

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Erki Arusoo 183108 YADM

TEHISNÄRVIVÕRKUDE KIRJELDAMINE VÄIKESE MAAILMA OMADUSE ABIL

Magistritöö

Juhendaja: Laura Päeske
MSc
Uljana Reinsalu
PhD

Tallinn 2020

Autorideklaratsioon

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor: Erki Arusoo

18.05.2020

Annotatsioon

Tänu paremale riistvarale ja ligipääsule suurematele andmebaasidele, on masinõppes viimasel paaril kümnendil järjest tähtsamaks saanud tehisnärvivõrgud. Mida suuremaks ja komplekssemaks aga tehisnärvivõrgud muutuvad, seda suuremaks kasvab niinimetatud „musta kasti“ efekt ehk tehisnärvivõrkude seletamatus. See seab aga takistusi nende suuremahulisema kasutamiseni valdkondades, kus on kaalul inimesed, nagu tervishoius, kus tehisnärvivõrkusid kasutatakse näiteks haiguste ennetamisel ja ravil. Viimastel aastatel on toimunud märgatav kasv publikatsioonide seas, mis tegelevad tehisnärvivõrkude tõlgendamise probleemiga. Autorile teadaolevatel andmetel pole hetkel aga kasutatud komplekssete võrkude analüüsimeetodeid tehisnärvivõrkude kirjeldamisel.

Käesoleva töö eesmärk oli selgitada, kas tehisnärvivõrkudel, mis on treenitud klassifitseerima keerukamaid andmeid, on kaalukoefitsiendid kõrgema väikese maailma mõõtme väärtusega. Eesmärgi täitmiseks püstitatud ülesanneteks oli treenida erinevad tehisnärvivõrgud järjest keerulisemate treeningandmetega, leida tehisnärvivõrkude kaalukoefitsientidel arvatud keskmised väikese maailma mõõtmed ning leida seos treeningandmete keerukuse ja väikese maailma mõõtmete vahel.

Treeningandmetena kasutati MNIST andmebaasi numbreid ning ühte andmebaasi, mis koosnes ühes Hiina haiglas tehtud laste rindkere röntgenülesvõtetest. Esimese meetodiga suurendati treeningandmete keerukust neid manipuleerides: roteerides, skaleerides, nihutades ning kõiki eelnevaid võtteid kombineerides. Teise meetodiga muudeti erinevate numbrite arvu, mida tehisnärvivõrgud klassifitseerima pidid. MNIST andmebaasil rakendati mõlemat eelpool kirjeldatud meetodit, röntgenülesvõtetel esimest meetodit. Kõikidest tehisnärvivõrkudest võeti välja kõikide täielikult ühendatud kihtide kaalukoefitsiendid. Kaalukoefitsientide korrelatsioonide põhjal arvutati igale kihile klasterdamise kordaja ja karakteristiklik teepikkus ning nende põhjal väikese maailma mõõde.

Käesoleva töö tulemustest selgus, et kõigi kasutatud tehisnärvivõrkude puhul suurenes väikese maailma mõõde oluliselt esimese täielikult ühendatud kihi kaalukoefitsientidel, kui treenitavad andmed muutusid keerulisemaks. Tulemustest võis järeldada, et enamasti bioloogilistel närvivõrkudel rakendatavatel komplekssete võrkude analüüsimeetoditel võib olla rakendusi ka tehisnärvivõrkude korral. Käesolevas töös kasutatud meetodil arvatud väikese maailma mõõtmeid võib tulevikus kaaluda näiteks lisaabinõuna tehisnärvivõrkude üldistusvõime hindamisel, kuid see nõuab siiski edasist uurimist. Lisaks tuleb märkida, et katsed teostati väga kontrollitud tingimustes ning muudeti vaid treeningandmeid. Tulevikus tasub kindlasti kontrollida, kuidas muutub väikese maailma mõõde teistsuguste tehisnärvivõrkude ja treeningandmetega ning tehisnärvivõrkude parameetreid muutes.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 39 leheküljel, 8 peatükki, 23 joonist, 10 tabelit.

Abstract

Describing artificial neural networks with small-worldness

In the field of artificial intelligence, artificial neural networks (ANN) have become more important during the last couple of decades due to better equipment and access to larger databases. However, the larger and more complex ANNs become, the bigger the problem of so called „black box“ effect becomes. In other words, ANNs become less explainable. This could cause obstacles in implementing them in a larger scale in the fields, where human lives could be at stake, for example, in healthcare, where ANNs are used to help prevent and cure illnesses. In the last couple of years, the number of publications discussing the problem of explainable artificial intelligence has grown remarkably. However, to the best of the author’s knowledge, complex network analysis methods have not been implemented in describing artificial neural networks to this day.

The aim of this study was to determine whether the weights of ANNs, which have been trained with more difficult data, have better small world properties. The tasks included training several ANNs with data of different complexities, finding the measures of small-worldness for weights of every ANN layer, and finding the connection between the measure of small-worldness and complexity of training data.

MNIST database numbers and a database, which consisted of Chinese childrens’ chest x-ray images, were used for training data. The first method consisted of increasing the complexity of training data by manipulating them: rotating, scaling, shifting and combining all of the above. The second method consisted of changing the number of different values the ANNs were trained to classify. The two methods described were applied on the MNIST database numbers and only the first method on the x-ray images. All of the weights from all of the fully-connected layers were extracted. Through clustering coefficients and characteristic path lengths, measures of small-worldness were calculated for every layer based on the correlation coefficients of the corresponding weights.

The results showed that when training data became more complex, the measures of small-worldness that were calculated on the weights of the first fully-connected layers became significantly bigger. The results were the same for all the different ANNs used. According to the result, it can be derived that complex network analysis methods that are generally popular in the field of biological neural networks could have applications in the field of ANNs. Measures of small-worldness found through the method described in this study could be considered as an extra measure to evaluate the generalisation capability of ANNs. However, it still needs further studies. Furthermore, the experiments were carried out in a very controlled environment and only training data were changed. In the future, it is recommended to test how the measure of small-worldness changes with different types of ANNs, training data and changing different parameters of ANN.

The thesis is in Estonian and contains 39 pages of text, 8 chapters, 23 figures, 10 tables.

Lühendite ja mõistete sõnastik

AP	<i>Anterior-posterior</i> , eest-taha (otse röntgenülesvõte)
ANN	<i>Artificial neural network</i> , tehisnärvivõrk
CAM	<i>Class activation maps</i> , klassi aktiveerimiskaardid
CNN	<i>Convolutional neural network</i> , konvolutsiooniline närvivõrk
DNN	<i>Deep neural network</i> , süvavõrk
MNIST Database	<i>Modified National Institute of Standards and Technology</i> andmebaas
ReLU	<i>Rectified linear unit</i> , mittenegatiivne lineaarfunktsioon

Sisukord

1 Sissejuhatus	13
2 Mõisted	15
2.1 Tehisnärvivõrk.....	15
2.2 Konvolutsiooniline võrk.....	17
2.3 Graaf	21
3 Teoreetilised alused	23
3.1 Tehisnärvivõrkude tõlgendamine	23
3.2 Väikese maailma omadustega võrgud	26
3.2.1 Väikese maailma omadused	26
3.2.2 Väikese maailma omaduste hindamine	27
3.3 Väikese maailma omadused tehisnärvivõrkudel	28
4 Metoodika.....	29
4.1 Kasutatavad andmed ja tehisnärvivõrgud.....	29
4.1.1 MNIST andmebaas	29
4.1.2 Rindkere röntgenülesvõtete andmebaas	29
4.1.3 Numbrite klassifitseerimiseks kasutatavad tehisnärvivõrgud	30
4.1.4 Röntgenülesvõtete klassifitseerimiseks kasutatav tehisnärvivõrk.....	30
4.2 Arvutusteks kasutatud tarkvara	30
4.2.1 <i>Brain Connectivity Toolbox</i>	31
4.3 Ülevaade protsessist	32
4.3.1 MNIST andmebaasi treeningandmete töötlemine ja tehisnärvivõrkude treenimine	32
4.3.2 Röntgenülesvõtete andmebaasi treeningandmete töötlemine ja konvolutsiooniliste võrkude trenimine	33
4.3.3 Väikese maailma mõõtme arvutamine	33
5 Tulemused	37
5.1 Manipuleeritud numbrite klassifitseerimine kasutades ühe peidetud kihiga tehisnärvivõrkusid	37
5.1.1 Väikese maailma mõõde.....	37

5.1.2 Naabusmaatriksite tihedus	38
5.2 Manipuleeritud numbrite klassifitseerimine kasutades kahe peidetud kihiga tehiskärvivõrkusid	38
5.2.1 Väikese maailma mõõde.....	39
5.2.2 Naabusmaatriksite tihedus	40
5.3 Erineva arvu numbrite klassifitseerimine kasutades ühe peidetud kihiga tehiskärvivõrkusid	41
5.3.1 Väikese maailma mõõde.....	41
5.3.2 Naabusmaatriksite tihedus	42
5.4 Erineva arvu numbrite klassifitseerimine kasutades kahe peidetud kihiga tehiskärvivõrkusid	43
5.4.1 Väikese maailma mõõde.....	43
5.4.2 Naabusmaatriksite tihedus	44
5.5 Röntgenülesvõtete klassifitseerimine konvolutsiooniliste võrkudega.....	45
5.5.1 Väikese maailma mõõde.....	45
5.5.2 Naabusmaatriksite tihedus	46
6 Arutelu	47
7 Järeldused	51
8 Kokkuvõte	52

Jooniste loetelu

Joonis 1. Tehisneuron	15
Joonis 2. Sigmoidfunktsioon	16
Joonis 3. Näide lihtsast tehishärvivõrgust	16
Joonis 4. Näide konvolutsioonilisest võrgust	18
Joonis 5. Näide konvolutsioonist.....	19
Joonis 6. Mittenegatiivne lineaarfunktsioon.....	20
Joonis 7. Näide maksimaalsest ahendusest.....	21
Joonis 8. Näide graafist	21
Joonis 9. Süvavõrgu (DNN) ennustuse „paat“ tõlgendus sisendi x korral	23
Joonis 10. LRP näide	24
Joonis 11. Näide CAM'ist	25
Joonis 12. Väikese maailma omadustega võrk.....	26
Joonis 13. Väljavõte MNIST andmebaasi numbritest	29
Joonis 14. Väljavõte manipuleeritud treeningandmetest	32
Joonis 15. Näide kaalukoefitsientide maatriksite leidmisest lihtsast tehishärvivõrgust. 34	
Joonis 16. Näide kihi naabusmaatriksi kaudu väikese maailma mõõtme arvutamisest 35	
Joonis 17. Manipuleeritud andmetega treenitud ühe peidetud kihiga tehishärvivõrkude a) 1. kihtide ja b) 2. kihtide kaalukoefitsientidel arvutatud keskmised väikese maailma mõõtmel	38
Joonis 18. Manipuleeritud andmetega treenitud kahe peidetud kihiga tehishärvivõrkude a) 1. kihtide, b) 2. kihtide ja c) 3. kihtide kaalukoefitsientidel arvutatud keskmised väikese maailma mõõtmel	40
Joonis 19. Erineva arvu numbritega treenitud ühe peidetud kihiga tehishärvivõrkude 1. kihtide kaalukoefitsientidel arvutatud keskmised väikese maailma mõõtmel.....	42
Joonis 20. Erineva arvu numbritega treenitud ühe peidetud kihiga tehishärvivõrkude 2. kihtide kaalukoefitsientidel arvutatud keskmised väikese maailma mõõtmel.....	42
Joonis 21. Erineva arvu numbritega treenitud kahe peidetud kihiga tehishärvivõrkude 1. kihtide kaalukoefitsientidel arvutatud keskmised väikese maailma mõõtmel.....	44

Joonis 22. Erineva arvu numbritega treenitud kahe peidetud kihiga tehisnärvivõrkude a) 2. kihtide ja b) 3. kihtide kaalukoefitsientidel arvatud keskmised väikese maailma mõõtmed	44
Joonis 23. Manipuleeritud andmetega treenitud konvolutsiooniliste võrkude täielikult ühendatud kihi kaalukoefitsientidel arvatud keskmised väikese maailma mõõtmed..	46

Tabelite loetelu

Tabel 1. Manipuleeritud andmetega treenitud ühe peidetud kihiga tehisnärvivõrkude keskmised klassifitseerimistäpsused	37
Tabel 2. Manipuleeritud andmetega treenitud tehisnärvivõrkude kihtide keskmised naabusmaatriksite tihedused.....	38
Tabel 3. Manipuleeritud andmetega treenitud kahe peidetud kihiga tehisnärvivõrkude keskmised klassifitseerimistäpsused	39
Tabel 4. Manipuleeritud andmetega treenitud tehisnärvivõrkude kihtide keskmised naabusmaatriksite tihedused.....	40
Tabel 5. Erineva arvu numbritega treenitud ühe peidetud kihiga tehisnärvivõrkude keskmised klassifitseerimistäpsused	41
Tabel 6. Erineva arvu numbritega treenitud tehisnärvivõrkude kihtide keskmised naabusmaatriksite tihedused.....	43
Tabel 7. Erineva arvu numbritega treenitud ühe peidetud kihiga tehisnärvivõrkude keskmised klassifitseerimistäpsused	43
Tabel 8. Erineva arvu numbritega treenitud tehisnärvivõrkude kihtide keskmised naabusmaatriksite tihedused.....	45
Tabel 9. Konvolutsiooniliste võrkude keskmised klassifitseerimistäpsused.....	45
Tabel 10. Konvolutsiooniliste võrkude täielikult ühendatud kihtide keskmised naabusmaatriksite tihedused.....	46

1 Sissejuhatus

Nagu teistes valdkondades, leiab ka meditsiinis üha enam kasutust tehisintellekt. Kuna kasvanud on suurte andmebaaside hulk, leiavad üha enam kasutust just tehisnärvivõrkude (ingl. k. *artificial neural network*) mudelid. Tehisnärvivõrkude mudeleid kasutatakse nii haiguste diagnoosimisel, klassifitseerimisel, ennustamisel kui ka elundite elujõulisuse analüüsimisel [1]. Eeliseid võrreldes traditsiooniliste meetoditega on mitmeid. Tehisnärvivõrkudel on parem võime analüüsida mittelineaarseid seoseid ja korreleerunud sõltumatuid muutujaid. Tänu heale üldistusvõimele võimaldavad tehisnärvivõrgud klassifitseerida ka keerukamaid andmeid [1, 2].

Tehisnärvivõrkudel on ka mõningaid puuduseid, eriti nende tõlgendamine on osutunud väljakutseks. Suureks teguriks on, et tehisnärvivõrkude disainimisel kasutatakse reeglina katse-eksitusmeetodit. Arendajal on kontroll sisendi ja väljundi üle, kuid õppeprotsess, mis toimub peidetud kihtides, jääb tihtipeale segaseks, kuna muutujate arv on väga suur [3]. Väga paljusid tehisnärvivõrkude klassifitseerimistäpsuse suurendamiseks kasutatavaid võtteid, mille töötamine on leidnud empiirilist tõesust, saame teoreetiliselt ainult oletada [4, 5]. Seetõttu on mõnikord raske selgitada peidetud kihtide toimimismehhanisme, kuigi võrk ise töötab hästi. Tihti ollakse raskustes järgmiste küsimustega: kuidas tehisnärvivõrgu mudel teeb taolisi ennustusi, miks eelistab mudel teatud parameetreid rohkem ning millises suunas peaks liikuma mudeli täiendamisel [3]? Tõlgendamise raskus on tekitanud tehisnärvivõrkudele niinimetatud „musta kasti“ efekti isegi juhul, kui me mõistame protsessides peituvat matemaatilist tausta [3, 6].

Aastatega on tehisnärvivõrgud üha paremaks läinud, kuid ka seda läbipaistmatumaks [7]. See on aga toonud endaga kaasa ka suurema nõudluse mõistmaks, mis tehisnärvivõrgu tehtud otsuste taga seisab. Eriti kriitilise tähtsusega on see valdkondades, kus on kaalul inimeste elu, nagu meditsiinis. Kui arstid toetuvad üha enam oma töös tehisnärvivõrkudele, peavad nad ka aru saama, kuidas ja miks tehisintellekt sellise otsuseni jõudis [6, 7].

Komplekssete võrkude analüüs on multidistsiplinaarne uurimisvaldkond, mille kaudu üritatakse kvantifitseerida kompleksseid võrke läbi üksikute lihtsate kuid oluliste

arvutuste [8]. See pärineb graafiteooriast ning käsitleb peamiselt kompleksseid võrke, mis ei ole ei regulaarsed ega täielikult juhuslikud [8, 9]. Komplekssed võrgud on topoloogiliselt hõredad, kuid väikese maailma omadustega: tipud on grupeerunud klastritesse, kuid keskmine lühim teepikkus üle kõikide tippude on võrdlemisi väike. Taolisteks võrkudeks loetakse muuhulgas ka näiteks internet, sotsiaalvõrgustikud ja aju närvivõrgud [10]. Komplekssete võrkude analüüsi käigus kasutatavate meetoditega hinnatakse komplekssetele võrkudele omaseid suuruseid, nende hulgas ka väikese maailma omadust kirjeldavaid suuruseid: klasterdamise kordaja, karakteristiklik teepikkus ja nende kaudu arvutatav väikese maailma mõõde [8].

