis). Klasifikavimui taikomas SVM metodas. Geriausias pasiektas AUC rezultatas – 0,9456.

Diagnozuojant tam tikras ligas, pavyzdžiui, diabetinę makulos edemą, yra poreikis išskirti akies tinklainę. A. Ben-Cohenas ir kt. [32] tinklainės segmentavimui atlikti išbandė U-Net tinklo architektūrą. Optinės koherentinės tomografijos (OCT) nuotraukose neuroninis tinklas atliko tinklainės sluoksnio segmentavimą, po to panaudojant Sobelio filtrą buvo nustatomos sluoksnio ribos. Galutinėje fazėje panaudojant Dijkstros algoritmą ribos buvo patikslinamos. Toks segmentavimo algoritmas segmentuojant tinklainės nervo pluošto sluoksnį pagal F-Score pasiekė rezultatą 0,95.

5 Išvados

Automatizuotas akies dugno nuotraukų tyrimas gali gerokai palengvinti gydytojų darbą diagnozuojant įvairius susirgimus. Semantinis vaizdų segmentavimas, kurio metu kiekvienas pikselis priskiriamas tam tikrai objektų klasei, leidžia iš fono išskirti tokias akies dalis kaip kraujagyslės, optinis diskas ar makula. Konvoliuciniai neuroniniai tinklai yra vienas efektyviausių metodų atlikti semantinį segmentavimą. Įvairių mokslininkų atliktų eksperimentų rezultatai rodo, jog tikslumas dažnai viršija 95 proc.

Biomedicininį vaizdų analizėje populiaru U-Net konvoliucinio tinklo architektūra. Viena priežasčių – ši architektūra buvo kuriama darbui su biomedicininiais vaizdais. Be to, šis neuroninis tinklas pasižymi tokiomis savybėmis kaip efektyvus skaičiavimo resursų išnaudojimas bei greitas ir tikslus mokymasis su mažomis vaizdų aibėmis. Tai pasiekama, nes tinklas, lyginant su dauguma kitų architektūrų, gerokai greičiau aptinka reikiamas tikslias detales. Skirtingų autorių darbai rodo, jog U-Net architektūra yra lengvai tobulinama ir pasiekiamas dar geresnis efektyvumas.

Konvoliuciniai tinklai leidžia išskirtose akies dalyse efektyviai aptikti įvairias anomalijas ir nustatyti susirgimus. Sprendžiant šiuos uždavinius susiduriama su įvairiomis problemomis, pvz., kraujagyslių susiliejimu ar išnykimu, per mažu ar per dideliu apšvietimu, netinkamu regėjimo lauku. Dalis šių problemų yra sunkiai sprendžiamos, tad reikalauja tolesnių intensyvių tyrimų, siekiant efektyvesnių rezultatų.

1 lentelė. Akies dugno nuotraukų segmentavimo algoritmų, naudojančių neuroninius tinklus, apibendrinimas

Autoriai	Architektūra	Uždavinys	Vaizdų duomenų bazė	Matavimo metrika	Rezultatas
Maji ir kt. (2015)	12-os CNN junginys	Kraujagyslių segmentavimas	DRIVE	AUC	0,947
Liskowski ir kt. (2016)	CNN	Kraujagyslių segmentavimas	DRIVE	AUC	0,972
			STARE	AUC	0,9605
Dasgupta ir Singh (2017)	CNN	Kraujagyslių segmentavimas	DRIVE	AUC	0,974
Brown ir kt. (2018)	GoogLeNet, U-Net	Diabetinės retinopatijos diagnozė	Privati	AUC	0,98
Laibacher ir kt. (2018)	M2U-Net	Kraujagyslių segmentavimas	DRIVE	AUC	0,9714
Alghamdi ir kt. (2016)	CNN	Optinio disko segmentavimas	DRIVE	Tikslumas	100 %
			STARE	Tikslumas	86,71 %
Mitra ir kt. (2018)	CNN	Optinio disko segmentavimas	MESSIDOR	Tikslumas	99,05 %
Maninis ir kt. (2016)	CNN DRIU	Optinio disko segmentavimas	DRIONS-DB	Accuracy	0,959
				F-Score	0,97
			RIM-ONE v.3	Tikslumas	0,959
		Kraujagyslių segmentavimas	DRIVE	Tikslumas	0,822
			STARE	Tikslumas	0,831
	Stack-U-Net		DRIONS-DB	F-Score	0,96

