

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Martin Joonas Pariis 204331IAPM

**Kujutiste masinõppele tugineva
automaatanalüüsi juurutamine radioloogide
töövoogu**

Magistritöö

Juhendaja: Juhan-Peep Ernits
PhD

Kaasjuhendaja: Julius Juurmaa
MD

Tallinn 2023

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: Martin Joonas Pariis

08.01.2023

Annotatsioon

Käesoleva töö eesmärk on pakkuda välja protsessid täiendavaks automaatanalüüside tegemiseks radioloogilistel kujutistel, kasutades selleks masinõppe tehnoloogiaid. Protsesside eesmärk on lihtsustada radioloogide tööd ja pakkuda patsientidele paremat teenust. Töö käigus kavandatud protsessid realiseeriti ja prooviti osaliselt läbi prototüüpide näol. Protsesside ja prototüübi asjalikkust hindas käesolevasse töösse kaasatud kaasjuhendaja.

Prototüüp realiseeriti peamiselt uudset MONAI tarkvara kasutades. Kõigepealt prooviti rakendada tarkvara töövoogudel eeltreenitud mudeleid. Lisaks prooviti treenimise töövoos abil ise treenida suurtel andmekogudel põhinevad mudelid, mis seejärel saadeti järeldamise töövoole kasutamiseks. Viimaks katsetati mudelite kasutatavust varasemate andmekogude näidiskujutistel ning ka Eestis tehtud radioloogilistel uuringutel.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 46 leheküljel, 7 peatükki, 27 joonist, 5 tabelit.

Abstract

Integration of AI Based Automatic Analysis of Images into the Workflow of Radiologists

The aim of this thesis is to propose processes for additional automatic analysis on radiological images using machine learning technologies. The aim of the processes is to simplify the work of radiologists and provide better service to patients. The proposed processes were partially realized and tested in the form of prototypes. The practicality of the processes and the prototype was evaluated by the co-supervisor involved in this thesis.

The prototype was mainly realized using the new MONAI software. Firstly, pre-trained models were applied to software workflows. In addition, it was attempted to self-train models based on large datasets using the training workflow, which were then sent to the inference workflow for use. Finally, the usability of the models was tested on sample images from earlier datasets and on radiological studies carried out in Estonia.

The thesis is in Estonian and contains 46 pages of text, 7 chapters, 27 figures, 5 tables.

Lühendite ja mõistete sõnastik

CNN	<i>Convolutional neural network</i> – konvolutsiooniline närvivõrk
CVAT	<i>Computer Vision Annotation Tool</i> , interaktiivne video- ja pilditöötlemise tööriist tehisenägemises
DICOM	<i>Digital Imaging and Communications in Medicine</i> , meditsiinilise andmevahetuse standard
GDPR	<i>The General Data Protection Regulation</i> , isikuandmete kaitse üldmäärus (IKÜM)
GPU	<i>Graphics Processing Unit</i> , graafikaprotessor
Katastroofiline unustamine	<i>Catastrophic forgetting</i> – tehislike närvivõrkude kalduvus unustada vana informatsiooni uue omandamisel
Otsinguga rahulolu	<i>satisfaction of search</i> , olukord kus kujutiselt ühe abnormaalsuse leidmine takistab teise leidmist
PACS	<i>Picture archiving and communication systems</i> , piltide arhiveerimise ja sidesüsteemid. Tervishoiu asutustes kasutatav tehnoloogia meditsiiniliste piltide ja aruannete hoiustamiseks ning edastamiseks
REST	<i>REpresentational State Transfer</i> , esinduslik olekuedastamine

Sisukord

1 Sissejuhatus	11
2 Radioloogilistele uuringutele täiendava automaatanalüüsi tegemise protsess	13
2.1 Üldine radioloogilise uuringu protsess	13
2.1.1 Õige mudeli töövoog käivitamine	16
2.2 Andmete kasutamine	17
2.2.1 Andmete hoiustamine	17
2.3 Mudeli treenimine.....	19
2.4 Mudeli kasutamine järeldamisel.....	20
2.4.1 Tagasiside andmine	21
3 Realiseeritud prototüüp	22
3.1 MONAI.....	22
3.2 Andmete hoiustamine	22
3.2.1 Orthanc	23
3.3 Piltide märgendamine	24
3.3.1 MONAI Label	24
3.3.2 Näidisrakendused MONAI Labelis	25
3.4 Mudelite treenimine MONAI Core'iga	27
3.5 Mudelite kasutamine järeldamiseks ja visualiseerimiseks MONAI Deploy abil ..	28
4 Eksperimendid prototüübil	30
4.1 Järelduste tegemine MONAI Deploy Expressi näitel.....	30
4.2 MONAI eeltreenitud mudelid.....	31
4.2.1 Eesnäärme segmenteerimine – näide ühe mudeli põhjal.....	31
4.2.2 Kõhunäärme ja maksa segmenteerimine - Mitme mudeli rakendamine	33
4.3 Mudeli treenimise ja järeldamise töövoogude prototüüp MedMNIST kopsupõletiku andmekogu näitel	34
4.3.1 Mudeli treenimine MedMNIST PneumoniaMNIST andmekogu põhjal.....	34
4.3.2 Pakendamine rakenduseks.....	35
4.3.3 Järeldamine testkujutistel	36

4.4 Mudeli treenimise ja järeldamise töövoogude prototüüp MedMNIST mitme kopsuhaiguse andmekogu näitel	36
4.4.1 Mudeli treenimine MedMNIST ChestMNIST andmekogu põhjal	36
4.4.2 Pakendamine rakenduseks	38
4.4.3 Järeldamine testkujutistel	38
4.5 Mudeli treenimise ja järeldamise töövoogude prototüüp ChestXRay2017 andmekogu näitel.....	40
4.5.1 Mudeli treenimine ChestXRay2017 andmekogu põhjal	40
4.5.2 Pakendamine rakenduseks	41
4.6 Kopsupõletiku ja -haiguste klassifitseerimise mudelite testimine Eestis tehtud röntgenite peal	41
4.6.1 Järeldamised ChestXRay2017 andmekogul põhineval mudelil enne ja pärast ekslike hinnangute parandamist	42
5 Alternatiivne prototüüp	44
5.1 Töövoog.....	44
5.2 Microsoft DICOM Server andmete hoiustamiseks	45
5.3 CVAT piltide märgendamiseks	46
5.4 nnU-Net mudelite treenimiseks ja järeldeste tegemiseks.....	48
6 Arutelu.....	49
6.1 Automaatanalüüsi tegemise protsess	49
6.2 Realiseeritud prototüüp.....	50
6.2.1 MONAI.....	50
6.2.2 Mudelite sooritusvõime	51
6.3 Edasised arendused.....	51
6.3.1 Potentsiaalsed probleemid mudelite kasutamisel	52
6.3.2 Seletatav mudel	52
6.3.3 Homomorfne krüpteerimine	53
7 Kokkuvõte	55
Kasutatud kirjandus	57
Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks	60
Lisa 2 – Eesnäärme segmenteerimise rakenduse failid	61
Lisa 3 – PneumoniaMNIST kopsupõletiku mudeli treenimiseks ja kasutamiseks loodud failid.....	62

Lisa 4 – ChestMNIST mitme kopsuhaiguse klassifitseerimise mudeli treenimiseks ja kasutamiseks loodud failid	63
Lisa 5 – ChestXRay2017 kopsupõletiku mudeli treenimiseks ja kasutamiseks loodud failid.....	64
Lisa 6 – DICOM konverteerija png ja NifTI failideks	65

Jooniste loetelu

Joonis 1. Uuringute protsess hetkel	13
Joonis 2. Pakutav protsess lühidalt	14
Joonis 3. Pakutav protsessi mudel	15
Joonis 4. Pakutav protsessi mudel, kus spetsialist on treener.....	16
Joonis 5. Uuringute saatmine märgendamisse.....	18
Joonis 6. Andmete liikumine järeldamise protsessis	19
Joonis 7. Märgendamise protsess	19
Joonis 8. Mudeli kasutamise protsess.....	21
Joonis 9. MONAI Radiology rakendus [18].....	25
Joonis 10. MONAI Pathology rakendus [18]	26
Joonis 11. MONAI Bundle rakendus [18].....	26
Joonis 12. MONAI Endoscopy rakendus [18]	27
Joonis 13. MONAI Core moodulid [20].....	28
Joonis 14. MONAI AI Deployment Workflow [3]	29
Joonis 15. Käsklus MAP seadistuse saatmiseks MONAI töövoog haldajale.....	30
Joonis 16. Eesnäärme segmenteerimise õnnestunud näide	32
Joonis 17. Eesnäärme segmenteerimise ebaõnnestunud näide.....	32
Joonis 18. Kõhunäärme ja maksa samaaegne segmenteerimine MONAI mudeli abil...	33
Joonis 19. Kopsupõletikku klassifitseeriva mudeli treenimise meetrika.....	34
Joonis 20. PneumoniaMNIST kopsupõletiku klassifitseerimise rakenduse väljund.....	36
Joonis 21. ChestMNIST kopsuhaiguseid klassifitseeriva mudeli treenimise meetrika..	37
Joonis 22. Mudeli väljund uue seeriana	39
Joonis 23. Kopsuhaiguste klassifitseerimise väljund	39
Joonis 24. ChestXRay2017 kopsupõletikku klassifitseeriva mudeli treenimise meetrika	40
Joonis 25. Esialgselt realiseeritud märgendamise töövoog	44
Joonis 26. Ülesande loomine CVATis	47
Joonis 27. Märgendamine OpenCVga.....	47

Tabelite loetelu

Tabel 1. PneumoniaMNIST kopsupõletiku klassifitseerimise mudeli täpsuse klassifikatsiooniraport	35
Tabel 2. ChestMNIST kopsuhaiguste klassifitseerimise mudeli täpsuse klassifikatsiooniraport	37
Tabel 3. ChestXRay2017 kopsupõletiku klassifitseerimise mudeli täpsuse klassifikatsiooniraport	41
Tabel 4. MedNMIST andmekogudel treenitud mudelite järelused Eesti röntgenkujutistel	42
Tabel 5. ChestXRay2017 andmekogul treenitud mudelite järelused Eesti röntgenkujutistel	43

1 Sissejuhatus

Digitaliseerimisega on meditsiinis tegeletud aastakümneid ja käesoleva töö ajaks on Eestis patsiendi terviseandmete vaatamiseks kasutuses patsiendiportaal. Lisaks on kõigis haiglates infosüsteemid, mis andmete vahetamiseks liidestuvad näiteks digiregistratuuri¹, digiloo² ja pildipangaga³, kuhu muuhulgas salvestatakse radioloogiliste uuringute tulemusi.

Käesolevas töös vaadeldakse, mis võimalusi pakub kaasaegne tehismõistuse lahendustele tuginev tehnoloogia radioloogiliste kujutiste analüüsiks ja kuidas on selliseid lahendusi võimalik integreerida haigla protsessidesse, et pakkuda patsientidele olemasolevate ressurssidega paremat teenust. Siinkohal tuleb meeles pidada, et terviseandmed kuuluvad eriliiki isikuandmete alla, mille töötlemist reguleeritakse Isikuandmete kaitse üldmäärusega IKÜM (mis on tihti tuntud ka inglise keelse akronüümi GDPR kaudu) [1]. Isikuandmete töötlemisega tegelev osapool peab järgima Isikuandmete töötleja üldjuhendit⁴.