Käesoleva magistr töö eesmärgiks oli selgitada, kas tehisnärvivõrkudel, mis on treenitud klassifitseerima keerukamaid andmeid, on kaalukoefitsiendid kõrgema väikese maailma mõõtmega väärtusega.

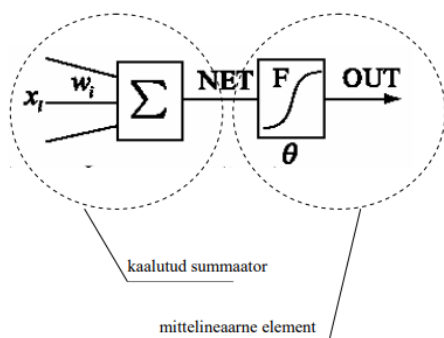
Eesmärgi täitmiseks püstitati järgmised ülesanded:

- 1) treenida erinevad tehisnärvivõrgud järjest keerulisemate treeningandmetega;
- 2) leida tehisnärvivõrkude kaalukoefitsientidel arvutatud keskmised väikese maailma mõõtmed;
- 3) leida seos treeningandmete keerukuse ning väikese maailma mõõtmete vahel.

2 Mõisted

2.1 Tehisnärvivõrk

Lihtsustatult öeldes koosneb tehisnärvivõrk suurel hulgal tehisneuronitest, mis on omavahel ühendatud, imiteerides nii bioloogilist närvivõrku. Tehisneuron (Joonis 1) koosneb kahest osast: kaalutud summaatorist ja mittelineaarsest elemendist. [2]



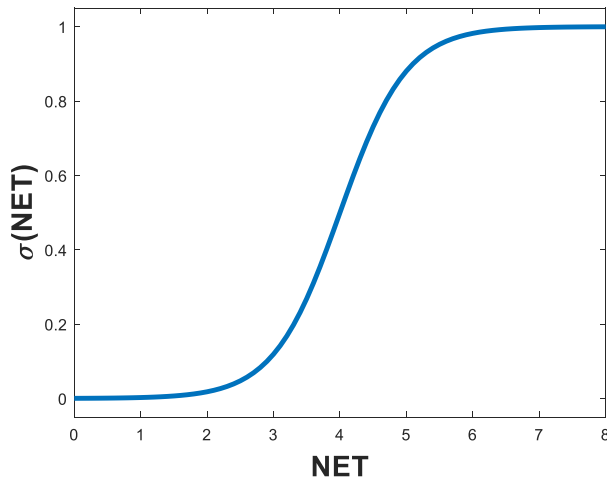
Joonis 1. Tehisneuron. Sisend x_i korrutatakse läbi kaalukoefitsiendiga w_i , liidetakse vabaliige ja saadakse kaalutud summaator NET. Kaalutud summaatorist arvutatakse aktiveerimisfunktsioon F väljundi OUT. [2]

Igal neuronil on n sisendit $x_1 \dots x_n$, mis korrutatakse läbi kaalukoefitsiendiga w_i ja liidetakse kokku. Kaalukoefitsiendid võivad olla nii positiivsed kui ka negatiivsed ja võimaldavad suurendada või vähendada sisendi tähtsust. Kaalutud summaatorile liidetakse vabaliige b , mis kujutab mingisugust läviväärtust ning saadakse summaatori väljund NET (Valem 1). [2, 4]

$$NET = W \cdot X + b = (w_1 x_1 + \dots + w_n x_n) + b \quad (1)$$

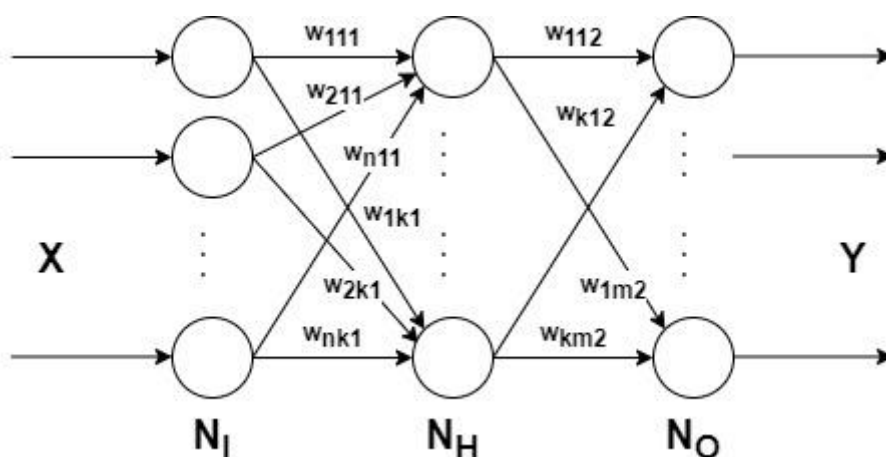
Sellest arvutab mittelineaarse elemendi funktsioon ehk aktiveerimisfunktsioon neuroni väljundi: $OUT = f(NET)$. Üks levinumaid aktiveerimisfunktsioone on sigmoid-funktsioon σ (Valem 2), mis tagab, et neuroni väljundi väärtus muutub sujuvalt 0 ja 1 vahel (Joonis 2). [2, 4]

$$OUT = \sigma(NET) = \frac{1}{1 + e^{-NET}} \quad (2)$$



Joonis 2. Sigmoidfunktsioon. Funktsioon tagab, et tehisneurooni väljund jääks 0 ja 1 vahele. Autori koostatud.

Tehisneuronid paiknevad tehisnärvivõrkudes kihtidena. Kõige lihtsam tehisnärvivõrk (Joonis 3) on otsesuunatud võrk ehk pärilevivõrk (ingl. k. *feedforward network*), millel on üks sisendkiht, üks väljundkiht ja üks või mitu täielikult ühendatud peidetud kihti (ingl. k. *fully connected layer*). Sisendkihti sisestatakse ning väljundkihist väljastatakse tulemus. Tehisnärvivõrku treenides kõrvutatakse saadud tulemust soovitud tulemusega ning leitakse klassifitseerimisviga, mida rakendatakse tagasilevis (ingl. k. *backpropagation*), et arvutada tehisnärvivõrgu uued kaalukoefitsiendid ja vabaliikmed. Lihtne närvivõrk on tsükliteta ehk info liigub ainult ühes suunas: sisendkihist läbi peidetud kihtide väljundkihti. [2, 4]



Joonis 3. Näide lihtsast tehisnärvivõrgust. Sisendvektori X väärtused moodustavad sisendkihi N_I . Sisendvektori väärtused korrutatakse läbi kaalukoefitsientide maatriksi W_1 väärtustega ja liidetakse vabaliikme B_1 vektori väärtused. OUT_1 aktiveerimisfunktsiooniga saadakse peidetud kihi N_H väljundid. Sama protsessi korratakse järgmises kihis ning saadakse lõpuks väljundkihi N_O väljundvektor Y . Autori koostatud.

Eelpool toodud joonisel on näide ühe peidetud kihiga ehk kahekihilisest tehisinärvivõrgust, mille matemaatilise funktsiooni võib välja kirjutada järgmiselt (Valem 3):

$$Y = OUT_2(W_2(OUT_1(W_1X + B_1)) + B_2), \quad (3)$$

kus $X = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$ on tehisinärvivõrgu sisendvektor;

$Y = \begin{bmatrix} y_1 \\ \dots \\ y_m \end{bmatrix}$ on tehisinärvivõrgu väljundvektor;

$W_1 = \begin{bmatrix} w_{111} & \dots & w_{n11} \\ \dots & \dots & \dots \\ w_{1k1} & \dots & w_{nk1} \end{bmatrix}$, $W_2 = \begin{bmatrix} w_{112} & \dots & w_{k12} \\ \dots & \dots & \dots \\ w_{1m2} & \dots & w_{km2} \end{bmatrix}$ on vastavalt esimese ja teise

(välis-) kihi kaalukoefitsientide maatriksid;

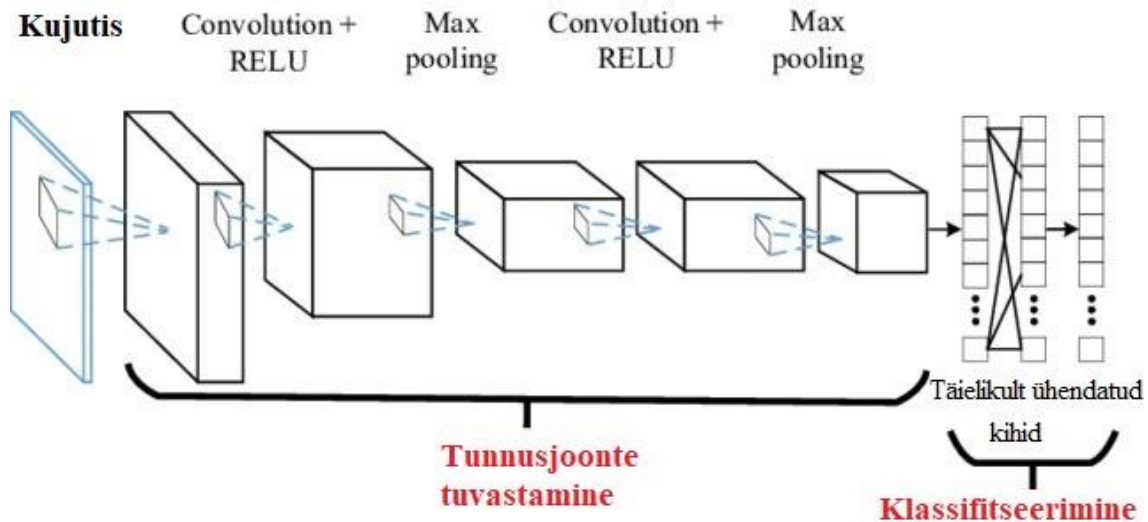
$B_1 = \begin{bmatrix} b_{11} \\ \dots \\ b_{k1} \end{bmatrix}$, $B_2 = \begin{bmatrix} b_{12} \\ \dots \\ b_{m2} \end{bmatrix}$ on vastavalt esimese ja teise (välis-) kihi vabaliikmete vektorid;

OUT_1 , OUT_2 on vastavalt esimese ja teise (välis-) kihi aktiveerimisfunktsioonid [2].

Sisend- ja väljundkihtide disain on üpris selgesõnaline, kuid peidetud kihtide disainimisel pole juhiseid võimalik paari lausega kokku võtta. Siin oleneb, millise probleemi lahendamiseks tehisinärvivõrku kasutatakse ja vastavalt sellele pannakse paika kihtide arv, nende tüüp, neuronite arv kihtides, kasutatavad aktiveerimisfunktsioonid, kulufunktsioon ja muud parameetrid [4].

2.2 Konvolutsiooniline võrk

Konvolutsioonilised võrgud (ingl. k. *convolutional neural networks*) on tehisinärvivõrgud, mis on väga edukalt rakendust leidnud pildi- ja videotuvastuses (Joonis 4). Nende roll on teisendada andmed väiksemaks nii, et neid oleks lihtsam töödelda, samas kaotamata tunnuseid, mis on olulised õige tulemuse saamiseks [11].



Joonis 4. Näide konvolutsioonilisest võrgust. Kujutise tunnusjooned tuvastatakse vaheldumisi paiknevate konvolutsiooniliste kihtide (*Convolution*) ja maksimaalsete ahenduskihtidega (*Max pooling*). Seejärel klassifitseeritakse, kasutades täielikult ühendatud kihte. RELU – tehisneuronite aktiveerimisfunktsioonina kasutatav mittenegatiivne lineaarfunktsioon. [12]

Nendes võrkudes eelneb lihtsatele peidetud kihtidele tavaliselt üks või mitu vaheldumisi paiknevat konvolutsioonilist kihti ja ahenduskihti (ingl. k. *pooling layer*) [4, 13]. Nagu nimigi viitab, kasutatakse konvolutsioonilistes kihtides konvolutsioonteisendust. Konvolutsioonteisendus (Valem 4) on integraal kahe funktsiooni korrutisest, kus ühe funktsiooni argumenti on peegeldatud ja nihutatud [13].

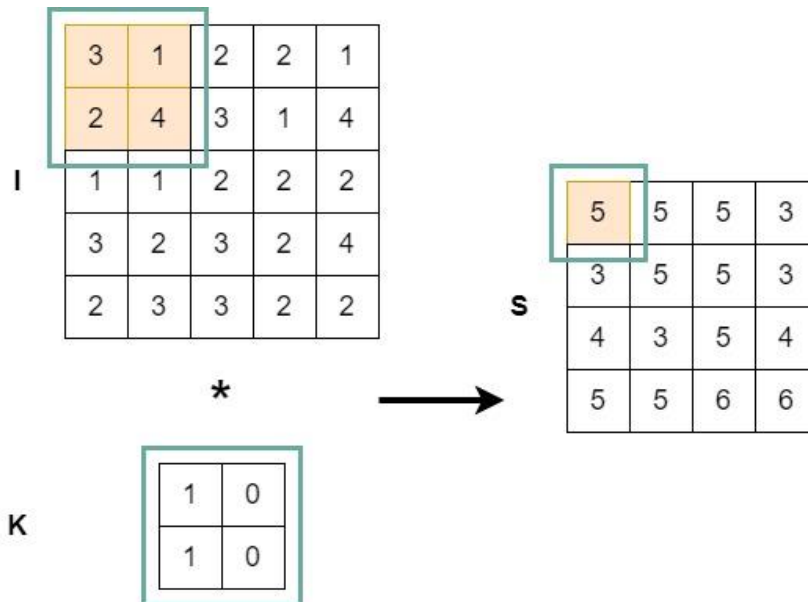
$$s(t) = (x * k)(t) = \int x(a)k(t - a)da \quad (4)$$

Konvolutsiooniliste võrkude kontekstis nimetatakse argumenti $x(a)$ sisendiks ja teist argumenti $k(t-a)$ kerneliks; väljundit $s(t)$ nimetatakse tunnuskaardiks (ingl. k. *feature map*). Tavaliselt kasutatakse konvolutsioonteisendust 2D-kujutistel ning rohkem kui üks

rida korruga (Joonis 5). Hoolimata nimest kasutatakse teisendustel rohkem ristkorrelatsiooni (Valem 5), mis on sarnane konvolutsiooniga. [13]

$$S(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i + m, j + n)K(m, n), \quad (5)$$

kus I on 2D-kujutis, K on kernel ja S tunnускаart.



Joonis 5. Näide konvolutsioonist. 2×2 suuruse kerneliga K liigutakse rea ja veeru kaupa üle sisendi I. Väljundi S esimene liige saadakse $3 \times 1 + 2 \times 1 + 1 \times 0 + 4 \times 0 = 5$, järgmine $1 \times 1 + 4 \times 1 + 2 \times 0 + 3 \times 0 = 5$ jne. Autori koostatud.

Konvolutsiooniliste kihtide eesmärgiks on filtreerida kujutiselt välja tähtsaid tunnusoone, nagu servad, nurgad või muud kujundid [4]. Üks konvolutsiooniline kiht koosneb mitmest tunnускаardist, iga tunnускаart keskendub teoorias erinevate tunnuste otsimisele [4, 13].

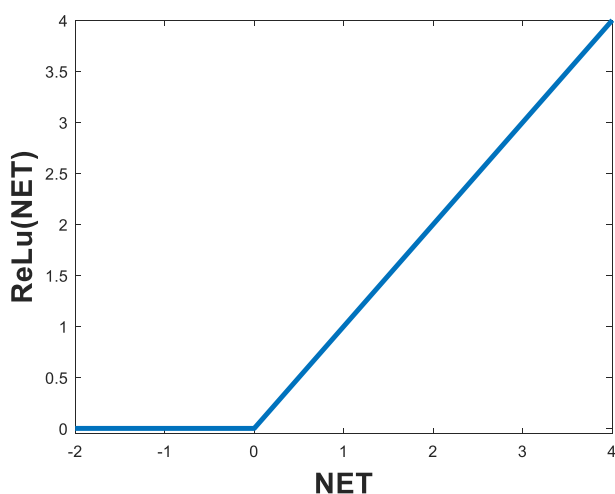
Ühe tunnускаardi kõik tehisneuronid jagavad samu kaalusid ehk kernelit ja vabaliiget [4]. Seega tehisneuroni väljund on kirjeldatud järgmiselt (Valem 6):

$$OUT_{ij} = f(NET_{ij}) = f(S(i, j) + b), \quad (6)$$

kus b on vabaliige ja f aktiveerimisfunktsioon.

Levinud aktiveerimisfunktsiooniks konvolutsiooniliste kihtide korral on mittenegatiivne lineaarfunktsioon (ingl. k. *rectified linear unit*, *ReLU*). See funktsioon (Joonis 6) tagab, et tunnускаardid on alati positiivsete väärtustega (Valem 7). [4]

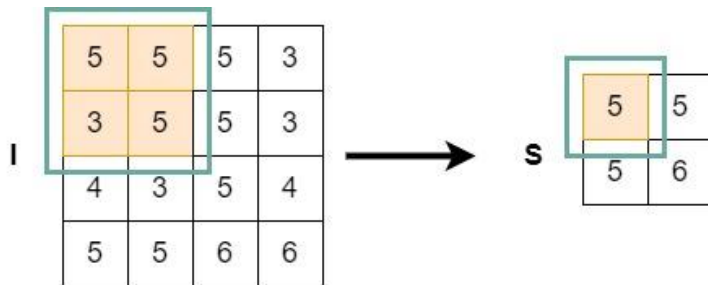
$$ReLU(NET) = \max(0, NET) \quad (7)$$



Joonis 6. Mittenegatiivne lineaarfunktsioon. Funktsioon tagab, et väljund on alati positiivne. Autori koostatud.