Sevastopolsky ir kt. (2018)		Optinio disko segmentavimas	RIM-ONE	F-Score	0,95
			DRISHTI-GS	F-Score	0,97
		Disko įdubos segmentavimas	RIM-ONE	F-Score	0,84
			DRISHTI-GS	F-Score	0,89
Pal ir kt. (2018)	G-EyeNet	Glaukomos diagnozavimas	HRF, RIM ONE v.3, DRISHTI-GS ir DRIONS-DB	AUC	0,923
Fu ir kt. (2018)	MCDN	Glaukomos diagnozavimas	Privati	AUC	0,9456
Ben-Cohen ir kt. (2017)	U-Net	Tinklainės segmentavimas	Privati	F-Score	0,95

Kadangi biomedicininį vaizdų prieiga yra viena esminių problematikų, tyrimai atliekami su labai mažomis duomenų imtimis, dažniausiai DRIVE ir STARE vaizdų bazėmis, todėl tyrimuose aprašytus metodus reikia validuoti su didesnėmis vaizdų bazėmis. Be to, į šias vaizdų bazes būna atrinktos kokybiškos nuotraukos, o realybėje taip būna ne visada ir tikslumas gali gerokai skirtis.

6 Literatūra

- [1] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, “GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification”, *Neurocomputing*, vol. 321, pp. 321–331, 2018.
- [2] W. Xiancheng, L. Wei, M. Bingyi, J. He, Z. Jiang, W. Xu, Z. Ji, G. Hong, and S. Zhaomeng, “Retina blood vessel segmentation using a U-net based Convolutional neural network”, in *Procedia Computer Science: International Conference on Data Science (ICDS 2018), Beijing, China, 8-9 June 2018*, 2018.
- [3] B. Y. Shen and S. Mukai, “A portable, inexpensive, nonmydriatic fundus camera based on the Raspberry Pi® Computer”, *Journal of ophthalmology*, vol. 2017, 2017.
- [4] E. Trucco, A. Ruggeri, T. Karnowski, L. Giancardo, E. Chaum, J. P. Hubschman, B. Al-Diri, C. Y. Cheung, D. Wong, M. Abramoff, *et al.*, “Validating retinal fundus image analysis algorithms: issues and a proposal”, *Investigative ophthalmology & visual science*, vol. 54, no. 5, pp. 3546–3559, 2013.
- [5] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp, “Automated assessment of diabetic retinal image quality based on clarity and field definition”, *Investigative ophthalmology & visual science*, vol. 47, no. 3, pp. 1120–1125, 2006.
- [6] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, *et al.*, “Gradient-based learning applied to document recognition”, *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks”, in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [8] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan,