Meditsiiniliste andmete vahetamise ühtlustamiseks arendab organisatsioon HL7 välja standardit FHIR⁵, mis peaks võimaldama terviseandmete vahetamist erinevate asjasse puutuvate osapoolte vahel. DICOM [2] on standard, mille kohased andmed tagastatakse digitaalsetest radioloogilistest aparatuuridest. Kuna suurem osa radioloogilistest andmetest on digitaliseeritud kujul, siis vaatleme käesolevas töös stsenaariume, kus radioloogilise uuringule, mis tehakse konkreetse seisundi täpsustamiseks, näiteks rindkereuuring kopsupõletiku tuvastamiseks, on võimalik sisse lülitada täiendav automaatanalüüs, mis saab vastava uuringu tulemuselt otsida võimalikke jälgi näiteks tuumoritest. Sellised

¹ <https://www.veebiregistratuur.ee/all>

² <https://www.digilugu.ee>

³ <http://pacs.ee/>

⁴ https://www.aki.ee/sites/default/files/dokumendid/isikuandmete_tootleja_uldjuhend.pdf

⁵ http://www.hl7.org/implement/standards/product_brief.cfm?product_id=491

täiendavad digitaalsed analüüsid annavad radioloogidele võimaluse kasutada täiendavat automaatskriiningut.

Käesoleva töö etapid koos uurimisküsimustega on järgnevad:

1. Näidistööprotsessi väljatöötamine ja prototüübi seadistamine - DICOM serveri ülesseadmine piltide hoiustamiseks ja märgendamiseks. See on eelduseks, et spetsialistid saaksid hakata märgendama pilte, mille põhjal mudeleid valideerida ja treenida.
 - a. Milline peaks olema üldine protsess uuringutelt patoloogiliste leidude tuvastamiseks ja vajadusel mudelile paranduste saatmiseks?
 - b. Milline peaks olema protsess radioloogiliste uuringute jaoks märgendatud andmete kogumiseks, mis vastaks andmekaitsereeglite ja haiglas kasutatavata infosüsteemide nõuetele?
 - c. Kuidas teha märgendamine mugavaks radioloogile?
2. Patoloogilisi leide tuvastavate mudelite valideerimine märgendatud piltide põhjal.
 - a. Kui hästi suudavad varem treenitud mudelid meie meditsiinisisüsteemi röntgenülesvõtetel patoloogilisi leide tuvastada?
 - b. Kas on võimalik parandada mudeli võimekust Eestis tehtud uuringutel, lisades neid treeningandmetesse?
3. Patoloogilisi leide tuvastavate mudelite ületreenimise ja uue valideerimise protsessi paika panemine.
 - a. Missugune peaks olema ületreenimise protsess ja kuidas hoida treenimise andmeid, säilitades kooskõla IKÜMiga?
 - b. Kuidas kombineerida erinevaid märgendamisviise?

Käesoleva töö käigus realiseeriti osaliselt kaks prototüüpi – üks eraldi rakendustest kokkupanduna ja teine MONAI raamistiku [3] toel. MONAI abil ülesseatud rakendustega prooviti läbi paari mudeli treenimise töövoog ning erinevate mudelitega järeldamise töövood.

2 Radioloogilistele uuringutele täiendava automaatanalüüsi tegemise protsess

Käesolevas peatükis pakutakse välja näidisprotsess meditsiinilistelt uuringutelt patoloogiliste leidude otsimiseks ja arstidele kuvamiseks. Protsess hõlmab endas ka asjakohase tagasiside kogumist mudeli parandamiseks. Lisaks kirjeldatakse andmekaitsega seonduvaid tähelepanekuid.

Eestis meditsiinis on digitaalsed andmed peamiselt reguleeritud kahe standardiga: HL7 ja DICOM (*Digital Imaging and Communications in Medicine*) [2]. Kuna HL7 on pigem dokumentide standard, siis keskendutakse käesolevas töös DICOM standardile, mida kasutatakse meditsiiniliste piltide ja nendega soetud andmete jagamiseks. Käesolev töö pakub lahenduse keskendudes radioloogilistele ülesvõtetele, mis arhiveeritakse DICOMi kujul.

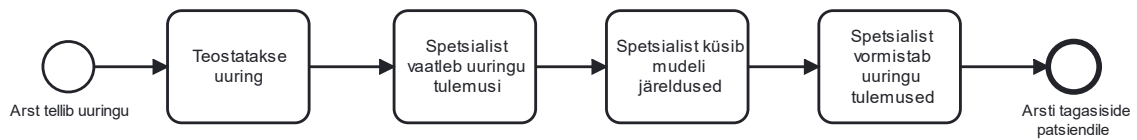
2.1 Üldine radioloogilise uuringu protsess

Järgnevalt kirjeldatakse lihtsustatud kujul lahti üldine protsess meditsiiniliste uuringute teostamisel (Joonis 1). Patsient läheb arsti juurde, kus arst selgitab välja, kas patsient vajab uuringut. Uuring teostatakse spetsialisti poolt, kes kirjutab üles leiud, ning tulemused saadetakse näiteks DICOM standardit kasutades patsiendiportaali. Arst vaatab uuringu tulemused üle ja annab tagasiside patsiendile ning määrab vajadusel ravi.



Joonis 1. Uuringute protsess hetkel

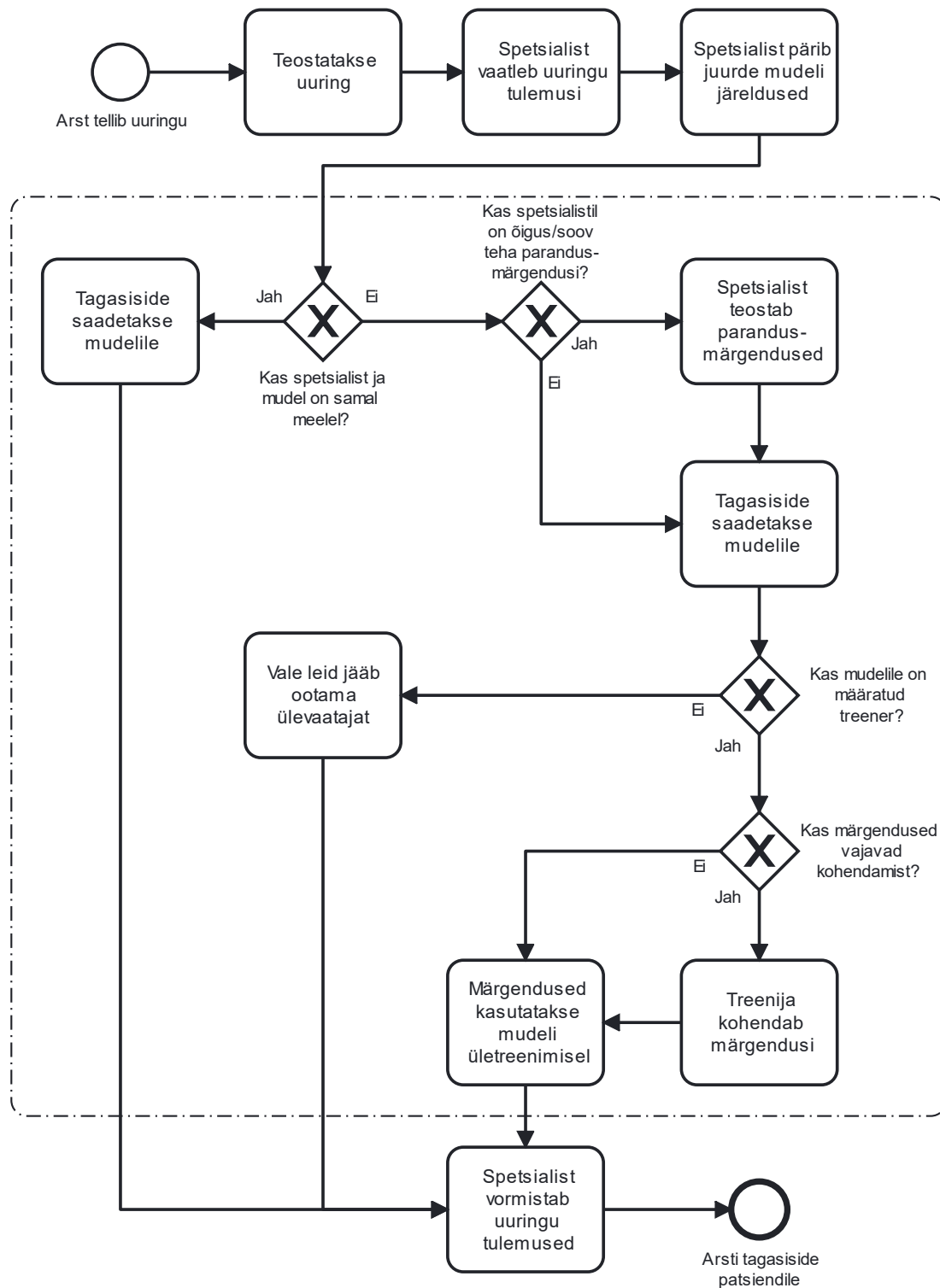
Käesoleva tööga välja pakutav näidisprotsess lisab tehisintellekti osa uuringu vaatlemise ja tulemuste vormistamise vahele (Joonis 2). Spetsialist küsib juurde mudeli järeldused ja teeb nende põhjal järeldused.



Joonis 2. Pakutav protsess lühidalt

Täpsemalt kirjeldab mudeli kasutamise protsessi Joonis 3. Kõigepealt peab spetsialist ise vaatlema uuringu tulemusi ning alles pärast seda saab ta pärida süvaõppemudelilt (edaspidi mudel) leidusid, kuigi automaatanalüüs on eelnevalt tehtud pärast uuringu masinast väljumist. See väldib olukorda, kus mudel tõmbab oma leidudega tähelepanu eemale ja seejuures võib jääda mõni teine leid märkamata. Radioloogias on sellise nähtuse jaoks kasutusel termin *satisfaction of search* (otsinguga rahulolu), mis kirjeldab olukorda, kus radioloogiliselt kujutisel ühe abnormaalsuse leidmine segab teiste leidmist [4].