Konvolutsioonilisele kihile järgneb tavaliselt ahenduskiht, mille ülesanne on lihtsustada konvolutsioonilise kihi väljundit. Põhimõtteliselt tehakse kontsentreeritud tunnускаardid: kõik tunnускаardid käiakse mingi aknaga läbi ja selles aknas jäetakse alles ainult üks väärtus. Maksimaalse väärtuse korral on tegemist maksimaalse ahendusega (ingl. k. *max pooling*) (Joonis 7), kui võetakse keskmine väärtus, on tegemist keskmistava ahendusega (ingl. k. *average pooling*). Ahenduskihi eesmärk on lihtsustada konvolutsiooniliste kihtide väljundeid veelgi. Mõte seisneb selles, et kokkuvõttes pole oluline teada

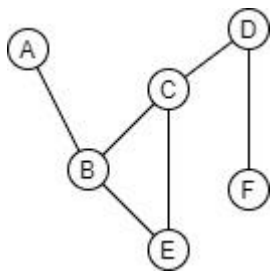
tunnusjoonte täpset asukohta, vaid nende paiknemist võrreldes teiste tunnusjoontega. Seega saame arvutusprotsessi nii kiirendada. [4, 13]



Joonis 7. Näide maksimaalsest ahendusest. 2×2 suuruse kerneliga käiakse üle sisendi I. Väljundi S esimene liige saadakse sisendi I esimese nelja liikme maksimaalse väärtuse pealt, väljundi järgmine liige sisendi järgmise kasutamata nelja liikme pealt jne. Autori koostatud.

2.3 Graaf

Graaf (Joonis 8) koosneb tippudest ja tippusid ühendavatest servadest. Tipud ja servad võivad tähistada näiteks linnasid ja neid ühendavaid teid, arvuteid ja nende ühendusi või inimesi ja nende omavahelisi tutvuseid. Sealjuures pole oluline servade pikkus, vaid tipud, mida serv omavahel ühendab. Definiitsiooni järgi on graaf „paar $G = (V, E)$, kus V on mittetühi hulk ning E hulk, mille elementideks on hulga V kaheelemendilised alamhulgad.“ Hulga V moodustavad tipud ja hulga E servad. Allpool toodud näite (Joonis 8) korral: $V = \{A, B, C, D, E, F\}$; $E = \{\{A, B\}, \{B, C\}, \{B, E\}, \{C, D\}, \{D, F\}\}$. [14]



Joonis 8. Näide graafist. Autori koostatud.

Eelpool toodud definiitsiooni korral ei tohi graafis olla silmuseid (serv, mis ühendab tippu iseendaga) ega kordseid servasid (kahte tippu ühendab omavahel rohkem kui üks serv). Sellisel juhul on tegemist multigraafiga. [14]

Graafe saab iseloomustada erinevate näitajatega, näiteks tipu aste, naabrusmaatriks ja ahela pikkus. Üks tähtsamaid näitajaid on tipu aste, mis näitab tipust väljuvate servade arvu. Eelpool kujutatud graafil (Joonis 8) on näiteks tipu B aste 3. Graafi, mille kõikide

tippude astmed on võrdsed, nimetatakse regulaarseks graafiks. Naabrusmaatriks on $n \times n$ maatriks $A = (a_{ij})$, kus $a = 1$, kui tippude v_i ja v_j on serv, ning $a = 0$, kui vastavate tippude vahel puudub serv. Eelpool kujutatud graafi (Joonis 8) naabrusmaatriks oleks järgmine:

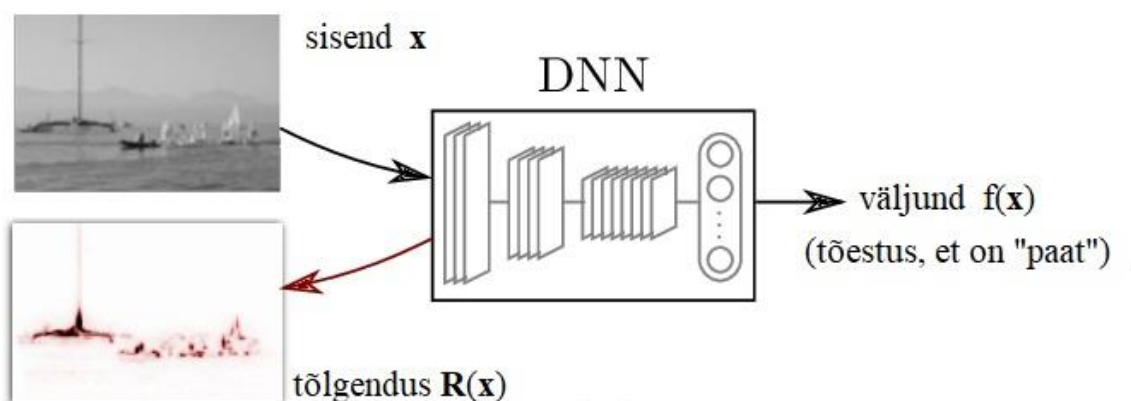
$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}. [14]$$

Ahel on tippude järjend v_0, v_1, \dots, v_k , kus iga kaks järjestikkust tippu on omavahel servaga ühendatud. Lihtahela korral on kõik ahela tipud erinevad. Ahela pikkuse määrab ära ahela servade arv. Näiteks eelpool kujutatud joonisel on ahela ABCDF pikkus 4. [14]

3 Teoreetilised alused

3.1 Tehisnärvivõrkude tõlgendamine

Tehisnärvivõrkude tõlgendamine on viimasel ajal üha enam populaarsust kogunud. Publikatsioone on viimase paari aasta jooksul ilmunud märkimisväärselt palju [3, p. 3]. Kuna teema on suhteliselt uus, on ka lähenemisi probleemile mitmesuguseid [3]. Levinud lähenemine on arvutada igale sisendile x_i nn. „olulisuse mõõde“ R_i . Kuna tehisnärvivõrgu väljundit y saab kirjeldada kui sisendi funktsiooni $f(x)$, arvutatakse R_i igale sisendile x_i vastavalt sellele, kui suurt rolli ta funktsiooni $f(x)$ tõlgendamisel mängis (Joonis 8Joonis 9) [15].



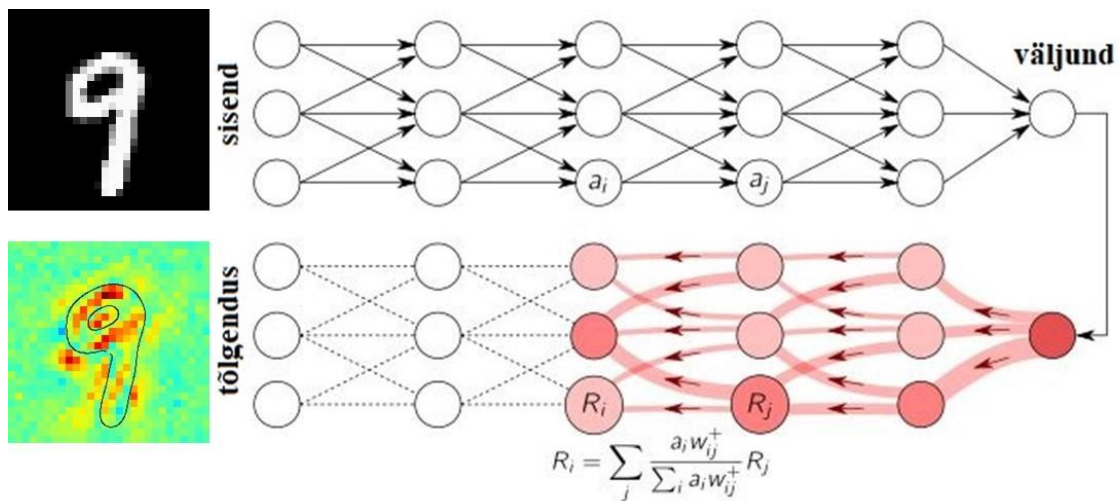
Joonis 9. Süvavõrgu (DNN) ennustuse „paat“ tõlgendus sisendi x korral. Sisendisse antakse pilt, mille süvavõrk klassifitseerib „paadiks“. Väljundkihi väljundil rakendatakse tagasilevi sisendisse ja kuvatakse soojuskaart, kus suurema R_i väärtusega pikslid on rohkem punased. [15]

Üks enim kasutatud meetoditest on olnud tundlikkuse analüüs (ingl. k. *sensitivity analysis*). See baseerub muutuse mõõtmisel tehisnärvivõrgu väljundis, kui muudetakse sisendikihti antavaid väärtuseid. Mida suurem on muutus väljundis, seda tundlikum on see antud muutusele sisendkihis [16]. Kõige suurema olulisuse mõõtmega R_i on need sisendid x_i , mille muutustele väljund kõige tundlikum on [15].

Tundlikkuse analüüs on varasemalt kasutatust leidnud ka meditsiinis [17, 18]. Hunter *et al.* [17] kasutasid tehisnärvivõrku, et hinnata patsientide ellujäämise tõenäosust nende anamneesi pealt. Seejärel rakendati tundlikkuse analüüsi, et välja selgitada, missugused andmed mõjutavad ellujäämise tõenäosust kõige rohkem. Leiti, et selleks on patsiendi vanus, samas kui patsiendi sugu ei mänginud mingit rolli. Khan *et al.* [18] kasutasid tehisnärvivõrku, et klassifitseerida pahaloomulised kasvaja spetsiifilistesse diagnostilistesse kategooriatesse kasvajate geenisignatuuride alusel. Seejärel leiti

tundlikkuse analüüsi kasutades geenid, mis osutasid kasvajate klassifitseerimisel määravaks. Niiviisi leiti need 96 geeni 2308 algsest, millega tehisnärvivõrku treenides esialgne klassifitseerimistäpsus säilis.

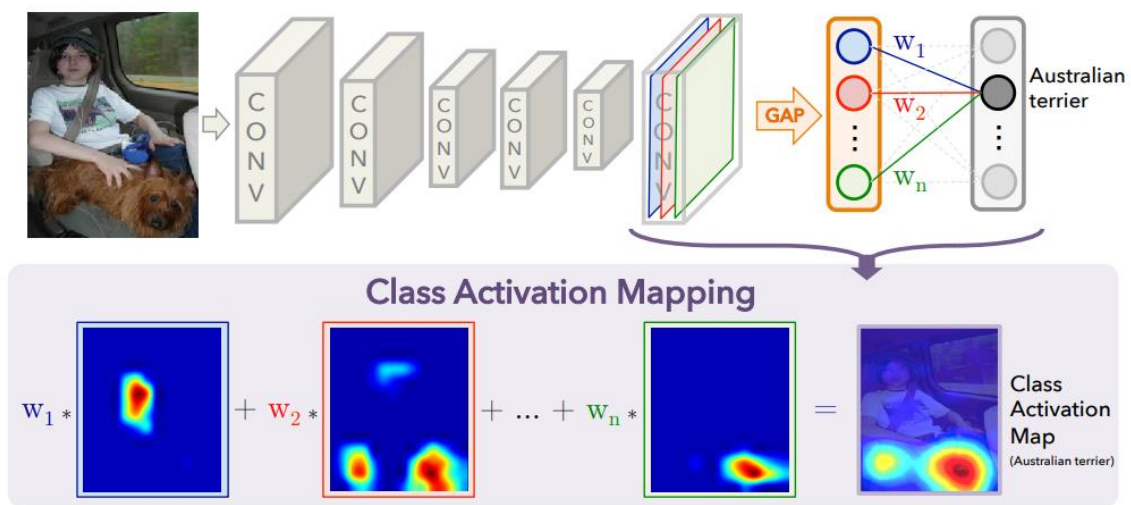
Siiski on näitlikustatud, et tundlikkuse analüüsil võivad olla omad puudused [15]. Meetodit on seega ka edasi arendatud ja jõutud muuhulgas palju kasutust leidvate meetoditeni nagu Tayloriga dekompositsiooni (ingl. k. *Taylor decomposition*) ja kihi kaupa olulisuse levi meetod (ingl. k. *layer-wise relevance propagation, LRP*) [15, 19]. Tayloriga dekompositsiooni meetodil on lisaks tundlikkusele oluline ka sisendi x_i suurus. See tähendab, et kui sisendi väärtus on võrdne nulliga, puudub ka olulisus, kuigi tundlikkuse analüüsiga võiks olulisus esineda [15]. Nagu nimigi viitab, leitakse LRP meetodil olulisuse mõõde sisendile väljundist kihi kaupa tagasi liikudes (Joonis 10). Niiviisi on tehisneuronitel, mis panustavad rohkem järgmiste kihtide tehisneuronitesse, ka suurem olulisuse mõõde [15, 19, 20]. Kuigi uuringuid, mis annaks objektiivse hinnangu, millist meetodit eelistada, ei ole palju, on toodud välja, et LRP meetodiga saadud soojuskaardid on vähem müra [20].



Joonis 10. LRP näide. Tehisneuroni olulisus R_i sõltub aktivatsioonist a_i , kaalukoefitsiendist w_{ij} ja kõrgema kihi tehisneuroni olulisusest R_j . Tõlgenduse soojuskaart näitab, millised pikslid (punased) said määravaks numbri klassifitseerimisel klassi „9“. [19]

Hiljuti on pildituvastuseks mõeldud konvolutsiooniliste võrkude tõlgendamisel kasutatud niinimetatud klassi aktiveerimiskaardi meetodit (ingl. k. *class activation maps, CAM*) ja selle variatsioone, seda ka meditsiinis [21, 22, 23, 24]. Vastava klassi aktiveerimiskaardil kujutatakse pildi piirkondi, millele kasutatav konvolutsiooniline võrk klassifitseerimisel tähelepanu pöörab (Joonis 11) [25]. Zhao *et al.* [22] kasutasid

CAM'i variatsioone, et kuvada soojuskaarte spetsiaalse meditsiiniseadmega (ingl k. *cellular electron cryotomography*) saadud makromolekulide piltidest. Kõrgema intensiivsusega piirkonnad soojuskaardil viitasid makromolekulide esinemisele. Sarnase lähenemisega otsisid Izadyyazdanabadi *et al.* [23] veidi teise CAM'i variatsiooniga soojuskaardidelt laser-endomikroskoobiga saadud piltidelt inimese aju kasvaja gliome. Leiti, et antud meetod tuvastas gliomide paiknemist sarnaselt neurokirurgile. Tang *et al.* [24] kasutasid järjekordset CAM'i variatsiooni näitamaks, et tehiskärvivõrgud klassifitseerivad pildidel patoloogiat hinnates kokkulepitud patoloogiliste tunnuste esinemist pildidel.



Joonis 11. Näide CAM'ist. Soojuskaart tuletatakse CNN'i globaalsest keskmisest ahenduskihist (ingl. k. *global average pooling layer, GAP*). Soojuskaart näitab, mille alusel klassifitseeris CNN pildil oleva Austraalia terjeri (*Australian terrier*). w_1, w_2, w_n on kaalukoefitsiendid. [25]

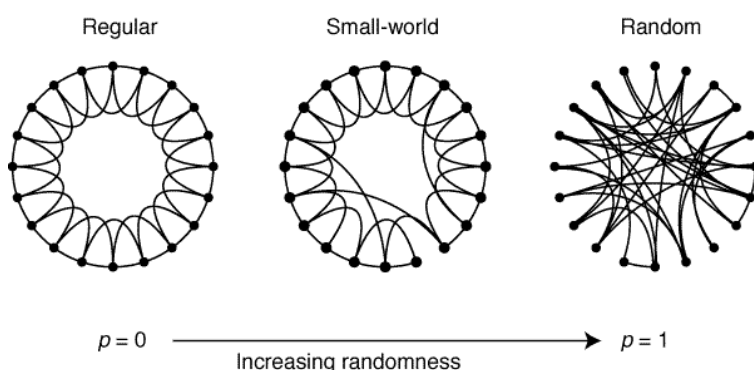
On mitmeid uuringuid, kus on analüüsitud tehiskärvivõrkude kaalukoefitsiente. Nendes uuringutes on hinnatud iga tehisneurooni suhtelist olulisust väljundkihi tulemustele vaadates otseselt kaalukoefitsientide suuruseid [26, 27, 28]. Olden *et al.* [26] ennustasid tehiskärvivõrgu abil liigirikust järvedes, tuginedes kaheksale muutujale. Lisaks tundlikkuse analüüsile kasutati ka Garsoni algoritmi, et selgitada, millised muutujad liigirikuse ennustamisel kõige suurema kaaluga olid. Garsoni algoritmi korral eraldatakse kaalukoefitsiendid ja määratakse nende absoluutväärtuste pealt iga muutuja suhteline olulisus väljundi kujunemisel. Statistiliselt oluliseks tulemuseks osutusid siiski ainult muutujad, mis polnud kolme kõige tähtsama seaski. Teises uuringus võrdlesid Olden *et al.* [27] erinevaid meetodeid, millega on võimalik hinnata sisendite olulisust väljundile. Meetodeid võrreldi andmetel, mis olid simuleeritud Monte Carlo meetodiga. Kõige täpsemalt suutis simuleeritud andmetel hinnata sisendite olulisust meetod, mis

määrab seda kaalukoefitsientide reaalsete väärtuste pealt. Kõige viletsama tulemuse saavutas aga Garsoni algoritm.