- V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
- [9] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition”, *arXiv preprint arXiv:1409.1556*, 2014.
- [10] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [11] V. Badrinarayanan, A. Handa, and R. Cipolla, “Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling”, *arXiv preprint arXiv:1505.07293*, 2015.
- [12] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation”, in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.
- [13] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-net: Fully convolutional neural networks for volumetric medical image segmentation”, in *2016 Fourth International Conference on 3D Vision (3DV)*, IEEE, 2016, pp. 565–571.
- [14] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [15] S. Sabour, N. Frosst, and G. E. Hinton, “Dynamic routing between capsules”, in *Advances in neural information processing systems*, 2017, pp. 3856–3866.
- [16] G. E. Hinton, A. Krizhevsky, and S. D. Wang, “Transforming auto-encoders”, in *International Conference on Artificial Neural Networks*, Springer, 2011, pp. 44–51.
- [17] R. Nekovei and Y. Sun, “Back-propagation network and its configuration for blood vessel detection in angiograms”, *Neural Networks, IEEE Transactions on*, vol. 6, pp. 64–72, Feb. 1995. doi: 10.1109/72.363449.
- [18] G. Stabingis, J. Bernatavičienė, G. Dzemyda, A. Paunksnis, L. Stabingienė, P. Treigys, and R. Vaičaitienė, “Adaptive eye fundus vessel classification for automatic artery and vein diameter ratio evaluation”, *Informatica*, vol. 29, no. 4, pp. 757–771, 2018.
- [19] Z. Yavuz and C. Köse, “Blood vessel extraction in color retinal fundus images with enhancement filtering and unsupervised classification”, *Journal of healthcare engineering*, vol. 2017, 2017.
- [20] D. Maji, A. Santara, S. Ghosh, D. Sheet, and P. Mitra, “Deep neural network and random forest hybrid architecture for learning to detect retinal vessels in fundus images”, in *2015 37th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2015, pp. 3029–3032.
- [21] P. Liskowski and K. Krawiec, “Segmenting Retinal Blood Vessels With Deep Neural Networks”, *IEEE Transactions on Medical Imaging*, vol. 35, no. 11, pp. 2369–2380, 2016, issn: 0278-0062. doi: 10.1109/TMI.2016.2546227.
- [22] A. Dasgupta and S. Singh, “A fully convolutional neural network based structured prediction approach towards the retinal vessel segmen-

- tation”, in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, IEEE, 2017, pp. 248–251.
- [23] J. M. Brown, J. P. Campbell, A. Beers, K. Chang, S. Ostmo, R. P. Chan, J. Dy, D. Erdogmus, S. Ioannidis, J. Kalpathy-Cramer, *et al.*, “Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks”, *JAMA Ophthalmology*, vol. 136, no. 7, pp. 803–810, 2018.
- [24] T. Laibacher, T. Weyde, and S. Jalali, “M2U-Net: Effective and Efficient Retinal Vessel Segmentation for Resource-Constrained Environments”, *arXiv preprint arXiv:1811.07738*, 2018.
- [25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.
- [26] H. S. Alghamdi, H. L. Tang, S. A. Waheeb, and T. Peto, “Automatic optic disc abnormality detection in fundus images: A deep learning approach”, 2016.
- [27] A. Mitra, P. S. Banerjee, S. Roy, S. Roy, and S. K. Setua, “The region of interest localization for glaucoma analysis from retinal fundus image using deep learning”, *Computer methods and programs in biomedicine*, vol. 165, pp. 25–35, 2018.
- [28] K.-K. Maninis, J. Pont-Tuset, P. Arbelaez, and L. Van Gool, “Deep retinal’ image understanding”, in *International conference on medical image computing and computer-assisted intervention*, Springer, 2016, pp. 140–148.
- [29] A. Sevastopolsky, S. Drapak, K. Kiselev, B. M. Snyder, J. D. Keenan, and A. Georgievskaya, “Stack-U-Net: Refinement Network for Image Segmentation on the Example of Optic Disc and Cup”, *arXiv preprint arXiv:1804.11294*, 2018.
- [30] A. Pal, M. R. Moorthy, and A. Shahina, “G-Eyenet: A Convolutional Autoencoding Classifier Framework for the Detection of Glaucoma from Retinal Fundus Images”, in *2018 25th IEEE International Conference on Image Processing (ICIP)*, IEEE, 2018, pp. 2775–2779.
- [31] H. Fu, Y. Xu, S. Lin, D. W. K. Wong, B. Mani, M. Mahesh, T. Aung, and J. Liu, “Multi-context deep network for angle-closure glaucoma screening in anterior segment oct”, in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2018, pp. 356–363.
- [32] A. Ben-Cohen, D. Mark, I. Kovler, D. Zur, A. Barak, M. Iglicki, and R. Soferman, “Retinal layers segmentation using fully convolutional network in OCT images”, *RSIP Vision*, 2017.