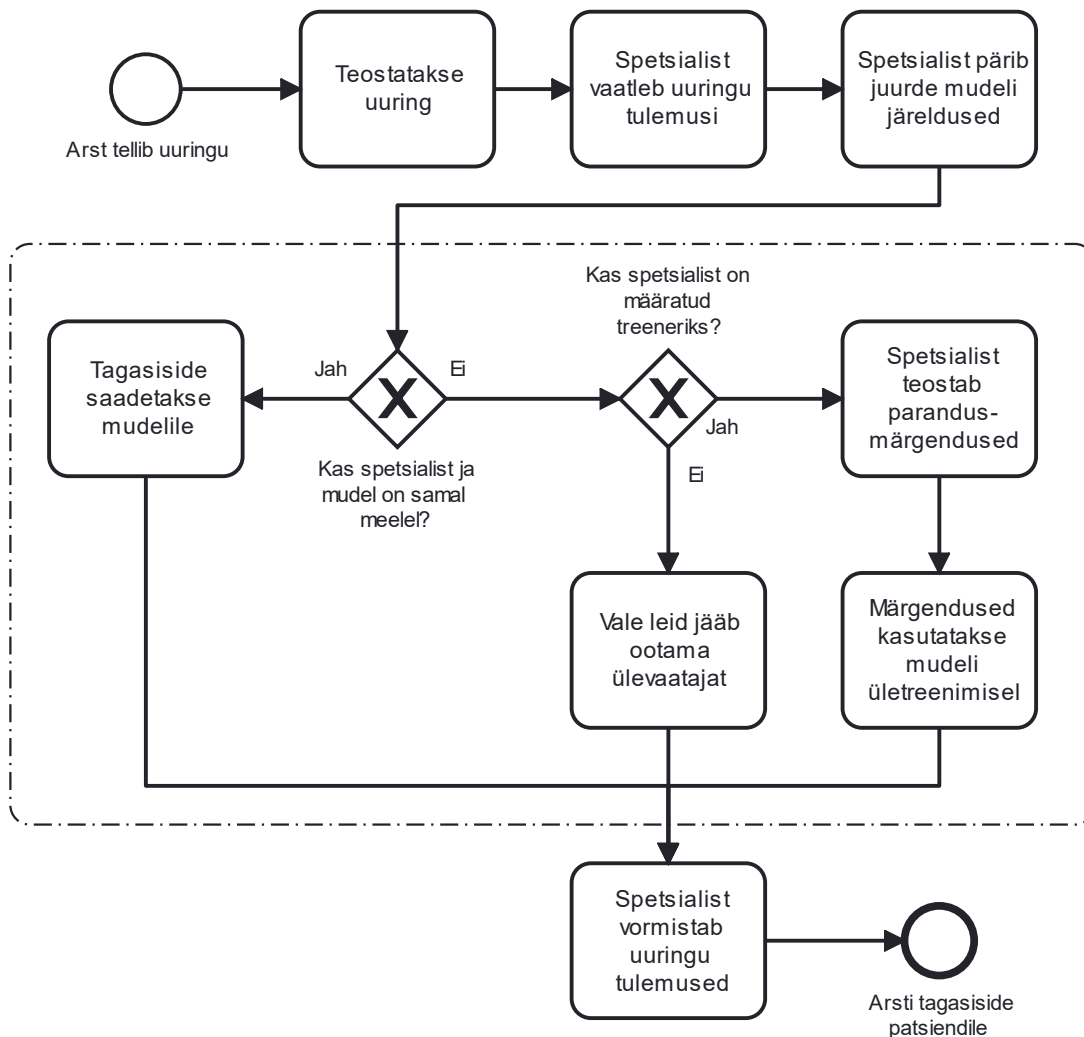
Spetsialist uurib üle mudeli leiud ja võrdleb neid enda omadega. Kui spetsialist on mudeli leidudega nõus, saadetakse positiivne tagasiside mudelile ja tehisintellekti osa lõppeb sellega. Kui spetsialist pole nõus mudeli tulemustega, siis sõltuvalt volitusest ja soovist saab ta teostada parandusi mudeli leidudes. Tagasiside ja parandusmärgendused (kui tehakse) saadetakse edasi mudelile. Kui mudelile on määratud treener, vaatab ta märgendused koos olemasolevate parandustega üle, teostab vajadusel omad parandused ja saadab need edasi mudeli ületreenimiseks. Kui on vaja lihtsamat varianti protsessi rakendamiseks, peaks tagasiside piirduma ainult tekstivormiga ja parandusmärgenduste tegemine jääks treeneri ülesandeks.



Joonis 3. Pakutav protsessi mudel

Alternatiivina pakutakse välja ka versioon protsessist, kus puudub eraldi mudeli treener (Joonis 4). Sellisel juhul on määratud spetsialistid, kellel on võimekus teostada ise parandusmärgendusi ja seejärel saata need otse mudelile ületreenimiseks.

Pärast leidude ülevaatamist vormistab spetsialist uuringu tulemused ning need saadetakse koos märgendustega patsiendiportaali, kust uuringule saatnud arst need kätte saab.



Joonis 4. Pakutav protsessi mudel, kus spetsialist on treener

2.1.1 Õige mudeli töövoog käivitamine

Mudelid on enamasti ühe kitsa leiu tuvastamiseks, seega on oluline protsessis on seadistada, millise mudeli töövoog käivitada millise kujutise korral. Seda saab hinnata DICOM kujutistes paiknevate metaandmete hulgast – näiteks silt (0018,0015) [5] tähistab vaadeldud kehaosa (*BodyPartExamined*), millel väärtused saavad näiteks olla rindkere (*CHEST*), kõht (*ABDOMEN*) ja rind (*BREAST*). Selle välja abil saab uuritava piirkonna järgi käivitada mudeleid. Silmas tuleb aga pidada, et igal juhul see ei sobi, sest näiteks nii mehel kui naisel on vaagen, aga eesnäärme mudelit on mõtet rakendada ainult mehel.

Alternatiiv oleks otsida võtmesõnu uuringu kirjelduse (StudyDescription) sildilt (0008,1030).

2.2 Andmete kasutamine

Isikuandmete kaitse seaduse [1] raames paika pandud andmekaitsemäärused rakenduvad ka meditsiiniteadusele. Andmekaitsemäärused kehtivad ainult isiklikele andmetele – andmetele, mille abil saab tuvastada indiviidi [6]. Nime ja aadressi kustutamisest tavaliselt ei piisa anonümiseerimiseks, näiteks USA 1990. aasta rahvaloenduse andmeid uurides leiti, et 87% juhtudel piisab andmete isikustamiseks postiindeksist, soost ja sünnikuupäevast [7].

Tõeliselt anonümiseeritud andmeid ei saa seostada indiviidiga, mis tähendab, et andmete verifitseerimine pole ka võimalik. Pseudonümiseeritud andmetel on tavaliselt isiku tuvastamist võimaldavad andmekirjete väljad asendatud ühe või mitme väljamõeldud identifikaatoriga [8]. Kuna seda saab veel kasutada isiku uuesti tuvastamiseks, klassifitseeruvad pseudonümiseeritud andmed samuti isikuandmeteks ja kuuluvad õigusakti reguleerimisalasse [9].

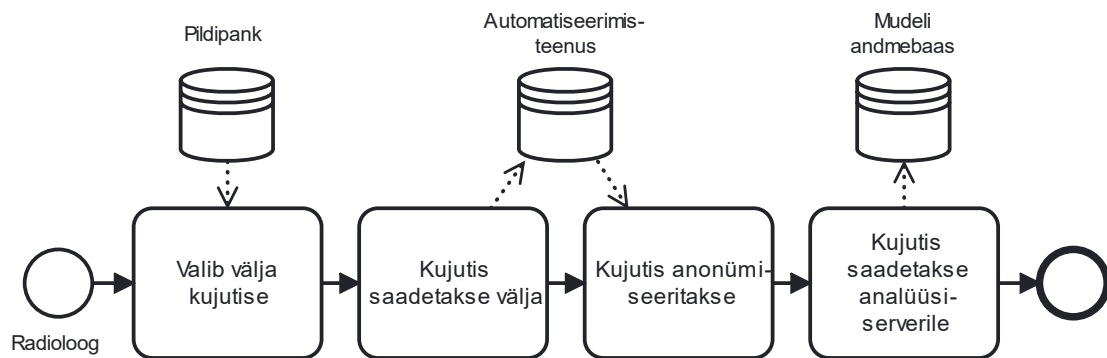
Euroopa Liidus on meditsiineseadmete definitsioonis loetletud muuhulgas tarkvara, mida kasutatakse meditsiinis eesmärgil haiguste, vigastuste, patoloogiliste protsesside uurimiseks ning diagnoosimiseks [10]. Radioloogias ja meditsiinis üldisemalt, peaks tehisintellekti meditsiineseadmed kasutatav süvaõpe pakkuma patsientide kohta infot ilma nende isikuandmeid vajamata. Isikuandmete kaitse üldmäärusest lähtudes peaks igasugune andmetöötlus olema *opt-in* ning tarbija nõusolek andmete kasutamiseks peab olema selge. See on kriitiline juhul, kui on vaja kasutada uuringuid mudelite treenimiseks või parandamiseks. Automaatanalüüsi tegemine ise oleks osa uuringu tegemiseks ja seeläbi ei vajaks eraldi nõusolekut.

2.2.1 Andmete hoiustamine

Eelnevat silmas pidades peab vältima isikustavate uuringute väljumist haiglasestest süsteemidest. Eristama peaks vähemalt nelja DICOM serverit – esiteks ühte, kus paiknevad tehtud uuringud, sellele vastab Eestis pildipank ehk PACS. Andmete vahendamiseks on vaja serverina automatiseerimisteenust, mis käitub kui marsruuter ülejäänud kolme serveri vahel. Lisaks on juurde vaja serverit, kuhu saadetakse uuringud

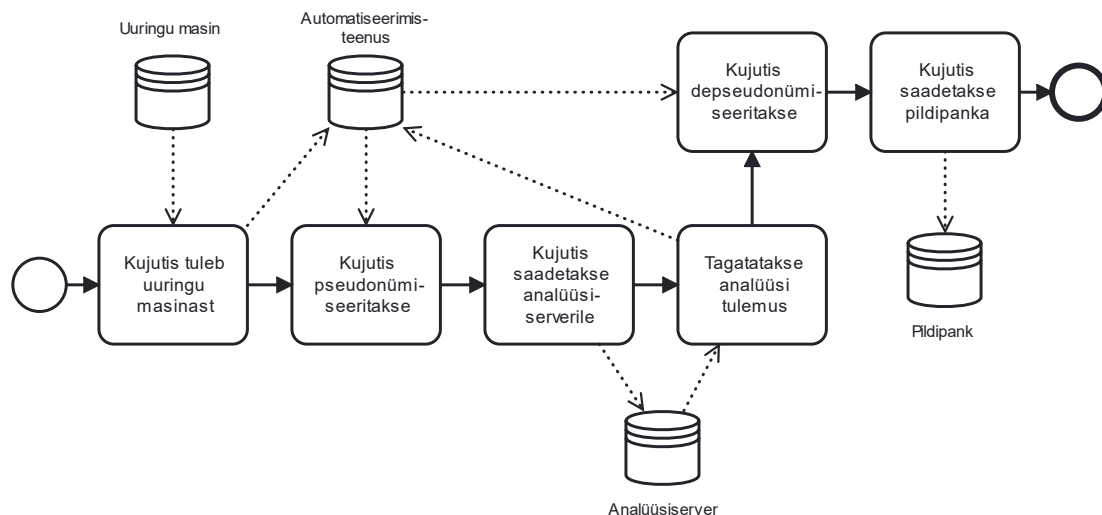
käitsi märgenduste tegemiseks, mida ka hiljem hoiustada mudeli täiendava treenimise jaoks. Et mitte vastuollu minna Isikuandmete kaitse üldmäärusega (IKÜM), peab mudelile saadetavad uuringud eelnevalt anonümiseerima haigla pinnal. Joonis 5 kirjeldab kahe esimese serveri kasutamist märgendamisele saatmise protsessis.

Esiolgsel märgendamisele saadetavate uuringute valik peaks olema radioloogi töökorralduslik otsus. Radioloog peab ise jälgima, kas tal on treeningandmetes piisavalt normi- ja patoloogilisi leidusid. Hilisemas faasis, kui mudel juba töötab, piisaks kui parandusmärgendamisse saata vaid vigaste hinnangutega uuringud. Kuna andmemahud on suured, peaks andmete hoiustamise realiseerima kohalikul kettal. Vajadusel on võimalik leida ka sobiv pilvepakkuja.