3.2 Väikese maailma omadustega võrgud

3.2.1 Väikese maailma omadused

Kompleksed võrgud on reeglina suured ja keerulised graafid, mille omadusi ei saa hinnata üksikute tippude või servade pealt [10]. Lisaks on täheldatud, et selliseid kompleksseid võrke ei saa modelleerida samal viisil nagu regulaarseid või täielikult juhuslikke graafe. Teljel (Joonis 12), kus viimased moodustavad kaks äärmust, jäävad sellised kompleksed võrgud nende keskele ja nimetatakse väikese maailma omadustega võrkudeks. Taolistel võrkudel on väikese maailma omadused: tipud on koondunud eraldi klastritesse (võrkudel on suur klasterdamise kordaja), kuid klastritevahelised ahela pikkused ja seeläbi ka võrkude keskmine lühima ahela pikkus (teisiseõnu karakteristik teepikkus) on väike. Esimest omadust peetakse omaseks regulaarsetele graafidele ja teist täielikult juhuslikele graafidele [29].



Joonis 12. Väikese maailma omadustega võrk (Small-world). Teljel (Increasing randomness), kus ühes otsas on täielikult regulaarne võrk (Regular) ja teises otsas täielikult juhuslik võrk (Random), jääb väikese maailma omadustega võrk keskele. [29]

Väikese maailma omadusi on täheldatud väga paljudes inimtekkelistes konstruktsioonides, nagu Veeb, signaalide ülekandevõrgud, metroovõrgustikud ning sotsiaalvõrgustikud Facebook ja Twitter [30, 31]. Nende võrkude iseorganiseerumine väikese maailma omadustega võrkudeks viitab info kiiremale levikule taolise konstruktsiooniga võrkudes [31].

Väga palju on teema kajastust leidnud aju närvivõrkude puhul. On täheldatud, et väikese maailma omadustega aju närvivõrgud esinevad nii kassidel, makaakidel [32] kui ka

inimestel [33, 34]. Veelgi enam, suurele hulgatele uuringutele toetudes on hetkel järeldatud, et väikese maailma omadus on peaaegu kõikidele närvisüsteemidele omane. On arutletud, et taoline nähtus on evolutsiooniliselt välja kujunenud, viidates sellega taoliste närvisüsteemide paremale kohanemisvõimele ja optimaalsemale andmete töötlusele [34, 35].

3.2.2 Väikese maailma omaduste hindamine

Muidugi tõstatub küsimus, mille suhtes võrrelda võrkude väikese maailma omadusi ning kui suur vahe on piisav [36]? Võrkude väikese maailma omadust saab hinnata, näiteks võrreldes võrkusid juhuslikustatud võrkudega. Võrreldes juhuslikustatud võrkudega on väikese maailma omadustega võrkudel sarnane karakteristiklik teepikkus ja palju suurem klasterdamise kordaja [8, 37, 29]. See tähendab, et kui uuritav võrk on väikese maailma omadustega, on tema normeeritud klasterdamise kordaja $\gamma \gg 1$ (Valem 8) ja normeeritud karakteristiklik teepikkus $\lambda \geq 1$ (Valem 9) [29, 37].

$$\gamma = \frac{C}{C_{rand}}, \quad (8)$$

kus γ on normeeritud klasterdamise kordaja, C on uuritava võrgu klasterdamise kordaja ja C_{rand} juhuslikustatud võrgu klasterdamise kordaja.

$$\lambda = \frac{L}{L_{rand}}, \quad (9)$$

kus λ on normeeritud karakteristiklik teepikkus, L on uuritava võrgu karakteristiklik teepikkus ja L_{rand} on juhuslikustatud võrgu karakteristiklik teepikkus.

Eelmise kahe valemi põhjal on defineeritud väikese maailma mõõde S (Valem 10), millega leitakse normeeritud klasterdamise kordaja ja karakteristikliku teepikkuse suhe [37].

$$S = \frac{\gamma}{\lambda} = \frac{C/C_{rand}}{L/L_{rand}}, \quad (10)$$

kus S on võrgu väikese maailma mõõde, C_{rand} ja L_{rand} on juhuslikustatud võrgu klasterdamise kordaja ja karakteristiklik teepikkus. Kui $S \gg 1$, siis on uuritav võrk väikese maailma omadustega [37].

3.3 Väikese maailma omadused tehisnärvivõrkudel

Töö autorile teadaolevatel andmetel ei ole kirjanduses viiteid, kas komplekssete võrkude analüüsimetodeid on kasutatud tehisnärvivõrkude tõlgendamisel. Varasemalt on kasutatud graafiteooriat tehisnärvivõrkudel selleks, et konstrueerida topoloogiliselt väikese maailma omadustega tehisnärvivõrkusid [38, 39]. Erkaymaz *et al.* [38] koostasid topoloogiliselt väikese maailma omadustega tehisnärvivõrgud nii, et kahe peidetud kihiga tehisnärvivõrgus ühendati lahti juhuslik arv ühendusi tehisneuronite vahel ning ühendati omavahel tehisneuronid, mis varasemalt polnud ühenduses. Niiviisi loodi tehisnärvivõrgud, kus osad tehisneuronid olid ühendatud ülejäärgmise või veelgi kaugemate kihtide tehisneuronitega. Autorid näitasid, et taolised tehisnärvivõrgud on parema klassifitseerimistäpsusega kui tavapärased lihtsad tehisnärvivõrgud.

Javaheripi *et al.* [39] kohendasid meetodikat nii, et seda saaks kasutada ka konvolutsioonilistel tehisnärvivõrkudel. Sellega näidati, et väikese maailma topoloogiaga konvolutsioonilised tehisnärvivõrgud saavutavad sarnase klassifitseerimistäpsuse teiste üldlevinud ehitusega konvolutsiooniliste tehisnärvivõrkudega, kuid seda 2 korda kiirema arvutuse või 10 korda vähemate parameetrite korral.

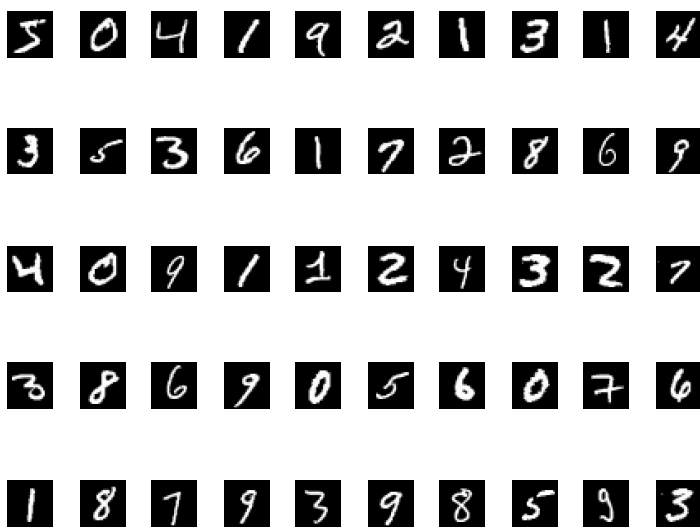
Kuigi meile teadaolevalt pole varasemalt tehisnärvivõrkude tõlgendamisel komplekssete võrkude analüüsimetodeid kasutatud, kasutatakse neid pidevalt bioloogiliste närvivõrkude uurimisel [34]. Sellest ajast, kui 1998. aastal mõistet tutvustati, on võrkude väikese maailma omadusi empiirilistes uuringutes kajastatud peaaegu sama tihedalt kui teisi põhilisi omadusi nagu võrgu suurus ja tihedus [36]. Tehisnärvivõrke on ajalooliselt koostatud, inspireerituna näidetest, kuidas bioloogilised närvivõrgud andmeid töötlevad [40]. Ajus kantakse signaale edasi ühelt neuronilt teisele sünapside kaudu, kus signaali kas erutatakse või pidurdatakse. Analoogiliselt antakse tehisnärvivõrgus info edasi ühelt neuronilt järgmisele, seda vahepeal vastavalt kaalukoefitsiendi suurusele võimendades või summutades. Seepärast võiksime eeldada, et omadused, nagu väikese maailma omadus võiksid esineda ka tehisnärvivõrkudel. Argumenteeritakse veel, et universaalsete tehisnärvivõrkude arendamisel oleme bioloogilistelt närvivõrkudelt õppinud kõik, mida võimalik. Sealjuures on ka diskuteeritud, et ükski tehisnärvivõrk ei saa kunagi olla täiesti universaalne, kuna nende modelleerimisel lähtutakse alati inimaju võimekusest ja inimajule sarnase võimekusega tehisnärvivõrk on seega paratamatult sarnaselt struktureeritud [40].

4 Metoodika

4.1 Kasutatavad andmed ja tehisnärvivõrgud

4.1.1 MNIST andmebaas

MNIST andmebaas [41] (ingl. k. *Modified National Institute of Standards and Technology database*) koosneb 28×28 pikslisuurustest must-valgetest piltidest, kus kujutatakse käsitsi kirjutatud numbreid vahemikus 0...9 (Joonis 13). Treeningkomplekt koosneb 60000 numbrist. 30000 numbrit on kogutud USA Statistikaameti (*U.S. Census Bureau*) töötajatelt, teine osa komplektist on kogutud USA keskkooli õpilastelt. Testkomplekt koosneb 10000 numbrist, millest pooled on samuti USA Statistikaameti töötajatelt ja pooled keskkooli õpilastelt kogutud. On tagatud, et treeningkomplekti ja testkomplekti numbrite autorid ei kattu. Andmebaas sobib erinevate mustrituvastusmeetodite katsetamiseks reaalse elu andmete peal, mis ei nõua kasutajalt andmete suurt eeltöötlust. Tihtipeale võetakse treeningkomplektist välja 10000 numbrit, et moodustada valideerimiskomplekt [4].



Joonis 13. Väljavõte MNIST andmebaasi numbritest. Autori koostatud.

4.1.2 Rindkere röntgenülesvõtete andmebaas

Käesolevas töös kasutatud rindkere ülesvõtete andmebaas [42] koosneb treening-, valideerimis- ja testandmeteks jaotatud 5863 rindkere AP (eest-taha otseülesvõte, lad. k. *anterior posterior*) röntgenülesvõttest, mis on jagatud kahte klassi: patoloogiata ja kopsupõletikuga. Ülesvõtted on saadud rutiinse töö käigus patsientidelt vanuses 1 kuni 5

aastat Guangzhou naiste ja laste haiglas Hiinas. Andmebaasist on välja jäetud viletsa kvaliteediga ja loetamatud ülesvõtted. Kõiki ülesvõtteid on hinnanud kaks radioloogit. Ülesvõtted on salvestatud *jpeg*-formaadis, neid on kokku 5856: 1583 patoloogiata ning 4273 kopsupõletiku diagnoosiga.

4.1.3 Numbrite klassifitseerimiseks kasutatavad tehisnärvivõrgud

Käesoleva töö raames treeniti MNIST andmebaasi numbreid klassifitseerima kaks tehisnärvivõrku. Esimesel tehisnärvivõrgul oli üks peidetud kiht 30 tehisneuroniga, teisel tehisnärvivõrgul [43] oli kaks peidetud kihti vastavalt 100 ja 80 tehisneuroniga kihis. Sisendkihis oli 784 tehisneuronit, mis vastas MNIST andmebaasi pildi pikslite arvule. Pildi piksli väärtused moodustasid sisendkihi. Väljundkihi tehisneuronite arv varieerus vahemikus 3...10, mis olenes sellest, mitut erinevat numbrit vastav tehisnärvivõrk klassifitseerima treeniti. Väikese maailma mõõtmete arvutamiseks kasutati treenitud tehisnärvivõrkude kaalukoefitsiente.

4.1.4 Röntgenülesvõtete klassifitseerimiseks kasutatav tehisnärvivõrk.

Käesoleva töö raames treeniti röntgenülesvõtteid klassifitseerima üks konvolutsiooniline võrk [44]. Kasutatud võrgu sisendkihi suurus oli 200×200 tehisneuronit. Selleks töödeldi ka kõik röntgenülesvõtted ümber, et pikslite arv vastaks sisendkihi tehisneuronite arvule. 200×200 suurusega röntgenülesvõtete piksli väärtused moodustasid sisendkihi. Edasi koosnes võrk kolmest konvolutsioonilisest kihist vastavalt 27, 9 ja 3 tunnuskaardiga vaheldumisi kolme maksimaalse ahenduskihiga. Konvolutsiooniliste kihtide kerneli suurus oli 3×3 ja ahenduskihtide kerneli suurus 2×2 . Neile järgnes 64 tehisneuroniga täielikult ühendatud peidetud kiht ning lõpuks 1 tehisneuroniga väljundkiht.

4.2 Arvutusteks kasutatud tarkvara

Röntgenpiltide klassifitseerimiseks mõeldud tehisnärvivõrgud treeniti Kaggle [45] veebikeskkonnas. MNIST andmebaasi numbrite klassifitseerimiseks mõeldud tehisnärvivõrkude treenimiseks ja kõikide kaalukoefitsientidelt saadud tulemuste arvutamiseks kasutati MATLAB [46] tarkvara. MATLAB on programmeerimisplatvorm, mis kasutab maatriksarvutustel põhinevat keelt. MATLAB'iga on võimalik teostada statistilist analüüsi, masinõpet, signaalitöötlust ja palju muud. Lisaks on MATLAB'i

võimalik liidestada teiste programmeerimiskeeltega. Käesolevas töös kasutati MATLAB'i versiooni R2020a.

4.2.1 Brain Connectivity Toolbox

Brain Connectivity toolbox [47] sisaldab MATLAB'is kasutatavaid funktsioone, millega on võimalik teostada kompleksset võrkude analüüsi. Käesolevas töös kasutati nendest funktsioone, millega saab arvutada igale graafi tipule (käesolevas töös tehisneuronile) väikese maailma mõõtme (Valem 11).

$$S = \frac{1}{n} \sum_{i \in N} S_i = \frac{C_i / C_{i \text{ rand}}}{L / L_{\text{rand}}}, \quad (11)$$

kus S on keskmine väikese maailma mõõde üle kõigi tippude, S_i on tipu i väikese maailma mõõde, C_i on tipu i klasterdamise kordaja, L on karakteristiklik teepikkus, $C_{i \text{ rand}}$ ja L_{rand} on vastava juhuslikustatud võrgu klasterdamise kordaja ja karakteristiklik teepikkus. Kui $S_i \gg 1$, siis on uuritav võrk väikese maailma omadustega. Tipu i klasterdamise kordaja saadi järgneva valemiga (Valem 12):

$$C_i = \sum_{i \in N} \frac{2t_i}{k_i(k_i - 1)}, \quad (12)$$

kus C_i on tipu i klasterdamise kordaja, t_i on kolmnurkade arv ümber tipu i (Valem 13) ja k_i on tipu aste (Valem 14).

$$t_i = \frac{1}{2} \sum_{j, h \in N} a_{ij} a_{ih} a_{jh}, \quad (13)$$

$$k_i = \sum_{j \in N} a_{ij}, \quad (14)$$

kus a on binaarne väärtus, mis näitab serva olemasolu kahe tipu vahel. Karakteristlik teepikkus saadi järgmise valemiga (Valem 15):

$$L = \frac{1}{n} \sum_{i \in N} L_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}}{n-1}, \quad (15)$$

kus L on karakteristiklik teepikkus, n on tippude arv, L_i on keskmine teepikkus tipu i ja kõigi teiste tippude vahel ja d_{ij} on lühim teepikkus tippude i ja j vahel (Valem 16).

$$d_{ij} = \sum_{a_{uv} \in g_{i \leftrightarrow j}} a_{uv}, \quad (16)$$

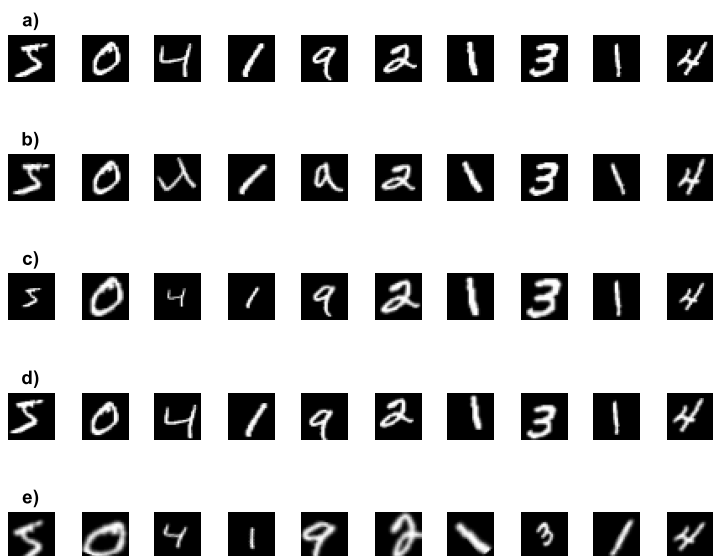
kus $g_{i \leftrightarrow j}$ on lühim tee tippude i ja j vahel.

4.3 Ülevaade protsessist

Käesolevas töös püstitati hüpotees: mida keerulisemaid andmeid on tehisnärvivõrk treenitud klassifitseerima, seda suurem on tema kihtide kaalukoefitsientidel arvutatud keskmine väikese maailma mõõde. Selleks, et hinnata püstitatud hüpoteesi, tuli treenida tehisnärvivõrkusid klassifitseerima järjest keerukamaid andmeid.