Joonis 5. Uuringute saatmine märgendamisse

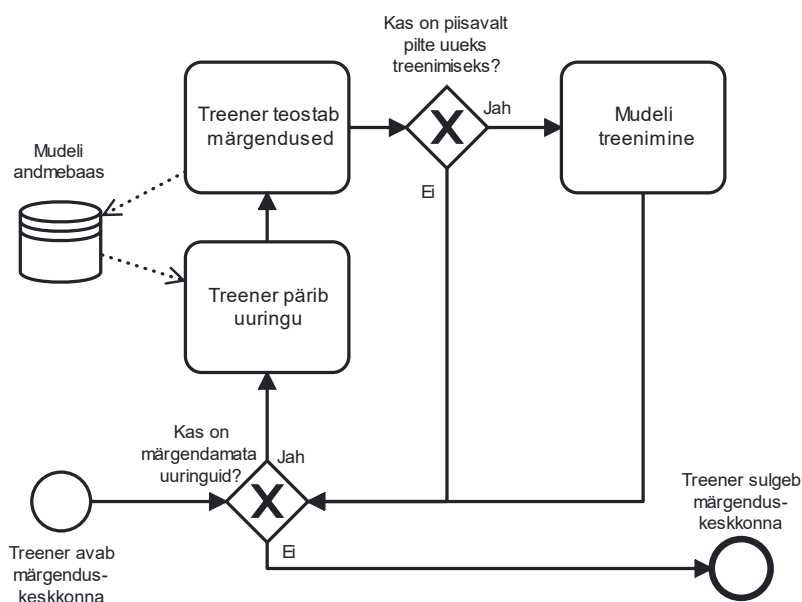
Viimaks on juurde vaja serverit, kuhu saadetakse kujutised automaatselt mudeli kasutamiseks. Joonis 6 kujutab andmete liikumist järeldamise protsessis. Uuringust tulevad andmed pseudonümiseeritakse haigla pinnal automatiseerimisteenus, seejärel saadetakse analüüsiserverile. Pärast analüüsi kustutatakse andmed analüüsiserverist ja tagastatakse kujutis automatiseerimisteenus, mis saadab depseudonümiseeritud kujutise koos analüüsiga pildipanka. Kui on eelnevalt saadud patsiendi nõusolek, võib andmed kasutada ka anonümiseeritud kujul väljaspool haiglat mudelite treenimiseks ja parandamiseks, kui näiteks mudel on vale hinnangu andnud.



Joonis 6. Andmete liikumine järeldamise protsessis

2.3 Mudeli treenimine

Kui anonümiseeritud treeningandmed on jõudnud märgendamiskeskonda, saab alustada märgendamisprotsessiga (Joonis 7). Treener avab märgendamiskeskonna ning pärib märgendamata uuringuid. Kui neid leidub, siis teostab ta märgendused, salvestab ning edastab mudelile treenimiseks. Mudeli täiendavaks treenimiseks tuleb märgendada piisav kogus pilte. Kirjeldatud protsessi saab kasutada nii mudeli esialgses treenimisfaasis kui ka töötava mudeli parandamisel.

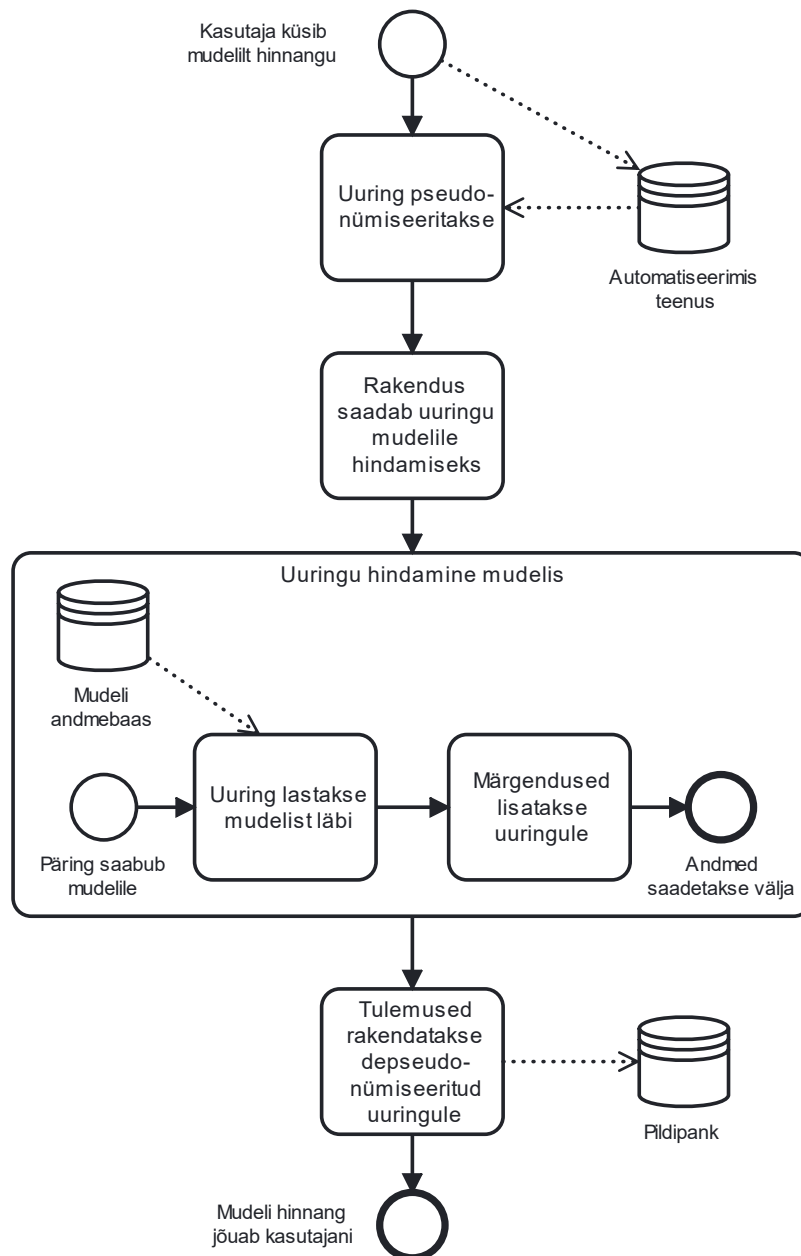


Joonis 7. Märgendamise protsess

2.4 Mudeli kasutamine järeldamisel

Kasutades mudelit järelduste tegemisel, tuleb samuti järgida IKÜMi. Kuna siinkohal on uuringut tehes patsient juba nõustunud oma andmete töötlemisega, siis piisab uuringu pseudonümiseerimisest mudeli hinnangu pärimiseks. Sel juhul ei ole masinõppe mudel võimeline uuringut isikustama, kuid haiglasüsteem suudab tagastatud uuringu tulemust seostada patsiendiga.

Protsessi tehisintellekti hinnangu küsimiseks kirjeldab Joonis 8. Pseudonümiseeritud uuring saadetakse päringuga mudelile. Mudel võtab selle vastu, laseb oma algoritmist läbi ning lisab juurde märgendused. Seejärel tagastab ta päringu tulemuse koos uuringuga ja see depseudonümiseeritakse haiga süsteemis. Pärast järelduste tegemist ja tulemuste tagastamist andmed kustutatakse.



Joonis 8. Mudeli kasutamise protsess

2.4.1 Tagasiside andmine

Mudel peab võimaldama ka tagasiside vastu võtta. Kuna haiglate vahel võivad olla kasutuses erinevad kujutiste vaaturid, peaks tagasiside andmine olema sõltumatu vaaturist. DICOM pildid ei toeta linkide lisamist, kuid see-eest toetavad piltide lisamist kihtidena. Seega pakutakse käesoleva tööga välja, et tagasiside andmiseks lisatakse hinnanguga DICOMi külge QR-koodiga pilt, mis suunab tagasiside andmise keskkonda.

3 Realiseeritud prototüüp

Järgnevas peatükis kirjeldatakse prototüübi tehnilisi nõudeid ning tarkvara MONAI, mille põhjal realiseeriti prototüüp eksperimentide tegemiseks. Alampeatükk 3.1 kirjeldab üldisemalt MONAI võimekusi, seejärel kaardistatakse nõuded ning selgitatakse süvitsi MONAI sisalduvaid komponente.

3.1 MONAI

Projekt MONAI algatati NVIDIA ja King's College Londoni poolt, et luua tehisintellekti teadlaste kogukond heade tavade arendamiseks ja jagamiseks [3] meditsiiniliste kujutiste analüüsiks. MONAI on avatud lähtekoodiga PyTorch'i peal üles ehitatud raamistike kogu, mille eesmärk on kiirendada innovatsiooni meditsiinilistes piltuuringutes. Tehisintellekti elutsükkel on jagatud kolmeks:

1. MONAI Label – intelligentne märgendamistöööriist, mis kasutab märgendamiseks kuluva aja ja vaeva vähendamiseks tehisintellekti. Mudel õpib pidevalt kasutaja interaktsioonide abil.
2. MONAI Core – MONAI lipulaeva teek, mis pakub domeeni-spetsiifilisi võimekusi meditsiiniliste mudelite treenimiseks. Võimekused sisaldavad endas meditsiini-spetsiifilisi pilditeisendusi, 3D segmentatsiooni algoritme nagu UNETR ja AutoML raamistik DiNTS.
3. MONAI Deploy – standard meditsiinilise tehisintellekti rakenduste jooksutamiseks kliinilises keskkonnas. MONAI Deploy loob sammud teadlastele ja arstidele iteratiivseks töövooks.

3.2 Andmete hoiustamine

Andmete hoiustamiseks tuleb kasutada DICOM failide salvestamist võimaldavat serverit. Server peaks toetama DICOMweb standardit. DICOMweb standard paneb paika kogumi RESTful teenustest, mis võimaldab veebiarendajatel kasutada standardseid tööriistaid [11]. REST on akronüüm *REpresentational State Transferist* (esinduslik

olekuedastamine), mis veebi API'de ja teenuste arhitektuuristiil koos oma põhimõtete ja piirangutega [12].

DICOMweb defineerib viis teenust [11]:

- Pärimine (*Query, QIDO-RS*), mille abil otsida DICOM objekte. Sellega saab pärida patsiendipõhiselt uuringuid, mille käigus saadetakse unikaalsed identifikaatorid, millega näiteks edasisi teenuseid kasutada [13].
- Võtmine (*Retrieve, WADO-RS*), mille abil võtta DICOM objekte; samuti ka üksikute DICOM instantside võtmiseks (*WADO-URI*).
- Salvestamine (*Store, STOW-RS*) DICOM objektide salvestamiseks.
- Töönimekiri (*Worklist, UPS-RS*), millega hallata töönimekirja kirjeid. Sellega kirjeldatakse ka, kuidas teavitused ja tellimused töötavad [14].
- Võimekused (*Capabilities*), millega pärida konkreetse DICOMwebi lõpp-punkti toetatud teenuseid [15].

Lisaks defineeritakse veel standardid andmete kodeerimiseks JSON ja XML formaatidesse.

3.2.1 Orthanc

MONAIga ühildub mugavalt DICOM serverina Orthanc [16]. See on avatud lähtekoodiga meditsiiniliste piltide ökosüsteem, mis võimaldab PACS¹ (*picture archiving and communication systems*, piltide arhiveerimise ja sidesüsteemid) süsteemide haldajatel automatiseerida ja optimeerida haiglapõhiseid piltide liikumisi. Orthanc on loodud arendajaid silmas pidades, võimaldades lihtsalt pistikprogramme ja välist tarkvara luua minimaalsete teadmistega DICOM standardist.

¹Tervishoiu asutustes kasutatav tehnoloogia meditsiiniliste piltide ja aruannete hoiustamiseks ning edastamiseks [40].

3.3 Piltide märgendamine

Enne mudeli treenimist tuleb andmed ette valmistada. Parimat täpsust on võimalik saavutada andmeid segmenteerides. See on küll ajakulukam kui klassifitseerimine või piirdekastide joonistamine, kuid segmentatsioonide põhjal treenitud mudel pakub selle võrra täpsemat hinnangut radioloogidele interpreteerimiseks. Järgnevas alampeatükis kirjeldatakse MONAI Labeli võimekusi segmenteerimise protsessi lihtsustamiseks.

3.3.1 MONAI Label

MONAI Label on avatud lähtekoodiga platvorm tehisintellekti rakendustele, mille eesmärk on 3D-meditsiiniliste kujutiste annoteerimisele kuluvat aega vähendada. [17] MONAI Label võimaldab teadlastel avaldada oma rakendusi teenustena, mida saavad arstid kasutada läbi oma eelistatud kasutajaliidese. MONAI Label toetab muuhulgas serverirakendust OHIF ning kohalikult paigaldatud rakendust 3DSlicer.

MONAI Label rakendused on kättesaadavad teenusena läbi MONAI Label Serveri, mis avab funktsionaalsuse läbi RESTful API [17]. Andmehaldus MONAI serveris toetab nii kohalikult salvestatud kui ka DICOMweb toega DICOM serveris paiknevaid andmekogusid, andes võimaluse integreerida see oma PACSiga.

MONAI Label toetas käesoleva töö tegemise käigus nelja annoteerimise viisi [17]. DeepGrow on interaktiivne segmenteerimise mudel, kus kasutaja juhendab segmenteerimist positiivsete ja negatiivsete klikkidega. Positiivsete klikkidega laiendatakse segmentatsiooni, kaasates huvipiirkondasid; negatiivsete klikkidega välistatakse segmentatsioonist huviväliseid piirkondi. DeepGrow lubab märgendada ainult ühte märgendit korraga.

DeepEdit laiendab DeepGrow'i klikipõhist segmentatsiooni, lubadest klikivaba segmentatsiooni järeldamist ja klikipõhist segmentatsiooni muutmist. See tähendab, et sihitakse kasutajakogemuse hõlbustamist, luues kasutajale esialgse ennustatud segmentatsiooni maski, mida saab vajadusel parandada.

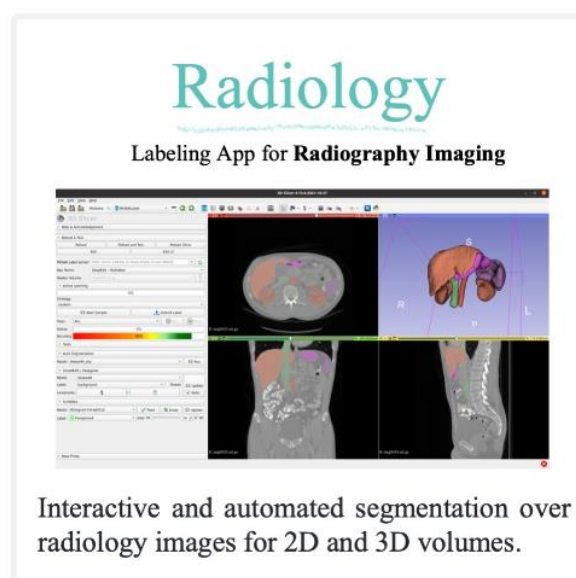
Scribbles-based Segmentation (kriitseldustel põhinev segmentatsioon) lubab kasutajal teostada vabakäelisi märgendusi. Seda saab kasutada kahel eri viisil:

1. Kritiseldustel põhinev veebipõhine tõenäosuslik segmentatsioon, kus kasutaja kritiseldab segmentatsiooni genereerimiseks. Kasutaja teeb kritiseldusi, mis määratlevad huviobjektid ja tausta ning nende põhjal ehitatakse tõenäosuse mudel, mis võimaldab piiritleda huviobjekte. Kritiselduste asemel on võimalik kasutada ka piirdekaste.
2. Kritiseldustel põhinev CNN (*Convolutional neural network* – konvolutsiooniline närvivõrk) segmentatsiooni täpsustamine, kus süvaõppemudeli segmentatsiooni täiendatakse kritiselduste põhjal. Eeltreenitud CNN mudeli kaasamine kritiselduste tegemisele vähendab oluliselt annoteerimisele kuluvat aega.

Automaatne segmenteerimine on mitte-interaktiivne algoritm, mis põhineb mõnel MONAI Core'is olemasoleval CNNil, näiteks UNet'il. See on sarnane DeepGrow ja DeepEdit struktuuriga, kuid ei võimalda sisendpildile anda lisakanalitega huviobjektide ja tausta infot. Siiski saab automaatset segmentatsiooni hiljem manuaalselt muuta tavaliste segmenteerimistööriistadega.

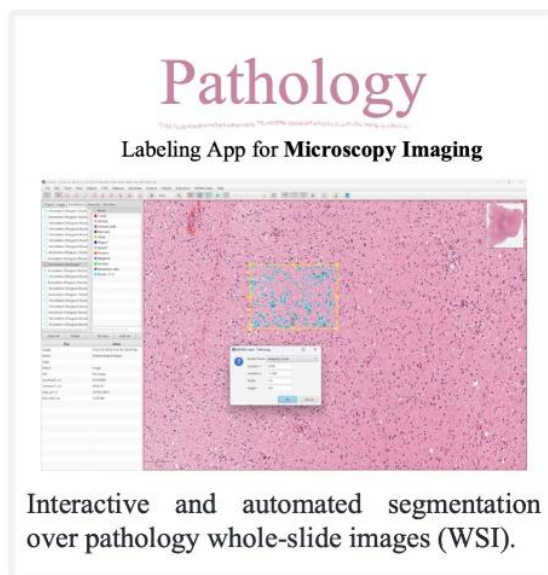
3.3.2 Näidisrakendused MONAI Labelis

Käesoleva töö tegemisel ajal oli MONAIs neli näidisrakendust [18]. Radiology (radioloogia, Joonis 9) rakenduses on näidismudelid interaktiivse ja automaatse segmenteerimise teostamiseks radioloogilistel kujutistel. Muu hulgas sisaldab see UNet'i ja UNETRi süvaõppemudeleid kõhuõõne organite segmenteerimiseks.



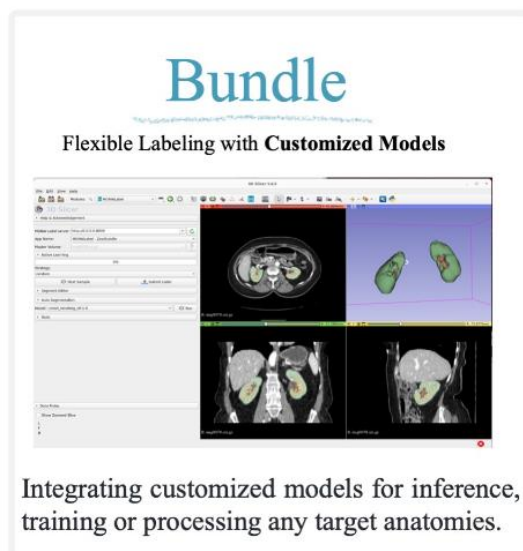
Joonis 9. MONAI Radiology rakendus [18]

Pathology (patoloogia, Joonis 10) rakenduses on näidismudelid interaktiivseks ja automaatseks segmenteerimiseks mikroskoopia piltidel (WSI ehk *DICOM Whole-Slide Images*, digitaliseeritud mikroskoobi slaidid [19]). Rakendus sisaldab endas *nuclei* mitme sildiga segmenteerimist epiteel-, side- ja pehmete kudede rakkudel, neoplastilistel, põletikulistele ja surnud rakkudel.



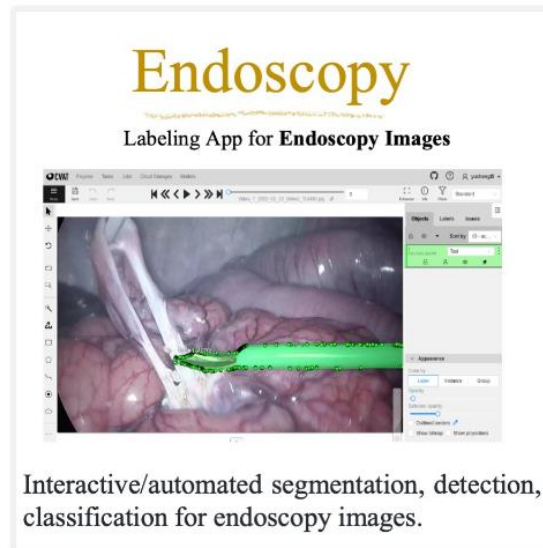
Joonis 10. MONAI Pathology rakendus [18]

Bundle (pundar, Joonis 11) rakendus võimaldab kasutada mudeleid treenimiseks ja järelduste tegemiseks igal anatoomial. Bundle kasutab Model-Zoo eeltreenitudid süvaõppemudeleid segmenteerimiseks.



Joonis 11. MONAI Bundle rakendus [18]

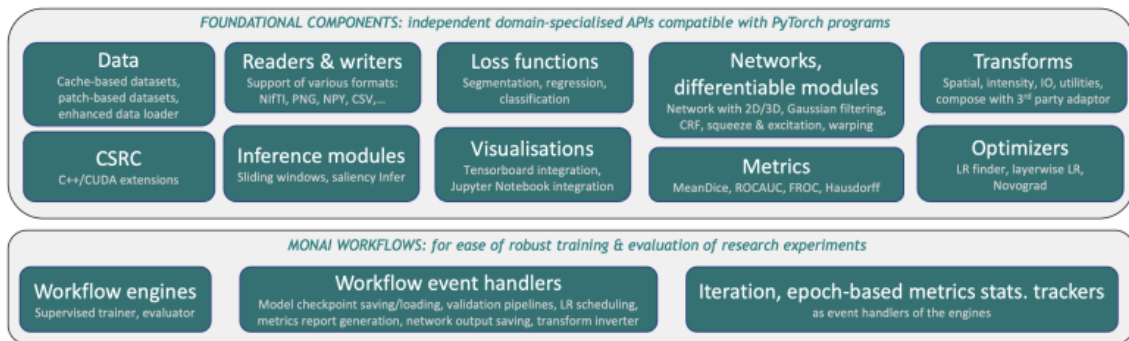
Endoscopy (endoskoopia, Joonis 12) rakendus spetsialiseerub interaktiivse ja automaatse segmenteerimise tegemisele 2D kujutistel endoskoopia kasutusjuhus. Koostöös CVATiga plaanitakse demonstreerida täielikult automatiseeritud aktiivse õppe töövoogu mudeli treenimiseks ja peenhäälestamiseks.



Joonis 12. MONAI Endoscopy rakendus [18]

3.4 Mudelite treenimine MONAI Core'iga

MONAI Core on PyTorch'i baasil põhinev tervishoiu süvaõppe raamistik ning on omaks võtnud põhimõtte, et võimekused peaksid olema arendatud, järgides PyTorch'i põhijuhiseid ja -mõtteid [20]. Tarkvara sisaldab endas komponente (Joonis 13) andmekogudele (*Data*), kirjutajatele ja lugeritele (*Readers ja writers*), kahjufunktsioonidele (*Loss functions*), võrkudele (*Networks, differentiable modules*), teisendustele (*Transforms*), C++ ja CUDA laiendustele (*CSRC*), järeldustele (*Inference modules*), visualiseerimisele (*Visualisations*), meetrikaks (*Metrics*) ja optimeerimiseks (*Optimizers*). Et erinevaid komponente integreerida, sisaldavad moodulid endas veel töövoogu mootoreid (*Workflow engines*), sündmuse haldajaid (*Workflow event handlers*) ja iteratsiooni ning epohhi-põhiseid meetrika jälgijaid (*Iteration, epoch-based metrics stats. trackers*).