4.3.1 MNIST andmebaasi treeningandmete töötlemine ja tehisnärvivõrkude treenimine

MNIST andmebaasi numbrite korral muudeti kasutatavaid treeningandmeid kahel meetodil. Esimese meetodiga üritati muuta tehisnärvivõrkude ülesanne raskemaks treeningandmetega manipuleerides. Esimese tehisnärvivõrgu jaoks kasutati algseid treeningandmeid, järgmise tehisnärvivõrgu jaoks roteeriti numbreid suvaliselt vahemikus -45° ... 45° , kolmanda jaoks muudeti numbrite suurust vahemikus $0,5$... $1,5$ korda, neljanda jaoks nihutati numbreid 3 piksli ulatuses nii horisontaalsel kui ka vertikaalsel teljel (Joonis 14). Viienda tehisnärvivõrgu korral kasutati andmete manipuleerimiseks kõiki eelpool kirjeldatud võtteid.



Joonis 14. Väljavõte manipuleeritud treeningandmetest: a) algseid; b) roteeritud; c) skaleeritud; d) nihutatud; e) kombineeritud. Autori koostatud.

Pärast andmetega manipuleerimist eraldati 10000 numbrit klassifitseerimistäpsuse hindamiseks. 50000 numbrit jäi treeningandmete hulgaks. Iga andmete komplektiga treeniti 20 tehisnärvivõrku, et saada statistiliselt usaldusväärsed tulemused. Iga

tehisnärvivõrgu treenimise jaoks manipuleeriti treeningandmeid uuesti. Seega treeniti kokku $20 \times 5 = 100$ tehisnärvivõrku.

Teise meetodiga muudeti erinevate numbrite arvu, mida tehisnärvivõrk klassifitseerima pidi. Kõigepealt treeniti tehisnärvivõrku klassifitseerima kolme erinevat numbrit, seejärel nelja ja nii edasi kuni kõikide numbriteni välja. Klassifitseeritavad numbrid valiti igakord juhuslikult. Väljavalitud numbritest eraldati $1/6$ klassifitseerimistäpsuse hindamiseks. Igat klassifitseeritavate numbrite hulka valiti 50 korda. Seega treeniti kokku $50 \times 8 = 400$ tehisnärvivõrku.

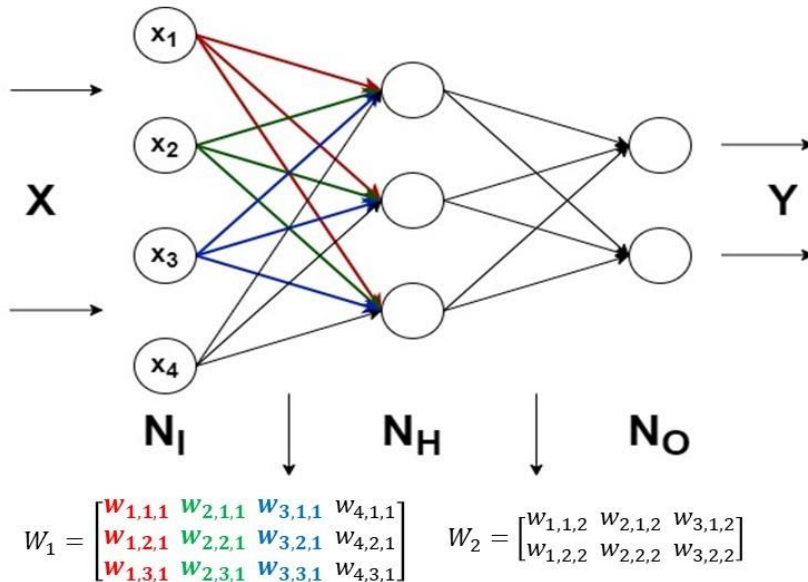
4.3.2 Röntgenülesvõtete andmebaasi treeningandmete töötlemine ja konvolutsiooniliste võrkude treenimine

Röntgenülesvõtete korral üritati konvolutsiooniliste võrkude ülesanne samamoodi raskemaks muuta treeningandmete manipuleerimisega. Enne andmete manipuleerimist eraldati kõikidest ülesvõtetest 20% testandmeteks ja allesjäänud 80%-st veelkord 20% valideerimisandmeteks. Esimese võrgu korral treeningandmetega ei manipuleeritud, teise võrgu korral roteeriti ülesvõtteid vahemikus $0 \dots 90$ kraadi. Kolmanda võrgu korral skaleeriti ülesvõtteid vahemikus $0,9 \dots 1,1$ korda, neljanda võrgu korral nihutati ülesvõtteid mööda horisontaalset ja vertikaalset telge 20 piksli ulatuses. Viienda võrgu korral kombineeriti kõiki eelnevalt kasutatud võtteid. Kõiki variante konvolutsioonilisi võrke treeniti 10 korda, iga kord manipuleeriti treeningandmeid uuesti. Kokku treeniti seega $10 \times 5 = 50$ konvolutsioonilist võrku.

4.3.3 Väikese maailma mõõtme arvutamine

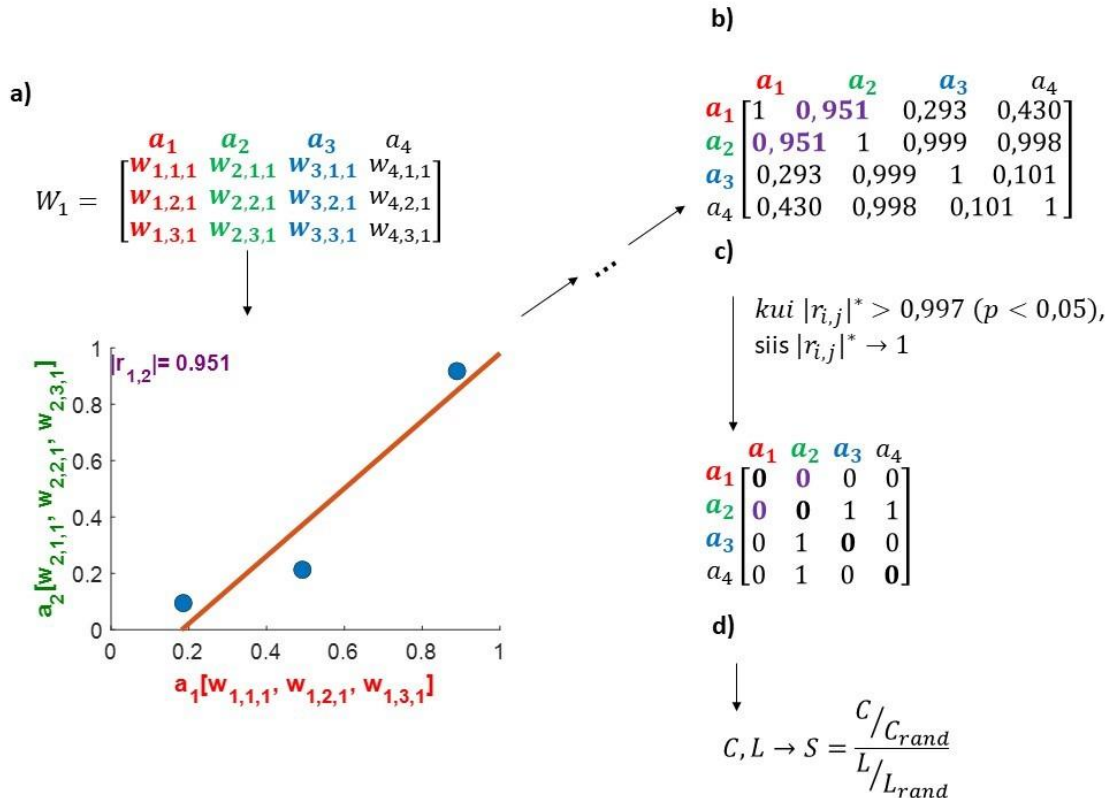
Tehisnärvivõrkudest võeti pärast nende treenimist välja kihtide kaalukoefitsiendid. Iga kaalukoefitsientide komplekti jaoks saadi kaalukoefitsientide maatriksid (Joonis 15). Esimese kihi puhul saadi maatriks suurusega $k \times n$, kus k näitas esimese peidetud kihi tehisneuronite arvu ja n sisendkihi tehisneuronite arvu (Joonis 3); teise kihi puhul saadi maatriks suurusega $m \times k$, kus m näitas väliskihi tehisneuronite arvu (kahe peidetud kihi korral teise peidetud kihi tehisneuronite arvu) ja k esimese peidetud kihi tehisneuronite arvu (Joonis 3, Joonis 15). Kahe peidetud kihiga tehisnärvivõrgu korral võeti ka kolmanda kihi kaalukoefitsientide maatriks, mille mõõtmed moodustasid vastavalt väliskihi tehisneuronite ja teise peidetud kihi tehisneuronite arv.

Röntgenülevõtteid klassifitseerima treenitud konvolutsioonilistel võrkudel võeti pärast treenimist välja viimase ahenduskihi ja täielikult ühendatud kihi vahelised kaalukoefitsiendid. Seega saadi kaalukoefitsientide maatriksid suurusega 64×1875 (vastavalt täielikult ühendatud kihi tehisneuronite arv ja ahenduskihi tehisneuronite arv).



Joonis 15. Näide kaalukoefitsientide maatriksite leidmisest lihtsast tehisnärvivõrgust. x_1, \dots, x_4 on sisendväärtused, W_1 ja W_2 on kaalukoefitsientide maatriksid (vt Joonis 3).

Saadud maatriksites kõrvutati veergude kaalukoefitsiendid (st. eelmise kihi tehisneuronite järgi jaotatult) ja leiti nendevahelised Pearsoni korrelatsioonikordajate absoluutväärtused (Joonis 16a). Sellega leiti tehisneuronid, mis on omavahel seotud, kuna nende kaalukoefitsiendid korreleeruvad ja väljundeid muudetakse ühtemoodi järgmise kihi kaalutud summaatorite jaoks. Saadi niinimetatud kaalutud naabusmaatriks, kus tippudeks olid tehisneuronid ja servadeks kaalukoefitsientide korrelatsioonikordajate absoluutväärtused (Joonis 16b).



Joonis 16. Näide kihi naabusmaatriksi kaudu väikese maailma mõõtme arvutamises: a) kaalukoefitsientide komplektid võetakse kahe veeru kaupa ja leitakse korrelatsioonikordaja absoluutväärtus; b) saadud korrelatsioonikordajate absoluutväärtused kantakse tabelisse ja saadakse nn. kaalutud maatriks; c) kaalutud maatriks teisendatakse binaarseks maatriksiks fikseeritud lävendi* meetodil, lävendiks on statistiliselt oluline korrelatsioonikordaja väärtus, mis sõltub katsepunktide arvust; d) saadud binaarse naabusmaatriksi põhjal arvutatakse klasterdamise kordaja ja karakteristiklik teepikkus ning nende kaudu väikese maailma mõõde. *Antud näite korral koosneb esimene kiht kolmest tehiseuronist, seega katsepunktide arv on kolm ja statistiliselt oluline $|r| > 0,997$.

Et arvutusprotsesse kiirendada, asendati korrelatsioonikordajate absoluutväärtused binaarsete väärtustega, et saada kaalutud graafi asemel binaarne (Joonis 16c). Kirjanduses [48, 49] on välja pakutud mitu meetodit ning rõhutatatakse, et ei ole olemas erapooletut varianti ja see oleneb suuresti võrgu topoloogiast. Üheks meetodiks on fikseeritud lävendi meetod (ingl. k. *fixed threshold*), mille käigus jäetakse alles servad, mille väärtus ületab etteantud lävendi. Lävendiks võib olla näiteks statistiline olulisus: $p < 0,05$. Selle meetodi eeliseks on, et graafi servasid ei moodusta juhuslikkusest tingitud väärtused. Teiseks populaarseks meetodiks on fikseeritud serva tiheduse meetod (ingl. k. *fixed edge density*), mille käigus jäetakse alati alles määratud protsent kõrgeima väärtusega servadest. Selle meetod tagab graafidevahelise võrreldavuse sama arvu servade läbi, kuid on oht, et jäetakse kõrvale statistiliselt olulised väärtused või kaasatakse statistiliselt mitteolulised väärtused. Kuna omavahel võrreldi ühesuuruseid graafe (st. kaalukoefitsientide

maatrikseid) ja tehisnärvivõrkude enda parameetreid ei muudetud, valiti fikseeritud lävendi meetod, et vältida juhuslikke servade sattumist graafidesse. Fikseeritud lävendiks oli statistiliselt oluline korrelatsioonikordaja, mis varieerus vastavalt katsepunktide arvule.

Kui saadud korrelatsioonikordaja absoluutväärtus oli statistiliselt oluline ($p < 0,05$), teisendati väärtus üheks, vastasel juhul aga nulliks (Joonis 16). 30 katsepunkti juures oli lävendiks näiteks 0,36, 64 katsepunkti juures 0,25, 100 katsepunkti juures 0,20, 80 katsepunkti juures 0,22 ja 10 katsepunkti juures 0,63. Saadi binaarne naabusmaatriks, kus tehisneuronid olid tipud ja ühed tähistasid servasid vastavate tehisneuronite vahel. Kuna tipul iseendaga ei saanud serva olla, tuli peadiagonaal asendada nullidega (Joonis 16c). Saadud binaarse naabusmaatriksi põhjal arvutati karakteristikliku tee pikkuse ning klasterdamise kordajate kaudu iga tipu väikese maailma mõõde. Kihi väikese maailma mõõtmise saamiseks arvutati keskmine väikese maailma mõõde üle kõigi tippude (Joonis 16d).

Väikese maailma mõõtmete arvutamiseks tuli genereerida igale naabusmaatriksile ka juhuslikustatud naabusmaatriksid ja leida nende põhjal juhuslikustatud graafi korrelatsioonikordajad ja karakteristiklik teepikkus ($C_{i\text{ rand}}$ ja L_{rand}). Seda tehti põhimõttel, et säiliksivad samad tipu astmed ja muudeti ainult servade asukohti [50]. Vastavat protseduuri korrati 100 korda ja võeti väikese maailma mõõtmete arvutamiseks keskmine.

Arvutatud väikese maailma mõõtmetest võeti aritmeetiline keskmine ja kuvati graafikul koos standardhälbega ja p-väärtustega. Statistiliseks testiks kasutati Wilcoxon'i astaksumma testi, mis on mõeldud sõltumatute parameetrite ja väikeste valimite testimiseks. Tulemusi on statistiliselt võrreldud esimesel meetodil (manipuleeritud treeningandmetega) algsete treeningandmetega treenitud tehisnärvivõrkudega. Teisel meetodil on statistiliselt võrreldud tehisnärvivõrkusid, millel klassifitseeritud erinevate numbrite arv erines ühe võrra. Lisaks arvutati tehisnärvivõrkude keskmised klassifitseerimistäpsused ja naabusmaatriksi tihedused.

5 Tulemused

5.1 Manipuleeritud numbrite klassifitseerimine kasutades ühe peidetud kihiga tehismärgivõrkusid

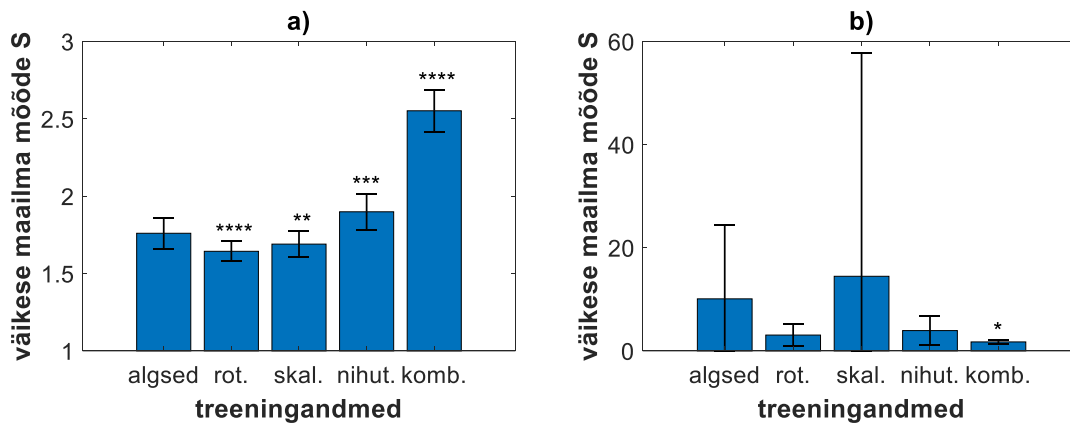
Tehismärgivõrke treeniti treeningandmetega neid roteerides, skaleerides, nihutades või eelnevaid võtteid kombineerides. Igat manipuleerimist korrati 20 korda ja seega saadi ka 20 treenitud tehismärgivõrku igat tüüpi treeningandmete korral. Treeningandmete hulk oli alati sama suur – 50000. Igast tehismärgivõrgu kihist võeti kaalukoefitsiendid, leiti korrelatsioonikordajate absoluutväärtused, mis teisendati binaarseteks naabusmaatriksiteks. Saadud naabusmaatriksitest leiti väikese maailma mõõtmel igale kihile, neist võeti aritmeetiline keskmine. Manipuleeritud treeningandmetega saadud tulemusi võrreldi statistiliselt algsete treeningandmetega treenitud tehismärgivõrkude väikese maailma mõõtmel. Statistilise testina kasutati Wilcoxon'i astak-summa testi. Järgnevas tabelis (Tabel 1) on kõigepealt toodud treenitud tehismärgivõrkude keskmised klassifitseerimistäpsused, mis on arvutatud valideerimisandmetega.