Joonis 13. MONAI Core moodulid [20]

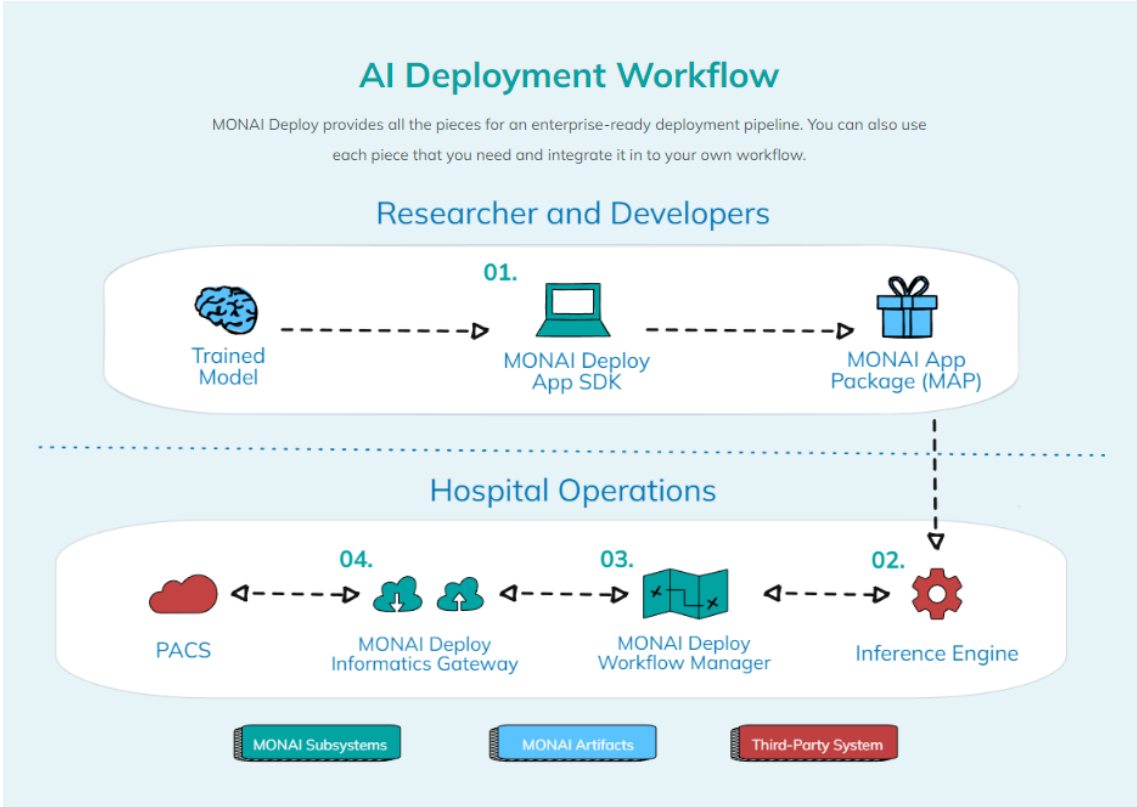
Andmeteisendus ja rikastamine on olulised süvaõppes, aidates kaasa andmetötlusele alates piltide laadimisest failist kuni keerukate geomeetriliste andmete teisendamiseni [20]. MONAI Core sihib pakkuda nii domeeni-spetsiifilisi kui ka meditsiiniliste kujutiste süvaõppes kasutatavaid teisendusi. Teisendused on mugavalt kokku pandavad töötlemisvoogu.

3.5 Mudelite kasutamine järeldamiseks ja visualiseerimiseks MONAI Deploy abil

Treenitud mudeleid saab seejärel kasutada järelduste (inglise keeles *inference*) tegemiseks. Et järeldustest oleks ka kasu radioloogile, tuleb neid visualiseerida, näiteks salvestades eraldi seeriana DICOMi peale.

MONAI Deploy sihib saada meditsiiniliste tehisintellekti rakenduste arendamise, testimise ja tarne standardiks [3]. Raamistik defineerib tehisintellekti juurutamise töövoogu (*AI Deployment Workflow*, Joonis 14) läbi nelja sammu:

1. *App Development* - Tehisintellekti rakenduste ja treenitud mudelite pakkimine MAPiks (*MONAI App Package*, MONAI rakenduse pakett).
2. *Inference Service* - Järelduste tegemise RESTful teenus MAPide jooksutamiseks.
3. *Workflow Manage* - Töövoos haldaja, käivitab järelduste tegemist vastavalt seadistatud reeglitele
4. *Informatics Gateway* - Informaatika värav tehisintellekti rakenduste ühendamiseks PACSidega.



Joonis 14. MONAI AI Deployment Workflow [3]

4 Eksperimendid prototüübil

Järgnevas peatükis kirjeldatakse MONAI tarkvara peal tehtud eksperimente. Kõigepealt prooviti üles seada töövoog kujutistel järelduste tegemiseks. Seejärel täiendati töövoogu MONAI eeltreenitud mudelite põhjal, proovides lisada nii ühe kui ka mitme mudeliga rakendust. Viimaks prooviti trennida MONAI Core abil ise mudel, pakkida see rakenduseks ja lisada see eraldi töövoogu. Ise trennitud kopsuhaiguste klassifikatsiooni mudelid katsetati läbi ka käesoleva töö kaasjuhendaja hangitud Eestis tehtud kopsuröntgenite peal.

4.1 Järelduste tegemine MONAI Deploy Expressi näitel

Mudelite katsetamiseks seati juhendi¹ abil serverisse üles MONAI Deploy Express keskkond Dockeri konteineri peal. Deploy Express on kompaktsem versioon MONAI Deploy töövoost, mille eesmärk on lihtsustada MONAI mudelite testimist ja valideerimist varajases töövoos ülesseadmise faasis, kus on olulised lihtne kasutus ja kiire startaeg. Näidisrakendus paljastab pordi 8042 peal Orthanc DICOM Serveri, mille peale laeti töö käigus üles näidisuuringud. Turvalisuse suurendamiseks lükati sisse Ortancis olemasolev autentimine. Seejärel valmistati ette Docker Image MONAI mudeli rakendusest (edaspidi MAP). MAP seadistus tehti json failis ning saadeti pordil 5000 olevale töövoos haldajale (*workflow manager*) käsu abil (Joonis 15), kus töövoog.json asendati vastava faili nimega.

```
curl --request POST --header 'Content-Type: application/json' --data
"@sample-workflows/toovoog.json" http://localhost:5001/workflows
```

Joonis 15. Käsklus MAP seadistuse saatmiseks MONAI töövoos haldajale

Seejärel navigeeriti Orthancis sobiva uuringu alla ja käivitati *Interact* menüüst käsklus *Send to DICOM modality*, mis itereeris läbi kõik seadistatud töövoos, mis omakorda

¹ <https://github.com/Project-MONAI/monai-deploy/tree/main/deploy/monai-deploy-express#running-a-monai-deploy-workflow>

hindasid oma parameetrite põhjal, kas teostada hinnangu andmist kujutisel. Rakendatud MAP hinnangud lisati DICOM kujutisele külge uue seeriana.

4.2 MONAI eeltreenitud mudelid

MONAI Model Zoo¹ pakub mitmeid eeltreenitud mudeleid, millest mõningad katsetati läbi käesoleva töö raames. Käesoleva töö tegemise ajal pakkus Model Zoo eelkõige 3D segmenteerimise mudeleid kompuutertomograafia uuringutele (CT), kuid ka mõningaid annoteerimise, MRT uuringute segmenteerimise, mikroskoopia uuringute ja muid mudeleid.

4.2.1 Eesnäärme segmenteerimine – näide ühe mudeli põhjal

MONAI Deploy Expressi töövoogu prooviti täiendada eesnäärme segmenteerimise mudeliga. MONAI Model Zoost võeti mudel (*Prostate mri anatomy*), mis oli treenitud UNet arhitektuuril T2w MRT kujutiste 3D segmenteerimiseks. Mudeli MAPiks tegemiseks kasutati juhendit². Loodi rakendus nimega `ai_prostate_seg_app`, mis pakiti koos mudeliga MAP Docker Image'iks nimega `prostate_seg_app`. Seejärel lisati mudeli konfiguratsioon `multiple-workflows.json` faili, kus `task_destinations` alla täpsustati tingimus, et DICOM kujutisel peab olema uuritud kehaosa silt (0018,0015) [5] väärtusega CHEST (rindkere).

Uuendatud töövoogu json saadeti töövoohaldajale. Seejärel saadeti testpildid järeldamisele. Testimiseks kasutati kujutisi Cancer Imaging Archive'ist³ ning järelduste visualiseerimiseks kasutati MITK Workbench rakendust⁴.

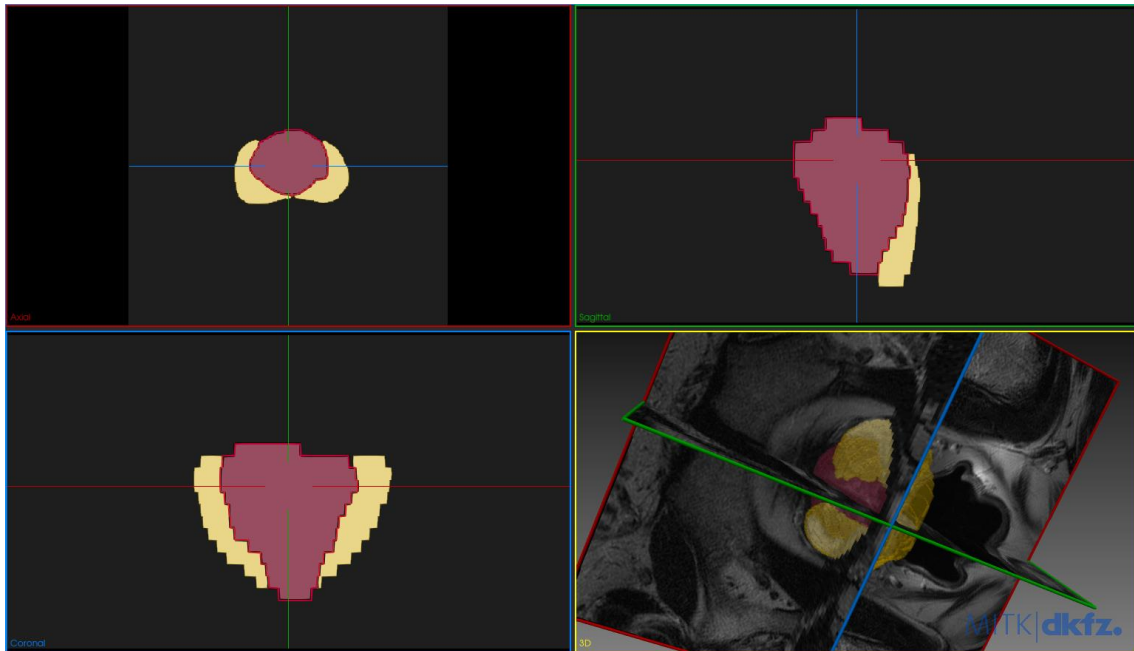
Joonis 16 visualiseerib esimest DICOM kujutist, millel järeldusi tehti. Kaasjuhendaja hinnangul oli segmentatsioon päris täpne.

¹ <https://monai.io/model-zoo.html>

² https://docs.monai.io/projects/monai-deploy-app-sdk/en/latest/getting_started/tutorials/monai_bundle_app.html

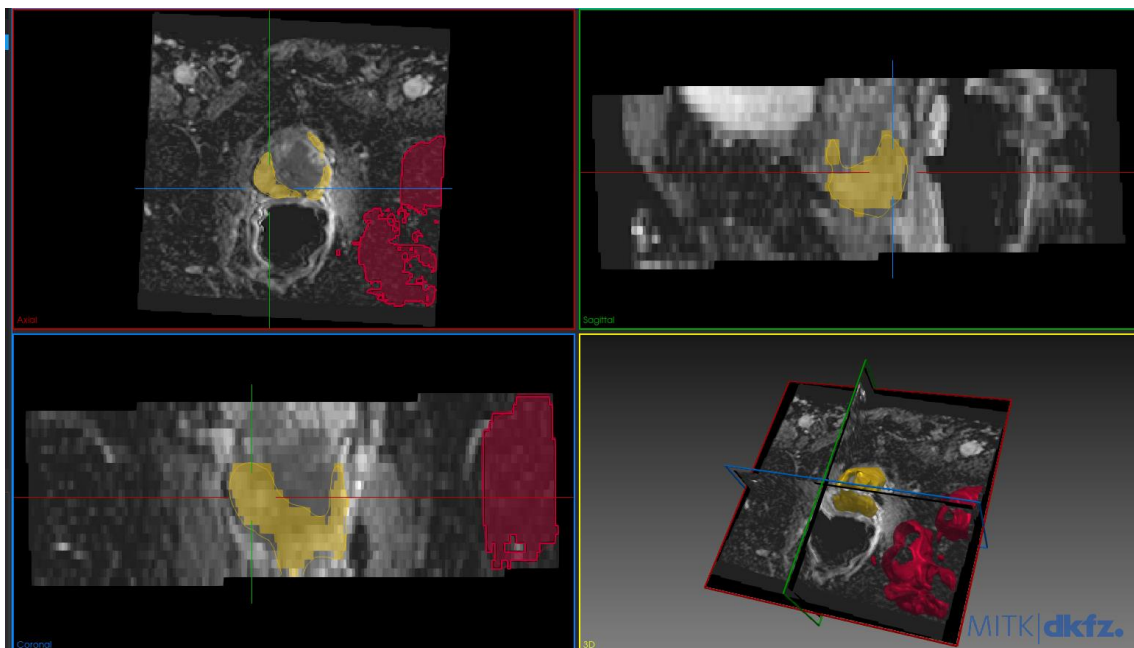
³ <https://www.cancerimagingarchive.net/nbia-search/?CollectionCriteria=PROSTATE-MRI>

⁴ [https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_\(MITK\)](https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_(MITK))



Joonis 16. Eesnäärme segmenteerimise õnnestunud näide

Joonis 17 kujutab teist eesnäärme MRT näidet, mille järeldamise töövoogu rakendati. Sellel on näha, et segmenteerimine on kaootiline ja ebatäpne.



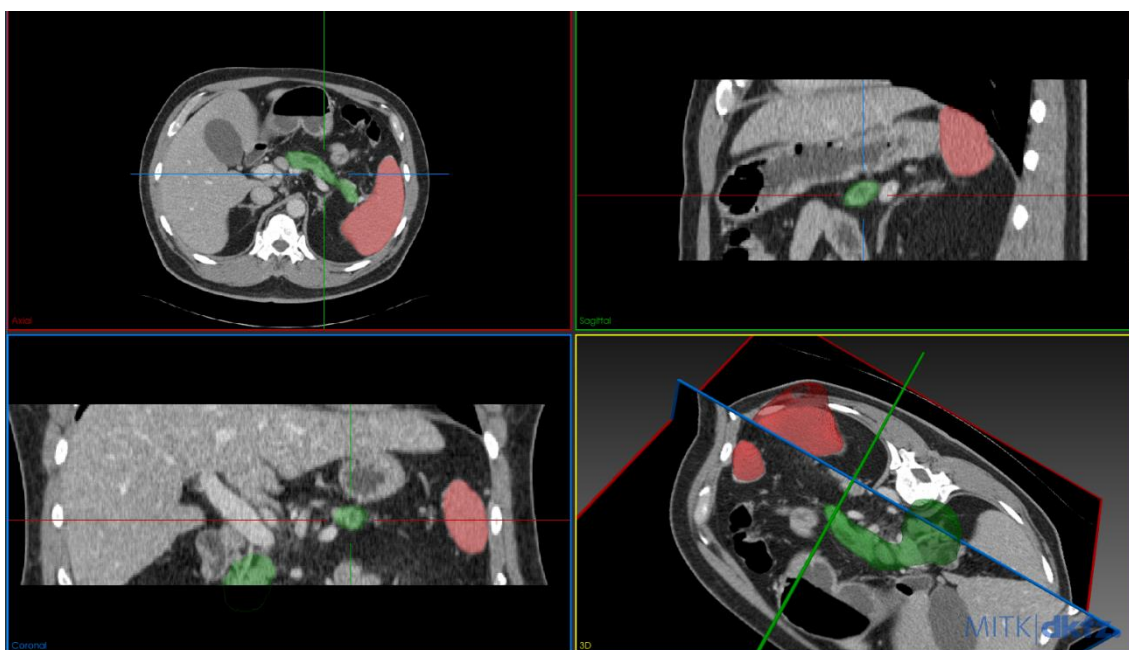
Joonis 17. Eesnäärme segmenteerimise ebaõnnestunud näide

Mudeli rakendus ja konfigureeritud töövoogu failid on kättesaadavad lisas 2 kirjeldatud lingiga.

4.2.2 Kõhunäärme ja maksa segmenteerimine - Mitme mudeli rakendamine

Lisaks prooviti lisada MONAI Deploy Expressi mudelite hulka MAP, mis põhines mitmel Model Zoos oleval mudelil. Selleks kasutati juhendit¹, mille abil tõmmati kõhunäärme ja maksa segmenteerimise mudelid koos näidis-DICOMiga arhiivis ai_multi_model_bundle_data.zip. Mudelistest ja monai-deploy-app-sdk repositooriumis olevast ai_multi_ai_app rakendusest loodi MAP Docker Image. Seejärel seadistati Deploy Expressi multiple-workflows.json fail. *Task_destinations* alla lisati mudel koos tingimusega, et DICOM kujutisel peab olema uuritud kehaosa silt (0018,0015) [5] väärtusega ABDOMEN (kõht). Ülejäänud seadistused tehti analoogselt näitega.

Seadistatud json fail saadeti töövoohaldajale, seejärel käivitati DICOM kujutisel *Send To Dicom Modality*. Näidiskujutisena kasutati arhiivis olevat DICOM kujutist. Tulemust visualiseeriti MITK Workbench² rakenduse abil (Joonis 18).



Joonis 18. Kõhunäärme ja maksa samaaegne segmenteerimine MONAI mudeli abil

¹ https://docs.monai.io/projects/monai-deploy-app-sdk/en/latest/getting_started/tutorials/multi_model_app.html

² [https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_\(MITK\)](https://www.mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_(MITK))

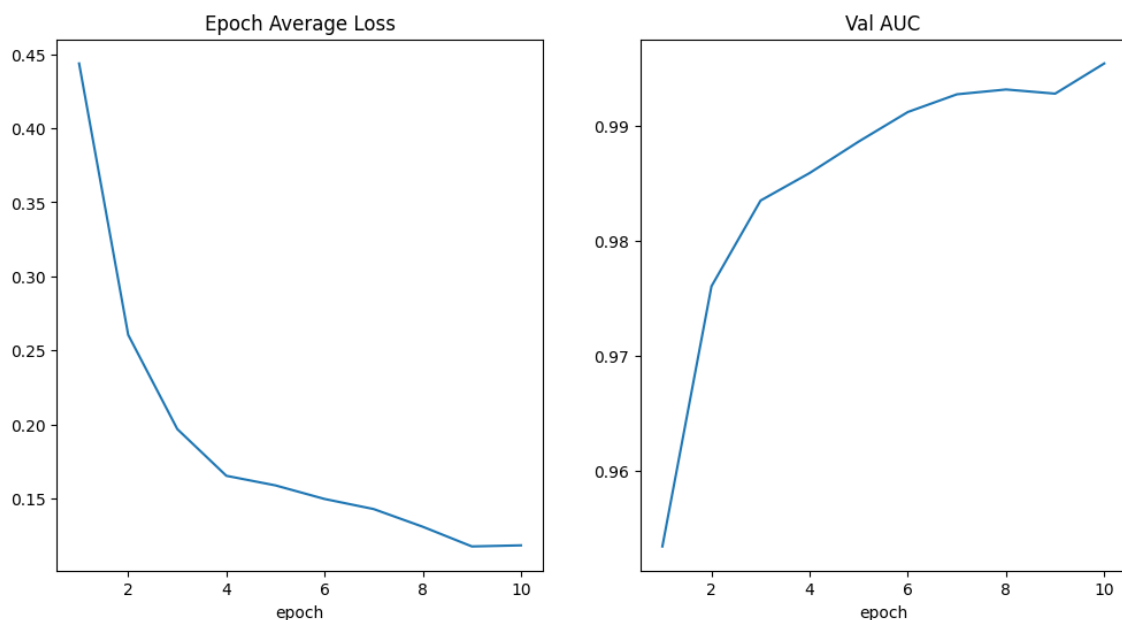
4.3 Mudeli treenimise ja järeldamise töövoogude prototüüp

MedMNIST kopsupõletiku andmekogu näitel

Seejärel prooviti MONAIs läbi ise mudeli treenimise töövoog. Valmis mudel pakendati MAP rakenduseks, konfigureeriti MONAI Deploy Expressis ja kasutati testkujutistel järelduste tegemiseks.

4.3.1 Mudeli treenimine MedMNIST PneumoniaMNIST andmekogu põhjal

Kõigepealt treeniti analoogselt juhendile¹ kopsupõletikku tuvastav mudel. Treenimiseks kasutati MedMNIST [21] [22] kopsuröntgenitel põhinevat kopsupõletiku andmekogu PneumoniaMNIST [23]. Kogu sisaldas endas 5856 kujutist tervetest ja põletikuga kopsudest. Andmekogu kujutistel tuli skaleerida intensiivsus, lisaks suurendada mõõtudele 64x64, kuna algselt olid nad suurusega 28x28 pikslit. Seejärel rakendati andmete rikastamise funktsioone, mis sisaldasid endas suvalisi keeramisi, pööramisi ning suurendamisi ja vähendamisi. DenseNet121 mudelit treeniti 10 epohhi, saavutades parima ROC-kõvera aluse pindala väärtusega 0.9954 10. epohhil (Joonis 19), mis on võrreldav MedMNIST enda tehtud võrdlusuuringute² parima väärtusega 0.991.



Joonis 19. Kopsupõletikku klassifitseeriva mudeli treenimise meetrika

¹ https://docs.monai.io/projects/monai-deploy-app-sdk/en/latest/notebooks/tutorials/02_mednist_app.html

² <https://medmnist.com/>

Mudeli treenimise vihk on kättesaadaval lisa 3 viidatud lingil. Tabel 1 kajastab mudeli täpsust andmekogu testhulga peal.

Tabel 1. PneumoniaMNIST kopsupõletiku klassifitseerimise mudeli täpsuse klassifikatsiooniraport

	precision	recall	f1-score	support
none	0.9464	0.6795	0.7910	234
pneumonia	0.8355	0.9769	0.9007	390
accuracy			0.8654	624
macro avg	0.8910	0.8282	0.8459	624
weighted avg	0.8771	0.8654	0.8596	624

4.3.2 Pakendamine rakenduseks

Sarnaselt eelnevate näidetega loodi seejärel mudeli kasutamiseks rakendus `ai_pneumonia_classifier_app`. Kuivõrd eelnevalt oli tegemist segmenteerimise rakendustega, mis väljastasid järeldamise käigus segmenteerimise maske, siis kopsupõletiku klassifitseerimise rakendus väljastas hinnangu, kas kujutisel leidis põletik või mitte. Tekstilise väljundi lisamist toetas MONAI klassi `DICOMTextSRWriterOperator` abil. Mudel pakiti koos rakendusega MAP Docker Image'iks nimega `pneumonia_app` ja konfigureeriti eralditöövoona failis `pneumonia-wf.json` (vt lisa 3). Rakendus konfigureeriti otsima CR-modaalsuseid ehk röntgeneid ning töövoogu konfiguratsiooni `task_destinations` alla täpsustati tingimus, et DICOM kujutisel peab olema uuritud kehaosa silt (0018,0015) [5] väärtusega CHEST (rindkere).

4.3.3 Järeldamine testkujutistel

Kõigepealt katsetati mudeli töötamist RSNA Pneumonia Detection Challenge¹ andmekogu paaril pildil, et näha, kuidas tulemusi kuvatakse. Kuna Orthanc ei toetanud sellise väljundi kuvamist, kasutati tulemuse visualiseerimiseks MicroDicom² rakendust. Joonis 20 kujutab klassifitseerimise väljundit, mille välja narratiivne kokkuvõte (*Narrative Summary*) väärtus *Pneumonia* ütleb, et pildil on kopsupõletik.