Tabel 1. Manipuleeritud andmetega treenitud ühe peidetud kihiga tehismärgivõrkude keskmised klassifitseerimistäpsused.

Treeningandmed	Klassifitseerimistäpsus (%)
algsed	93
roteeritud	86
skaleeritud	88
nihutatud	82
kombineeritud	56

5.1.1 Väikese maailma mõõde

Väikese maailma mõõde erines mõningate manipuleeritud treeningandmetega treenides oluliselt võrreldes algsete treeningandmetega esimese peidetud kihi (Joonis 17a) kaalukoefitsientide korral ($p < 0,05$). Väljundkihi (Joonis 17b) kaalukoefitsientide väikese maailma mõõtmel erinevuses puudus statistiline olulisus ($p > 0,05$).



Joonis 17. Manipuleeritud andmetega treenitud ühe peidetud kihiga tehismärgivõrkude a) 1. kihtide ja b) 2. kihtide kaalukoefitsientidel arvutatud keskmised väikese maailma mõõtmised standardhälbega (* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$; **** $p < 0,0001$). Manipuleeritud treeningandmetega saadud tulemusi on statistiliselt võrreldud algsete treeningandmetega treenitud tehismärgivõrkude tulemustega (rot. – roteeritud; skal. – skaleeritud; nihut. – nihutatud; komb. – kombineeritud).

5.1.2 Naabusmaatriksite tihedus

Korrelatsioonikoefitsiendi absoluutväärtuse tase, millest kõrgemad väärtused olid statistiliselt olulised ($p < 0,05$) ja seega teisendati ühtedeks, oli esimese kihi korral 0,36 ja teise kihi korral 0,63. Allpool tabelis (Tabel 2) on toodud binaarsetes naabusmaatriksites esinenud ühtede keskmine esinemissagedus.

Tabel 2. Manipuleeritud andmetega treenitud tehismärgivõrkude kihtide keskmised naabusmaatriksite tihedused. $|r| > \dots$ näitab lüvendit, millest suurema väärtusega korrelatsioonikordajate absoluutväärtused teisendati ühtedeks ja madalama väärtusega nullideks.

Treeningandmed	Maatriksi tihedus (%)	
	1. kiht ($ r > 0,36$)	2. kiht ($ r > 0,63$)
algsed	11	4
roteeritud	13	5
skaleeritud	16	5
nihutatud	14	6
kombineeritud	11	13

5.2 Manipuleeritud numbrite klassifitseerimine kasutades kahe peidetud kihiga tehismärgivõrkusid

Järjekordselt treniti iga tüüpi treeningandmetega 20 tehismärgivõrku, mille kaalukoefitsientidel arvutatud väikese maailma mõõtmest võeti aritmeetiline keskmine. Treeningandmete hulk oli iga kord sama suur – 50000. Manipuleeritud treeningandmetega saadud väikese maailma mõõtmist võrreldi statistiliselt algsete treeningandmetega saadud väikese maailma mõõtmist saadud tulemustega. Statistilise testina

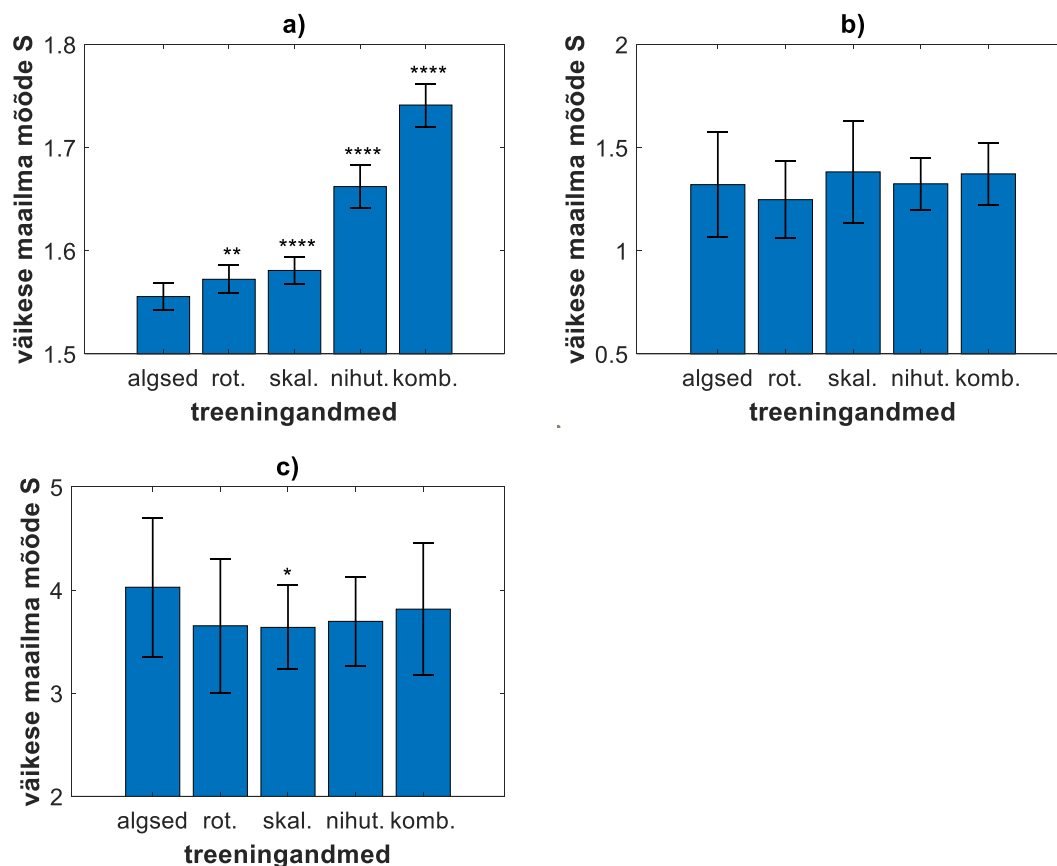
kasutati Wilcoxon'i astak-summa testi. Järgnevas tabelis (Tabel 3) on kõigepealt toodud treenitud tehismärgivõrkude keskmised klassifitseerimistäpsused, mis on saadud valideerimisandmetega.

Tabel 3. Manipuleeritud andmetega treenitud kahe peidetud kihiga tehismärgivõrkude keskmised klassifitseerimistäpsused.

Treeningandmed	Klassifitseerimistäpsus (%)
algsed	97
roteeritud	95
skaleeritud	96
nihutatud	95
kombineeritud	85

5.2.1 Väikese maailma mõõde

Väikese maailma mõõde erines kõikide manipuleeritud treeningandmetega treenides oluliselt võrreldes algsete treeningandmetega esimese peidetud kihi (Joonis 18a) kaalukoefitsientide korral ($p < 0,05$). Teise peidetud kihi (Joonis 18b) kaalukoefitsientide väikese maailma mõõtmete erinevuses puudus statistiline olulisus ($p > 0,05$). Samuti puudus statistiline olulisus väljundkihi (Joonis 18c) kaalukoefitsientide väikese maailma mõõtmete erinevuses ($p > 0,05$).



Joonis 18. Manipuleeritud andmetega treenitud kahe peidetud kihiga tehisnärvivõrkude a) 1. kihtide, b) 2. kihtide ja c) 3. kihtide kaalukoefitsientidel arvatatud keskmised väikese maailma mõõtmised standardhälbega (* $p < 0,05$; ** $p < 0,01$; **** $p < 0,0001$). Manipuleeritud treeningandmetega saadud tulemusi on statistiliselt võrreldud algsete treeningandmetega treenitud tehisnärvivõrkude tulemustega (rot. – roteeritud; skal. – skaleeritud; nihut. – nihutatud; komb. – kombineeritud).

5.2.2 Naabusmaatriksite tihedus

Korrelatsioonikoefitsiendi absoluutväärtuse tase, millest kõrgemad väärtused olid statistiliselt olulised ($p < 0,05$) ja seega teisendati ühtedeks, oli esimese kihi korral 0,20, teise kihi korral 0,22 ja kolmanda kihi korral 0,63. Allpool tabelis (Tabel 4) on toodud binaarsetes naabusmaatriksites esinenud ühtede keskmine esinemissagedus.

Tabel 4. Manipuleeritud andmetega treenitud tehisnärvivõrkude kihtide keskmised naabusmaatriksite tihedused. $|r| > \dots$ näitab ländent, millest suurema väärtusega korrelatsioonikordajate absoluutväärtused teisendati ühtedeks ja madalama väärtusega nullideks.

Treeningandmed	Maatriksi tihedus (%)		
	1. kiht ($ r > 0,20$)	2. kiht ($ r > 0,22$)	3. kiht ($ r > 0,63$)
algsed	5	6	5
roteeritud	5	6	5
skaleeritud	5	6	5
nihutatud	6	6	5
kombineeritud	6	7	5

5.3 Erineva arvu numbrite klassifitseerimine kasutades ühe peidetud kihiga tehisnärvivõrkusid

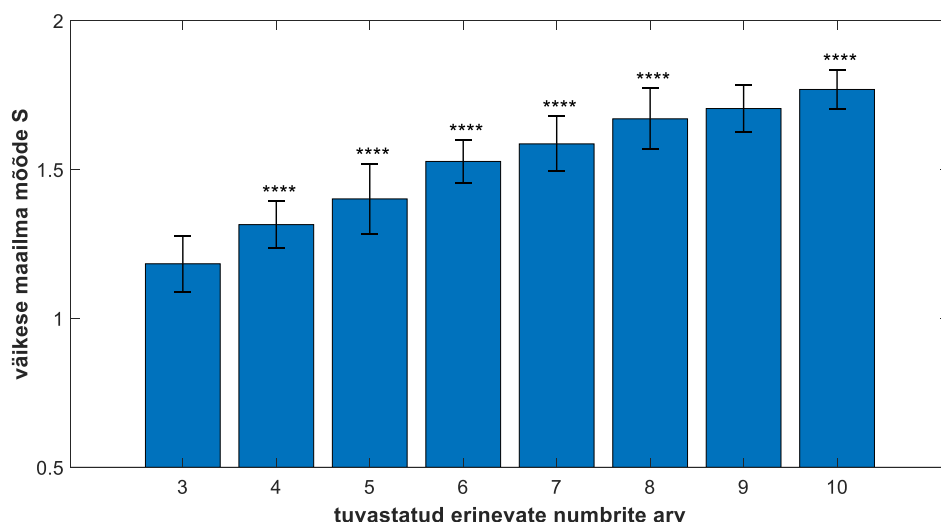
Tehisnärvivõrke treeniti nii, et iga grupp tehisnärvivõrkusid klassifitseeris erineva arvu numbreid: esimene grupp 3, teine grupp 4 ja nii edasi kuni 10 numbrini välja. Igas grupis treeniti 50 tehisnärvivõrku, mille kaalukoefitsientidel arvutatud väikese maailma mõõtmetest võeti aritmeetiline keskmine. Iga tehisnärvivõrgu jaoks valiti numbrid juhuslikult. Statistiliselt võrreldi tulemusi tehisnärvivõrkude vahel, mis olid treenitud klassifitseerima erinevaid numbreid ühe võrra rohkem eelmisest. Statistilise testina kasutati Wilcoxon'i astak-summa testi. Järgnevas tabelis (Tabel 5) on kõigepealt toodud treenitud tehisnärvivõrkude keskmised klassifitseerimistäpsused.

Tabel 5. Erineva arvu numbritega treenitud ühe peidetud kihiga tehisnärvivõrkude keskmised klassifitseerimistäpsused.

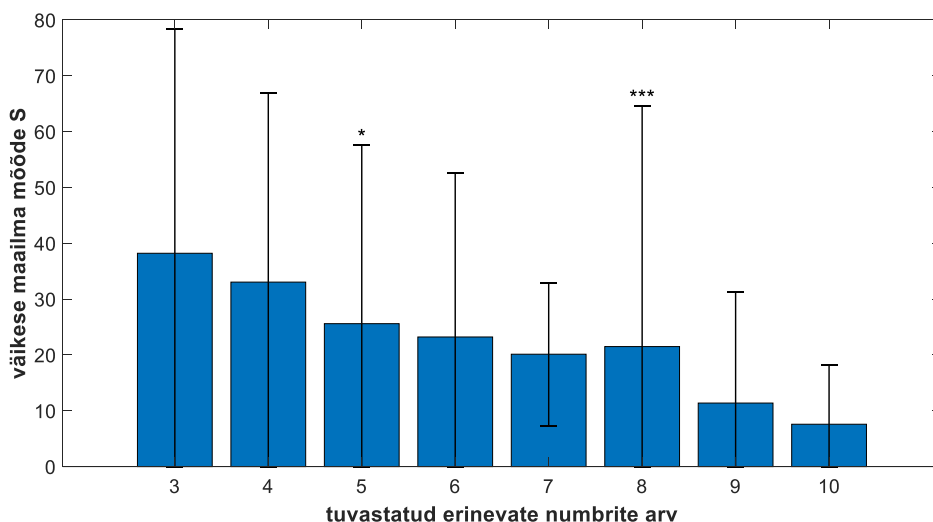
Treeningandmed	Klassifitseerimistäpsus (%)
3 numbrit	99
4 numbrit	98
5 numbrit	97
6 numbrit	97
7 numbrit	96
8 numbrit	95
9 numbrit	94
10 numbrit	93

5.3.1 Väikese maailma mõõde

Väikese maailma mõõde kasvas oluliselt esimese peidetud kihi kaalukoefitsientide korral ($p < 0,05$), mida rohkem erinevaid numbreid tehisnärvivõrk oli treenitud klassifitseerima (Joonis 19). Statistiline olulisus esines mõningatel juhtudel ka väljundkihi kaalukoefitsientide väikese maailma mõõtmete erinevuses (Joonis 20).



Joonis 19. Erineva arvu numbritega treenitud ühe peidetud kihiga tehisnärvivõrkude 1. kihtide kaalukoefitsientidel arvatud keskmised väikese maailma mõõtmed standardhälbega (**** $p < 0,0001$). Statistiliselt on võrreldud tulemusi tehisnärvivõrkude vahel, kus klassifitseerima treenitud erinevate numbrite arv oli ühe võrra väiksem.



Joonis 20. Erineva arvu numbritega treenitud ühe peidetud kihiga tehisnärvivõrkude 2. kihtide kaalukoefitsientidel arvatud keskmised väikese maailma mõõtmed standardhälbega (* $p < 0,05$; *** $p < 0,001$). Statistiliselt on võrreldud tulemusi tehisnärvivõrkude vahel, kus klassifitseerima treenitud erinevate numbrite arv oli ühe võrra väiksem.

5.3.2 Naabusmaatriksite tihedus

Korrelatsioonikoefitsiendi absoluutväärtuse tase, millest kõrgemad väärtused olid statistiliselt olulised ($p < 0,05$) ja seega teisendati ühtedeks, oli esimese kihi korral 0,36. Teise kihi korral olenes tase väljundkihi tehisneuronite arvust. Allpool tabelis (Tabel 6) on toodud binaarsetes naabusmaatriksites esinenud ühtede keskmine esinemissagedus.

Tabel 6. Erineva arvu numbritega treenitud tehismärvivõrkude kihtide keskmised naabusmaatriksite tihedused. $|r| > \dots$ näitab lävendit, millest suurema väärtusega korrelatsioonikordajate absoluutväärtused teisendati ühtedeks ja madalama väärtusega nullideks.

Treeningandmed	Maatriksi tihedus (%)	
	1. kiht ($ r > 0,36$)	2. kiht ($ r > \dots$)
3 numbrit	15	5 (0,997)
4 numbrit	14	5 (0,95)
5 numbrit	13	5 (0,88)
6 numbrit	13	4 (0,81)
7 numbrit	12	4 (0,75)
8 numbrit	12	4 (0,71)
9 numbrit	11	4 (0,67)
10 numbrit	11	4 (0,63)

5.4 Erineva arvu numbrite klassifitseerimine kasutades kahe peidetud kihiga tehismärvivõrkusid

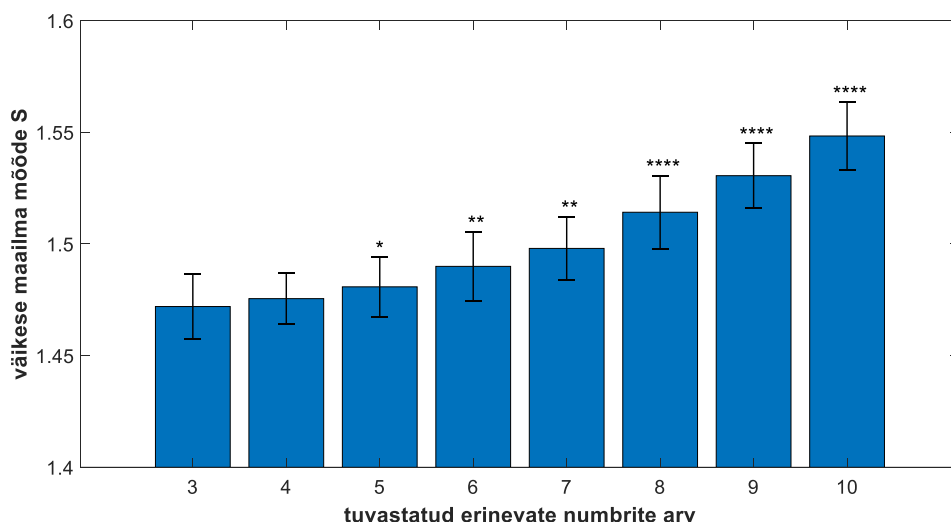
Järjekordselt treeniti igat tüüpi treeningandmetega 50 tehismärvivõrku, mille kaalukoefitsientidel arvatud väikese maailma mõõtmetest on võetud aritmeetiline keskmine. Numbrid, mida iga võrk pidi klassifitseerima, valiti juhuslikult. Statistiliselt on võrreldud tulemusi tehismärvivõrkude vahel, mis olid treenitud klassifitseerima erinevaid numbreid kahe võrra rohkem eelmisest. Statistilise testina kasutati Wilcoxon'i astak-summa testi. Järgnevas tabelis (Tabel 7) on kõigepealt toodud treenitud tehismärvivõrkude keskmised klassifitseerimistäpsused.

Tabel 7. Erineva arvu numbritega treenitud ühe peidetud kihiga tehismärvivõrkude keskmised klassifitseerimistäpsused.

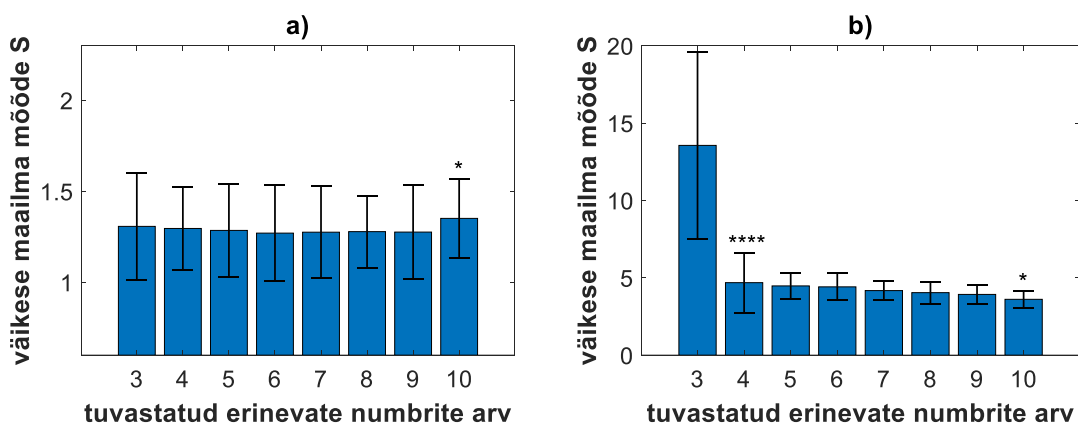
Treeningandmed	Klassifitseerimistäpsus (%)
3 numbrit	99
4 numbrit	99
5 numbrit	99
6 numbrit	98
7 numbrit	98
8 numbrit	98
9 numbrit	97
10 numbrit	97

5.4.1 Väikese maailma mõõde

Väikese maailma mõõde kasvas oluliselt esimese peidetud kihi kaalukoefitsientide korral ($p < 0,05$), mida rohkem erinevaid numbreid tehismärvivõrk oli treenitud klassifitseerima (Joonis 21). Teise peidetud kihi (Joonis 22a) ja väljundkihi (Joonis 22b) kaalukoefitsientide korral väikese maailma mõõtmete erinevuses statistiline olulisus peaaegu puudus ($p > 0,05$).



Joonis 21. Erineva arvu numbritega treenitud kahe peidetud kihiga tehismärgvõrkude 1. kihtide kaalukoefitsientidel arvatud keskmised väikese maailma mõõtmed standardhälbega (* $p < 0,05$; ** $p < 0,01$; **** $p < 0,0001$). Statistiliselt on võrreldud tulemusi tehismärgvõrkude vahel, kus klassifitseerima treenitud erinevate numbrite arv oli ühe võrra väiksem.



Joonis 22. Erineva arvu numbritega treenitud kahe peidetud kihiga tehismärgvõrkude a) 2. kihtide ja b) 3. kihtide kaalukoefitsientidel arvatud keskmised väikese maailma mõõtmed standardhälbega (* $p < 0,05$; **** $p < 0,0001$). Statistiliselt on võrreldud tulemusi tehismärgvõrkude vahel, kus klassifitseerima treenitud erinevate numbrite arv oli ühe võrra väiksem.

5.4.2 Naabusmaatriksite tihedus

Järgnevas tabelis (Tabel 8) on toodud binaarsetes naabusmaatriksites esinenud ühete keskmine esinemissagedus.

Tabel 8. Erineva arvu numbritega treenitud tehiskärvivõrkude kihtide keskmised naabusmaatriksite tihedused. $|r| > \dots$ näitab lävendit, millest suurema väärtusega korrelatsioonikordajate absoluutväärtused teisendati ühtedeks ja madalama väärtusega nullideks.

Treeningandmed	Maatriksi tihedus (%)		
	1. kiht ($ r > 0,20$)	2. kiht ($ r > 0,22$)	3. kiht ($ r > \dots$)
3 numbrit	5	5	14 (0,997)
4 numbrit	5	5	6 (0,95)
5 numbrit	5	5	6 (0,88)
6 numbrit	5	5	6 (0,81)
7 numbrit	5	5	6 (0,75)
8 numbrit	5	5	6 (0,71)
9 numbrit	5	5	5 (0,67)
10 numbrit	5	5	5 (0,63)

5.5 Röntgenülesvõtete klassifitseerimine konvolutsiooniliste võrkudega

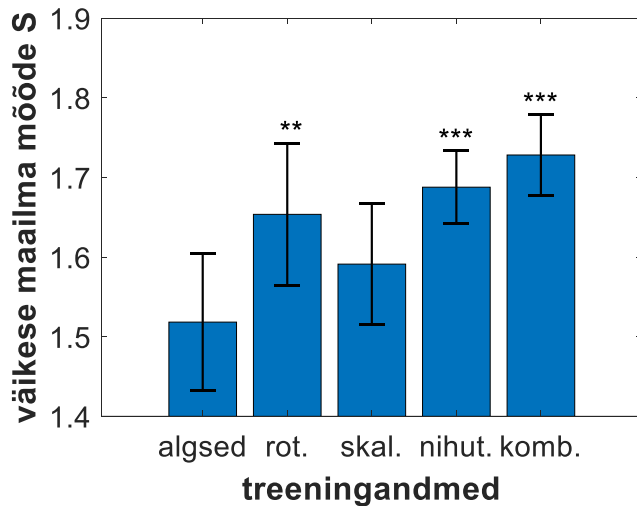
Konvolutsioonilisi võrkusid treeniti treeningandmetega neid roteerides, skaleerides, nihutades või eelnevaid võtteid kombineerides. Igat manipuleerimist korrati 10 korda ja saadi siis ka 10 konvolutsioonilist võrku igat tüüpi treeningandmetega. Kõikide konvolutsiooniliste võrkude korral võeti viimase ahenduskihi ja täielikult ühendatud kihi vahelised kaalukoefitsiendid. Nende põhjal leiti korrelatsioonikordajad, mis teisendati binaarseteks naabusmaatriksiteks. Saadud naabusmaatriksitest leiti väikese maailma mõõtmel igale kihile, millest võeti aritmeetiline keskmine. Manipuleeritud treeningandmetega saadud tulemusi on statistiliselt võrreldud algsete treeningandmetega treenitud konvolutsiooniliste võrkude kaalukoefitsientidel arvutatud väikese maailma mõõtmel. Statistilise testina kasutati Wilcoxon'i astak-summa testi. Järgnevas tabelis (Tabel 9) on kõigepealt toodud treenitud konvolutsiooniliste võrkude keskmised klassifitseerimistäpsused.

Tabel 9. Konvolutsiooniliste võrkude keskmised klassifitseerimistäpsused

Treeningandmed	Klassifitseerimistäpsus (%)
algset	94
roteeritud	92
skaleeritud	94
nihutatud	94
kombineeritud	91

5.5.1 Väikese maailma mõõde

Väikese maailma mõõtmel (Joonis 23) erinesid mõningate manipuleeritud treeningandmetega treenides oluliselt võrreldes algsete treeningandmetega täielikult ühendatud kihi kaalukoefitsientide korral ($p < 0,05$).



Joonis 23. Manipuleeritud andmetega treenitud konvolutsiooniliste võrkude täielikult ühendatud kihi kaalukoefitsientidel arvatud keskmised väikese maailma mõõtmised standardhälbega (** $p < 0,01$; *** $p < 0,001$). Manipuleeritud treeningandmetega saadud tulemusi on statistiliselt võrreldud algsete treeningandmetega treenitud konvolutsiooniliste võrkude tulemustega (rot. – roteeritud; skal. – skaleeritud; nihut. – nihutatud; komb. – kombineeritud).

5.5.2 Naabusmaatriksite tihedus

Korrelatsioonikoefitsiendi absoluutväärtuse tase, millest kõrgemad väärtused olid statistiliselt olulised ($p < 0,05$) ja seega teisendati ühtedeks, oli esimese täielikult ühendatud kihi korral 0,25. Allpool tabelis (Tabel 10) on toodud binaarsetes naabusmaatriksites esinenud ühede keskmine esinemissagedus.

Tabel 10. Konvolutsiooniliste võrkude täielikult ühendatud kihtide keskmised naabusmaatriksite tihedused. $|r| > \dots$ näitab lävendit, millest suurema väärtusega korrelatsioonikordajate absoluutväärtused teisendati ühtedeks ja madalama väärtusega nullideks.

Treeningandmed	Maatriksi tihedus, % ($ r > 0,25$)
algsed	6
roteeritud	6
skaleeritud	6
nihutatud	5
kombineeritud	5

6 Arutelu

Käesolevas töös oli püstitatud hüpotees: mida keerulisemaid andmeid on tehisnärvivõrk treenitud klassifitseerima, seda suurem on tema keskmine väikese maailma mõõde. Manipuleeritud numbritega treenitud tehisnärvivõrkudel vastas püstitatud hüpotees tõele osade esimese kihi kaalukoefitsientidel arvatud tulemuste korral. Nii ühe (Joonis 17) kui ka kahe (Joonis 18) peidetud kihiga tehisnärvivõrkudel täheldati statistiliselt olulist väikese maailma mõõtmete suurenemist esimese kihi korral, kui treeningandmeid oli nihutatud. Treeningandmeid nihutati 4. ja 5. komplekti tehisnärvivõrkude treenimisel.

Erineva arvu numbreid klassifitseerima treenitud tehisnärvivõrkudel vastas püstitatud hüpotees samuti tõele esimese kihi kaalukoefitsientidel arvatud tulemuste korral. Nii ühe (Joonis 19) kui ka kahe (Joonis 20) peidetud kihiga tehisnärvivõrkudel täheldati statistiliselt olulist väikese maailma mõõtmete suurenemist pea kõikidel juhtudel, kui treeningandmetes suurendati erinevat numbrite hulka ühe võrra. Kuna treeningandmed ehk erinevad numbrid valiti iga kord juhuslikult, välistati sellega võimalus, et muutus olenes sellest, milliseid numbreid tehisnärvivõrk klassifitseerima oli treenitud. Kuigi absoluutarvudes (just kahe peidetud kihiga tehisnärvivõrkude korral) on muutused väikesed, on väikese maailma mõõtmed siiski statistiliselt oluliselt suurenenud.

Hüpotees vastas tõele ka röntgenpilte klassifitseerima treenitud konvolutsiooniliste võrkude esimese täielikult ühendatud kihi kaalukoefitsientidel arvatud tulemuste korral (Joonis 23). Sarnaselt manipuleeritud numbritega treenitud tehisnärvivõrkudele täheldati statistilist olulist väikese maailma mõõtmete suurenemist, kui treeningandmete manipuleerimine hõlmas nende nihutamist (4. ja 5. variant).

Kui kaalukoefitsiendid omavahel korreleeruvad, näitab see, et nende väärtused muutuvad koos. Seega sisendväärtusi, mida esimeses peidetud kihis vaadatakse, võimendatakse või summutatakse ühtemoodi (või vastupidi) mööda tehisneuroneid liikudes ja nende roll järgmise peidetud kihi aktivatsiooni väärtuse kujunemisel on sarnase seaduspärasusega jaotatud. Kui numbreid on juhuslikult pildil nihutatud, on piirkonnad, millele tehisnärvivõrgud numbrite klassifitseerimisel peavad keskenduma, pildil rohkem hajutatud. Kuna tehisneuroneid on sama palju, nõuab see tehisnärvivõrkudelt paremat optimeeritust sisendväärtuste töötlemisel. Väikese maailma mõõtmete suurenemist nendel juhtudel võib seletada asjaoluga, et sisendväärtuste niinimetatud

vastutuspiirkonnad jagatakse peidetud kihi (st. täielikult ühendatud kihi) tehisneuronite vahel rohkem gruppidesse.

Kui MNIST andmebaasi treeningandmeid roteeriti või skaleeriti, erinesid väikese maailma mõõtmel samuti oluliselt võrreldes algsega, kuid ühe peidetud kihiga tehisnärvivõrkude korral olid mõõtmel algsest väiksemad (Joonis 17), kahe peidetud kihiga tehisnärvivõrkude korral algsest suuremad (Joonis 18). Seetõttu on raske nende andmete põhjal järeldusi teha. Mõlemal puhul oli mõõtmel muutus suhteliselt väike võrreldes algsega. Võimalik, et roteeritud ja skaleeritud treeningandmed pole võrreldes algsete andmetega piisavalt keerulised olnud ning tulemused on pigem juhuslikud.

Kõigi ülejäänud täielikult ühendatud kihtide kaalukoefitsientidel arvutatud väikese maailma mõõtmel muutusid erinevate treeningandmetega treenides pealtnäha juhuslikult. Seda ilmestas lisaks statistilise olulisuse puudumisele ka suur standardhälve. Pealtnäha juhuslikud väikese maailma mõõtmel võivad näidata, et tehisnärvivõrkude struktuur on siiski liialt kompleksne, et leida lihtne kvantifitseeritav mõõde sellisel meetodil, mis võimaldaks kirjeldada täielikult niinimetatud „mustas kastis“ toimuvat. Võimalik, et esimeste kihtide korral muutuvad väikese maailma mõõtmel süsteemselt, sest sisendkihiks on piltide pikselväärtused. Järgmistes kihtides aga kaob samasugune pikslite loogika ja sellepärast ei ole ka sarnaseid tulemusi. Tekib küsimus, kas saaksime esimeste kihtide korral ikka samad tulemused, kui kasutaksime teistsuguseid andmeid piltide asemel.

Erineva arvu numbrite klassifitseerimiseks kasutatud tehisnärvivõrkude korral tuleb mees pidada, et väljundkihi tehisneuronite arv sõltus, mitut erinevat numbrit tehisnärvivõrk klassifitseerima oli treenitud. See tähendab, et katsepunktide arv naabusmaatriksite koostamisel oli erinev. Samas vaadates naabusmaatriksite tihedusi, jäävad need kõigil juhtudel samasse suurusjärku. Seega võiks spekuloida, et antud juhul ei ole pealtnäha juhuslikkuse probleem tingitud väljundkihi tehisneuronite arvust.

Võis näha trendi, et klassifitseerimistäpsus langes, kui treeningandmed muutusid keerulisemaks. Drastiline erinevus klassifitseerimistäpsustes oli märgatavalt väiksem suuremate tehisnärvivõrkude korral. Siit tuleneb ka loogiline järeldus, et keerukamate probleemide lahendamiseks on tarvis komplekssemaid tehisnärvivõrkusid. Kui võrrelda ühe ja kahe peidetud kihiga tehisnärvivõrgu klassifitseerimistäpsusi MNIST andmebaasi

numbrite manipuleerimise korral (Tabel 1, Tabel 3), on näha, et kahe peidetud kihiga tehiskärgvõrgu klassifitseerimistäpsused on oluliselt paremad. Erineva arvu numbrite klassifitseerimiseks kasutatud tehiskärgvõrkude üldistusvõime on mõlemal juhul päris hea (Tabel 5, Tabel 7). See on ka loogiline, kuna probleem, mida lahendati oli küllaltki lihtne. Siiski täheldati ka seekord klassifitseerimistäpsuse langemist, kui tehiskärgvõrgule anti treeningandmete näol raskem ülesanne.

Konvolutsiooniliste võrkude korral võis näha, et teatud juhtudel, kui treeningandmeid keerukamaks muudeti, jäi klassifitseerimistäpsus samaks (Tabel 10). Nihutatud treeningandmete korral aga suurenes oluliselt ka väikese maailma mõõde. Klassifitseerimistäpsused olid siiski arvatud erinevalt. MNIST andmebaasi numbrite korral kasutati klassifitseerimistäpsuse arvutamiseks treeningandmetega koos manipuleeritud valideerimisandmeid, kuid röntgenülesvõtete korral kasutatud testandmeid ei manipuleeritud.

Klassifitseerimistäpsuse languse trend esines suuremal või vähemal määral kõigi käesolevas töös kasutatud tehiskärgvõrkude ja konvolutsiooniliste võrkude korral. Eriti ilmekalt tuli see esile ühe peidetud kihiga tehiskärgvõrkude juures. See on kohati ilmne, sest me ei saa oodata, et samasuguste parameetritega tehiskärgvõrk suudab keerulisemate probleemidega sama hästi või paremini hakkama saada. Samas kui mõelda intuiitselt, kumb tehiskärgvõrk on parema üldistusvõimega, võiks see olla tehiskärgvõrk, mis on näinud keerulisemaid või rohkem erinevatesse klassidesse kuuluvaid andmeid.

Naabusmaatriksite tihedus (Tabel 2, Tabel 4, Tabel 6, Tabel 8) jäi MNIST andmebaasi numbrite klassifitseerimiseks kasutatud tehiskärgvõrkude esimestel kihtidel suurusjärku vastavalt 10% ja 5%. Naabusmaatriksi tiheduse taoline märgatav vähenemine võib osutada problemaatiliseks, kui analüüsida peidetud kihte, kus on rohkem tehiskärgneuroneid, sest käesolevas tööst nähtub, et statistiliselt oluliste korrelatsioonikordajate suhtarv väheneb katsepunktide arvu suurenedes (esimese tehiskärgvõrgu puhul 30, teise puhul 100). Üheks lähenemiseks võiks olla kasutada fikseeritud lävendi asemel bioloogiliste kärgvõrkude puhul sagedasti kasutatud tiheduse lävendit [49]. Sel juhul tõstatub aga probleem, kus ühtedeks teisendatakse korrelatsioonikordajate absoluutväärtused, mis pole statistiliselt olulised. See aga omakorda võib mõjutada saadud väikese maailma mõõtmete usaldusväärsust.

Konvolutsiooniliste võrkude korral jäid naabrusmaatriksite tihedused 64 katsepunkti juures samuti suurusjärku 5% (Tabel 10). Võimalik, et taoline suurusjärk ongi taoliste andmete korral tavapärane ja just väiksemate kihtide puhul näeme lihtsalt hüppelist suurenemist. Selle tõestuseks on aga vaja rohkem uuringuid. Samuti tekib küsimus, kas me ikka saame konvolutsiooniliste võrkude taolisi näitajaid võrrelda lihtsate tehisnärvivõrkudega.

Käesolevas töös kasutatud konvolutsioonilistel võrkudel saigi väikese maailma mõõtmeid arvutada ainult ühe kihi kaalukoefitsientidel, kuna ainsale täielikult ühendatud peidetud kihile järgnes väljundkiht ühe tehisneuroniga. Konvolutsiooniliste kihtide kaalukoefitsientidel käesolevas töös väikese maailma mõõtmete arvutamist ei arutatud. Põhjus seisnes selles, et konvolutsioonilised kihid on teise põhimõtte järgi üles ehitatud. Konvolutsioonilised kihid on mõeldud pigem tunnusjoonte tuvastamiseks täielikult ühendatud kihtide jaoks, mille ülesanne on lõpuks klassifitseerimine. Teiseks on konvolutsiooniliste kihtide kaalukoefitsiendid määratud teistsuguse põhimõttega ning käesolevas töös kasutatud meetod nõuaks teistsugust lähenemist.

7 Järeldused

Käesolevas töös leiti, et esimese täielikult ühendatud peidetud kihi kaalukoefitsientidel esinevad suuremad väikese maailma mõõtmed, kui tehisnärvivõrkusid oli treenitud keerulisemate andmetega. Esimese meetodiga saadi keerulisemad treeningandmed pildil olevaid objekte nihutades. Teise meetodiga suurendati klasside arvu, mille vahel tehisnärvivõrk treeningandmeid klassifitseerima pidi. Töös kasutati kokku kolme erinevat tehisnärvivõrku ja neil kõigil täheldati väikese maailma mõõtmete kasvu esimese täielikult ühendatud kihi kaalukoefitsientidel.

Tehisnärvivõrgud on inspireeritud bioloogilistest närvivõrkudest ja omavad seega ka teatavaid sarnasusi. Seda tõdemust arvestades julgustavad saadud tulemused kasutama bioloogilistel närvivõrkudel kasutatavaid analüüsimeetodeid ka tehisnärvivõrkudel. Käesolevas töös kasutatud meetodil arvutatud väikese maailma mõõtmeid võib tulevikus kaaluda näiteks lisaabinõuna tehisnärvivõrkude üldistusvõime hindamisel. Kuna sellist meetodit pole autorile teadaolevalt varasemalt kasutatud, nõuab see kindlasti edasist uurimist.

Käesolevas töös tehtud katsed on tehtud väga kontrollitud tingimustes ning uuriti väikese maailma mõõtmete muutust ainult treeningandmete muutmisel. Hüpoteesi paikapidavust tuleks kindlasti uurida rohkem erinevatel tehisnärvivõrkudel. Samuti tuleks meetodit katsetada ka tehisnärvivõrkudel, mida kasutatakse teistsuguste andmete kui piltide tuvastamisel. Lisaks tasuks tulevikus uurida, kuidas mõjutab erinevate parameetrite, nagu aktiveerimisfunktsiooni, õpisammu, regulariseerimismeetodi, epohhide arvu vms muutmine kaalukoefitsientidel arvutatud väikese maailma mõõdet.

8 Kokkuvõte

Käesolevas töös rakendati tehisnärvivõrkude kaalukoefitsientidel komplekssete võrkude analüüsimeetodit, mida üldiselt kasutatakse bioloogiliste närvivõrkude korral. Täpsemalt seati hüpotees: mida keerulisemaid andmeid on tehisnärvivõrk treenitud klassifitseerima, seda suurem on tema kihtide kaalukoefitsientidel arvatud keskmine väikese maailma mõõde.

Käesolevas töös püstitatud hüpotees osutus tõseks väikese maailma mõõtmetel, mis olid arvatud tehisnärvivõrkude esimese täielikult ühendatud kihi kaalukoefitsientidel. Sealjuures olid vastuolulised tulemused treeningandmetega, mida oli roteeritud või skaleeritud. Ühe peidetud kihiga tehisnärvivõrkude korral väikese maailma mõõtmed nendel puhkudel vähenesid, kahe peidetud kihiga tehisnärvivõrkude korral aga suurenesid. Samas statistiline olulisus esines mõlemal korral. Konvolutsiooniliste võrkudel skaleeritud treeningandmete korral jällegi statistilist olulisust ei esinenud. Nende vastuolude pealt võib väita, et roteeritud ja skaleeritud treeningandmed ei erinenud märgavalt võrreldes algsete treeningandmetega.

Erineva arvu numbreid klassifitseerima treenitud tehisnärvivõrkude korral esines samuti positiivne trend väikese maailma mõõtmete suurenemises esimese kihi kaalukoefitsientide korral. Hüpotees pidas paika enamiku juhtude korral, kui võrreldi omavahel tehisnärvivõrkusid, mille klassifitseerima treenitud numbrite arv erines ühe võrra. Ainsad korrad, kui ei täheldatud statistiliselt olulist suurenemist, oli ühe peidetud kihiga tehisnärvivõrkude korral 8 ja 9 numbri vahel ning kahe peidetud kihiga tehisnärvivõrgu korral 3 ja 4 numbri vahel.

Ülejäänud kihtide kaalukoefitsientidel arvatud väikese maailma mõõtmete muutuses puudus süstemaatiline statistiline olulisus. Pealtnäha juhuslikud väärtused esinesid mõlema meetodiga saadud treeningandmete ning kõigi töös kasutatud tehisnärvivõrkude korral. Sellest tekib kahtlus, et väikese maailma mõõtmed suurenesid statistiliselt oluliselt esimese kihi kaalukoefitsientide korral, sest sisendkihiks olid piltide pikselväärtused. Võimalik, et teistsuguste andmete korral ei pruugi ka esimeses kihis enam süstemaatiline statistiline olulisus esineda.

Käesolevas töös kasutatud meetod võiks tulevikus rakendust leida näiteks tehisnärvivõrkude üldistusvõime hindamisel. Kuna katsed viidi läbi väga kontrollitud

tingimustes, on mõistlik tulevikus antud teemat edasi uurida. Tuleks katsetada rohkem nii erinevaid tehisnärvivõrkusid kui ka teist tüüpi andmeid. Samuti tuleks vaadata, kuidas muutuvad väikese maailma mõõtmed, kui muuta tehisnärvivõrgu parameetreid.

Kasutatud kirjandus

- [1] V. Renganathan, „Overview of artificial neural network models in the biomedical domain,“ *Bratislava Medical Journal*, kd. 120, nr 7, pp. 536-540, 2019.
- [2] E. Petlenkov, *Tehisnärivõrgud ja nende rakendused*. Õppematerjal, Tallinn: TTÜ, 2004.
- [3] F. Fan, J. Xiong ja G. Wang, „On Interpretability of Artificial Neural Networks,“ arXiv preprint, 2020.
- [4] M. A. Nielsen, *Neural Networks and Deep Learning*, San Fransisco: Determination Press, 2015.
- [5] Y. LeCun, *AMA: Yann LeCun*. [Intervjuu]. 15 mai 2014.
- [6] A. Holzinger, C. Biemann, C. S. Pattichis ja D. B. Kell, „What do we need to build explainable AI systems for the medical domain?,“ arXiv preprint, 2017.
- [7] A. Holzinger, G. Langs, H. Denk, K. Zatloukal ja H. Müller, „Causability and explainability of artificial intelligence in medicine,“ *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, kd. 9, nr 1312, pp. 1-13, 2019.
- [8] M. Rubinov ja O. Sporns, „Complex network measures of brain connectivity: Uses and interpretations,“ *Neuroimage*, kd. 52, nr 3, pp. 1059-1069, 2010.
- [9] L. da Fontoura Costa, „What is a Complex Network?,“ Sao Carlos Institute of Physics, Sao Carlos, 2018.
- [10] M. van Steen, „Graph Theory and Complex Networks. An Introduction,“ Maarten van Steen, 2010.
- [11] S. Saha, „A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way,“ *Towards Data Science*, 15 12 2018. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>. [Kasutatud 14 05 2020].
- [12] P. Kamencay, M. Benco, T. Mizdos ja R. Radil, „Advances in Electrical and Electronic Engineering,“ *Digital Image Processing and Computer Graphics*, kd. 15, nr 4, pp. 663-672, 2017.
- [13] I. Goodfellow, Y. Bengio ja A. Courville, „Deep Learning,“ MIT Press, Cambridge, 2016.
- [14] R. Palm, *Diskreetse matemaatika elemendid*, Tartu: TÜ Arvutiteaduste Instituut, 2009.
- [15] G. Montavon, W. Samek ja K.-R. Müller, „Methods for Interpreting and Understanding Deep Neural Networks,“ arXiv preprint, 2017.
- [16] J. J. Montano ja A. Palmer, „Numeric sensitivity analysis applied to feedforward neural networks,“ *Neural Computing & Applications*, kd. 12, pp. 119-125, 2002.
- [17] A. Hunter, L. Kennedy, J. Henry ja I. Ferguson, „Application of neural networks and sensitivity analysis to improved prediction of trauma survival,“ *Computer Methods and Programs in Biomedicine*, kd. 62, nr 1, pp. 11-19, 2000.

- [18] J. Khan, J. S. Wei ja P. S. Meltzer, „Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks,“ *Nature Medicine*, kd. 7, nr 6, pp. 673-679, 2001.
- [19] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Müller ja W. Samek, „On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation,“ *PLOS ONE*, kd. 10, nr 7, pp. 1-46, 2015.
- [20] W. Samek, A. Binder, G. Montavon, S. Bach ja K.-R. Müller, „Evaluating the visualization of what a Deep Neural Network has learned,“ arXiv preprint, 2015.
- [21] E. Tjoa ja C. Guan, „A Survey on Explainable Artificial Intelligence (XAI): Towards Medical XAI,“ arXiv preprint, 2019.
- [22] G. Zhao, B. Zhou, K. Wang, R. Jiang ja M. Xu, „Respond-CAM: Analyzing Deep Models for 3D Imaging Data by Visualizations,“ arXiv preprint, 2018.
- [23] M. Izadyazdanabadi, E. Belykh ja Y. Yang, „Weakly-Supervised Learning-Based Feature Localization for Confocal Laser Endomicroscopy Glioma Images,“ *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Lecture Notes in Computer Science*, kd. 11071, 2018.
- [24] Z. Tang, K. V. Chuang ja C. DeCarli, „Interpretable classification of Alzheimer’s disease pathologies with a convolutional neural network pipeline,“ *Nature Communications*, kd. 10, nr 2173, 2019.
- [25] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva ja A. Torralba, „Learning Deep Features for Discriminative Localization,“ arXiv preprint, 2015.
- [26] J. D. Olden ja D. A. Jackson, „Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks,“ *Ecological Modelling*, kd. 154, pp. 135-150, 2002.
- [27] J. D. Olden, M. K. Joy ja R. G. Death, „An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data,“ *Ecological Modelling*, kd. 178, pp. 389-397, 2004.
- [28] Z. Zhang, M. W. Beck, D. A. Winkler, B. Huang, W. Sibanda ja H. Goyal, „Opening the black box of neural networks: methods for interpreting neural network models in clinical applications,“ *Annals of Translational Medicine*, kd. 6, nr 216, pp. 1-11, 2018.
- [29] D. J. Watts ja S. H. Strogatz, „Collective dynamics of ‘small-world’ networks,“ *Nature*, kd. 393, nr 6684, pp. 440-2, 1998.
- [30] J. Wohlgemuth ja M. T. Matache, „Small-World Properties of Facebook Group Networks,“ *Complex Systems*, kd. 23, nr 3, pp. 197-225, 2014.
- [31] E. Ch’ng, „Local Interactions and the Emergence of a Twitter Small-World Network,“ *Social Networking*, kd. 4, nr 2, pp. 33-40, 2015.
- [32] C. C. Hilgetag ja M. Kaiser, „Clustered Organization of Cortical Connectivity,“ *Neuroinformatics*, kd. 2, nr 3, pp. 353-360, 2004.
- [33] D. S. Bassett, A. Meyer-Lindenberg, S. Achard, T. Duke ja E. Bullmore, „Adaptive reconfiguration of fractal small-world human brain functional networks,“ *Proceedings of the National Academy of Sciences of the United States of America*, kd. 103, nr 51, pp. 19518-19523, 2006.
- [34] D. S. Bassett ja E. T. Bullmore, „Small-World Brain Networks Revisited,“ *The Neuroscientist*, kd. 23, nr 5, pp. 499-516, 2017.

- [35] D. S. Bassett ja E. T. Bullmore, „Small-World Brain Networks,“ *The Neuroscientist*, kd. 12, nr 6, pp. 512-523, 2007.
- [36] Z. P. Neal, „How small is it? Comparing indices of small worldliness,“ *Network Science*, kd. 5, nr 1, pp. 30-44, 2017.
- [37] M. D. Humphries ja K. Gurney, „Network ‘Small-World-Ness’: A Quantitative Method for Determining Canonical Network Equivalence,“ *PLOS ONE*, kd. 3, nr 4, pp. 1-10, 2008.
- [38] O. Erkaymaz, M. Ozer ja M. Perc, „Performance of small-world feedforward neural networks for the diagnosis of diabetes,“ *Applied Mathematics and Computation*, kd. 311, pp. 22-28, 2017.
- [39] M. Javaheripi, B. D. Rouhani ja F. Koushanfar, „SWNet: Small-WorldNeuralNetworksandRapidConvergence,“ arXiv preprint, 2019.
- [40] A. M. Zador, „A critique of pure learning and what artificial neural networks can learn from animal brains,“ *Nature Communications*, kd. 10, nr 3770, pp. 1-7, 2019.
- [41] Y. LeCun, C. Cortes ja C. J. Burges, „The MNIST database of handwritten digits,“ [Võrgumaterjal]. <http://yann.lecun.com/exdb/mnist/>.
- [42] D. Kermany, K. Zheng ja M. Goldbaum, „Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification, v2,“ Mendeley Data, 06 01 2018. [Võrgumaterjal]. <https://data.mendeley.com/datasets/rsbjbr9sj/2>. [Kasutatud 12 05 2020].
- [43] J. Langelaar, *MNIST neural network training and testing*, MATLAB Central File Exchange, 2019.
- [44] X. Bian, „Pneumonia Detection Using Deep Learning, Version 20.,“ 08 05 2020. [Võrgumaterjal]. <https://www.kaggle.com/therealcyberlord/pneumonia-detection-using-deep-learning/data?fbclid=IwAR0keRYcQYWfqY3E0Xm7TPRNPROS-X8OA07aCWO2qXAPmuQ2r6bqJa6o1nw>. [Kasutatud 12 05 2020].
- [45] Kaggle, „kaggle.com,“ 2019. [Võrgumaterjal]. <https://www.kaggle.com/>. [Kasutatud 13 05 2020].
- [46] Mathworks, „MATLAB,“ Mathworks, 2020. [Võrgumaterjal]. https://se.mathworks.com/products/matlab.html?s_tid=hp_ff_p_matlab.
- [47] M. Rubinov ja O. Sporns, „Brain Connectivity Toolbox,“ 2010. [Võrgumaterjal]. brain-connectivity-toolbox.net.
- [48] B. C. M. van Wijk, C. J. Stam ja A. Daffertshofer, „Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory,“ *PLOS ONE*, kd. 5, nr 10, pp. 1-13, 2010.
- [49] C. Garcia Forlim, S. Haghiri, S. Düzel ja S. Kühn, „Efficient small-world and scale-free functional brain networks at rest using k-nearest,“ bioRxiv preprint, 2019.
- [50] S. Maslov ja K. Sneppen, „MATLAB programs. Degree-preserving random rewiring and correlation profile of a complex network,“ 2002. [Võrgumaterjal]. <http://maslov.bioengineering.illinois.edu/matlab.htm>.