Patient: ea872d0d-d42c-4330-9c09-799bc3b0ab09 (male, #ea872d0d-d42c-4330-9c09-799bc3b0ab09)
Series: Pneumonia evaluation (#5178)
Manufacturer: MONAI Deploy App SDK (DICOM SR Writer)
Completion Flag: PARTIAL
Verification Flag: UNVERIFIED
Content Date/Time: 2022-12-26 20:47:49

Diagnostic Imaging Report

Narrative Summary:
Pneumonia

Joonis 20. PneumoniaMNIST kopsupõletiku klassifitseerimise rakenduse väljund

4.4 Mudeli treenimise ja järeldamise töövoogude prototüüp

MedMNIST mitme kopsuhaiguse andmekogu näitel

Olles veendunud lihtsama klassifikatsioonimudeli kasutatavuses, treeniti seejärel mudel 14 erineva kopsuhaiguse klassifitseerimiseks. Valmis mudel pakendati MAP rakenduseks, konfigureeriti MONAI Deploy Expressis ja kasutati testkujutistel järelduste tegemiseks.

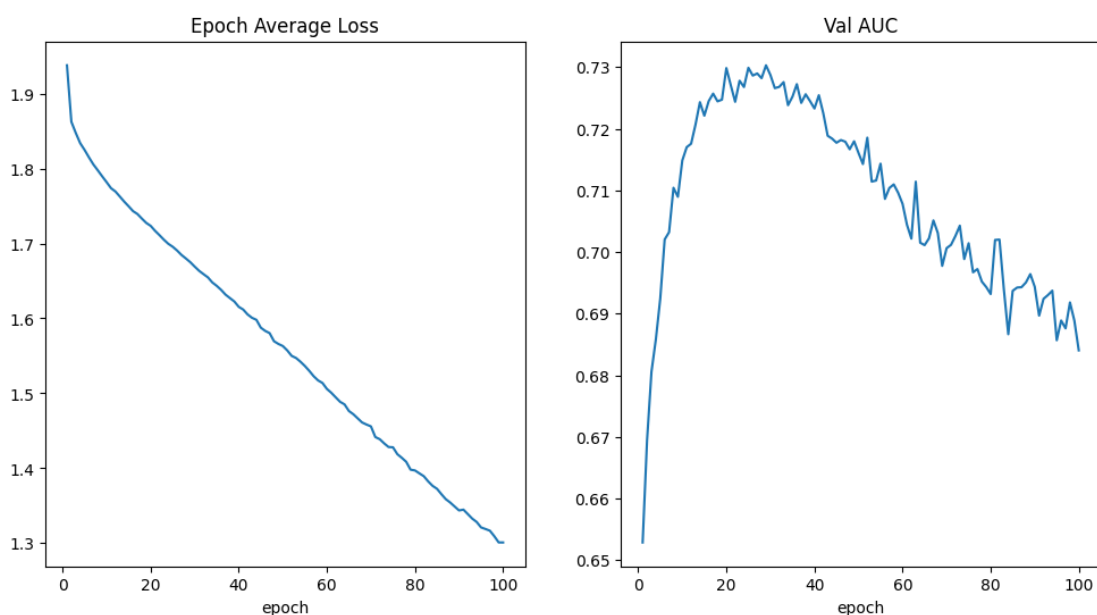
4.4.1 Mudeli treenimine MedMNIST ChestMNIST andmekogu põhjal

Sarnaselt eelnevas alampeatükis kirjeldatule treeniti seekord 14 erinevat kopsuhaigust tuvastav mudel. Treenimiseks kasutati MedMNIST [21] [22] kopsuröntgenitel põhinevat kopsupõletiku andmekogu ChestMNIST [24]. Kogu sisaldas endas 112 120 kujutist

¹ <https://www.rsna.org/education/ai-resources-and-training/ai-image-challenge/rsna-pneumonia-detection-challenge-2018>

² <https://www.microdicom.com/>

erinevate leidudega kopsudest, kusjuures ühel pildil võis olla mitu või mitte ühtegi leidu. Andmekogu kujutistel tuli skaleerida intensiivsus, lisaks suurendada mõõtudele 64x64, kuna algselt olid nad suurusega 28x28 pikslit. Seejärel rakendati andmete rikastamise funktsioone, mis sisaldasid endas suvalisi keeramisi, pööramisi ning suurendamisi ja vähendamisi. DenseNet121 mudelit trenitati 100 epohhi, saavutades parima ROC-kõvera aluse pindala väärtusega 0.7303 29. epohhil (Joonis 21), mis on võrreldav MedMNIST enda tehtud võrdlusuuringute¹ parima väärtusega 0.778.



Joonis 21. ChestMNIST kopsuhaiguseid klassifitseeriva mudeli trennimise meetrika

Mudeli trennimise vihik on kättesaadaval lisas 4 viidatud lingil. Tabel 2 kajastab mudeli erinevate haiguste tuvastamise täpsust andmekogu testhulga peal.

Tabel 2. ChestMNIST kopsuhaiguste klassifitseerimise mudeli täpsuse klassifikatsiooniraport

	precision	recall	f1-score	support
atelectasis	0.2052	0.0880	0.1232	2420
cardiomegaly	0.2547	0.0704	0.1104	582
effusion	0.2638	0.2923	0.2773	2754
infiltration	0.2433	0.2115	0.2263	3938
mass	0.2092	0.0600	0.0933	1133
nodule	0.3333	0.0007	0.0015	1335
pneumonia	0.0000	0.0000	0.0000	242
pneumothorax	0.1817	0.1111	0.1379	1089

¹ <https://medmnist.com/>

consolidation	0.0000	0.0000	0.0000	957
edema	0.1282	0.0121	0.0221	413
emphysema	0.1034	0.0118	0.0212	509
fibrosis	0.0000	0.0000	0.0000	362
pleural	0.3333	0.0014	0.0027	734
hernia	0.0000	0.0000	0.0000	42
none	0.5300	0.8737	0.6597	11928
accuracy			0.4401	28438
macroavg	0.1857	0.1155	0.1117	28438
weightedavg	0.3475	0.4401	0.3575	28438







4.4.2 Pakendamine rakenduseks

Analoogselt kopsupõletiku klassifitseerijaga loodi mudeli kasutamiseks rakendus `ai_chest_classifier_app`. See rakendus väljastas hinnangu, kas ja milline kopsuhaigus võib kujutisel olla. Lisaks konfigureeriti rakendus nii, et väljundil oleksid kõikidele haigustele antud hinnangud. Tekstilise väljundi esitamine realiseeriti taaskord MONAI klassi `DICOMTextSRWriterOperator` abil. Mudel pakiti koos rakendusega MAP Docker Image'iks nimega `chest_classifier` ja konfigureeriti eralditöövoona failis `chest-classification-wf.json` (vt lisa 4). Rakendus konfigureeriti otsima CR-modaalsuseid ehk röntgeneid ning töövoogu konfiguratsiooni `task_destinations` alla täpsustati tingimus, et DICOM kujutisel peab olema uuritud kehaosa silt (0018,0015) [5] väärtusega CHEST (rindkere).

4.4.3 Järeldamine testkujutistel

Kõigepealt katsetati mudeli töötamist sama RSNA Pneumonia Detection Challenge¹ andmekogu paaril pildil, et näha, kuidas tulemusi kuvatakse. Mudeli järeldus lisati eraldi seeriana uuringu alla (Joonis 22).

¹ <https://www.rsna.org/education/ai-resources-and-training/ai-image-challenge/rsna-pneumonia-detection-challenge-2018>

Saturday, December 31, 2022	
Chest X-Ray evaluation Status: <i>Unknown</i> Modality: SR BodyPartExamined: CHEST SeriesInstanceUID: 1.2.826.0.1.3680043.8.498.93953010523860371799441454117347096608 SeriesNumber: 5403	 
Pneumonia evaluation Status: <i>Unknown</i> Modality: SR BodyPartExamined: CHEST SeriesInstanceUID: 1.2.826.0.1.3680043.8.498.50679441572157572209986943273088297727 SeriesNumber: 4402	 
?	
view: AP Status: <i>Unknown</i> Modality: CR BodyPartExamined: CHEST SeriesInstanceUID: 1.2.276.0.7230010.3.1.3.8323329.1488.1517874291.205429 SeriesNumber: 1	 

Joonis 22. Mudeli väljund uue seeriana

Tulemuse visualiseerimiseks kasutati taaskord MicroDicom¹ rakendust. Joonis 23 kujutab klassifitseerimise väljundit, mille välja narratiivne kokkuvõte (*Narrative Summary*) väärtus *predicedresult: none* ütleb, et pildilt ei suudetud kopsupõletikku tuvastada.

Diagnostic Imaging Report

Narrative Summary:
 Predicted result:: none
 atelectasis:: 0.46762994
 cardiomegaly:: -2.0958745
 effusion:: 0.5455918
 infiltration:: 1.3054401
 mass:: -2.5430744
 nodule:: -2.7591002
 pneumonia:: -4.301213
 pneumothorax:: -2.3293078
 consolidation:: 0.5200945
 edema:: -3.8921022
 emphysema:: -1.9232446
 fibrosis:: -7.388486
 pleural:: -3.4501626
 hernia:: -10.787163
 none:: 2.147748

Joonis 23. Kopsuhaiguste klassifitseerimise väljund

¹ <https://www.microdicom.com/>

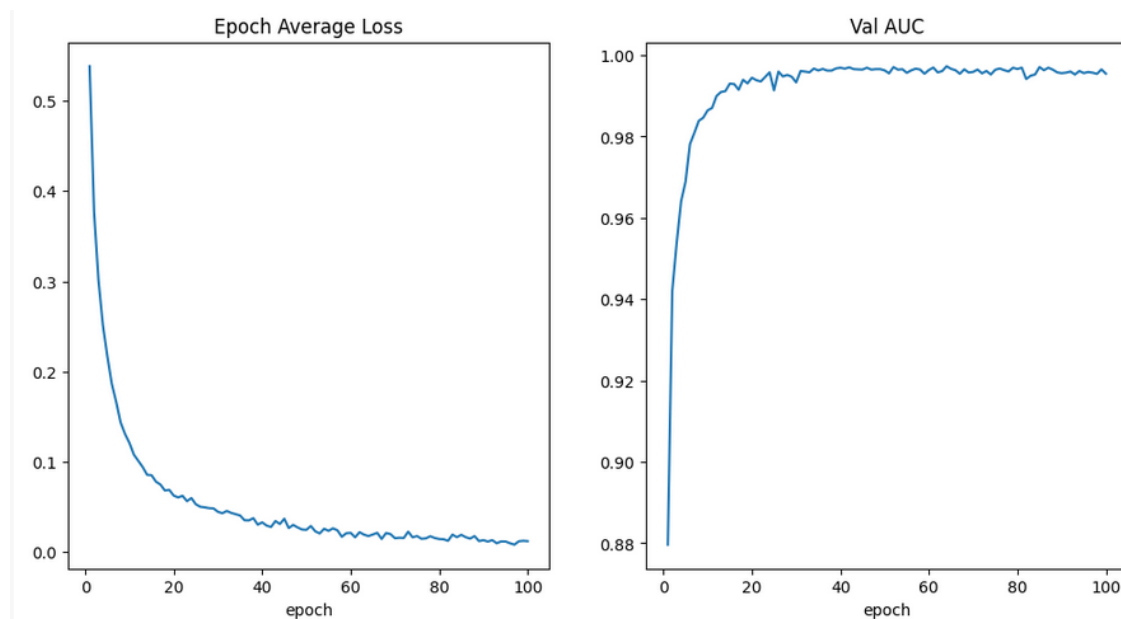
4.5 Mudeli treenimise ja järeldamise töövoogude prototüüp

ChestXRay2017 andmekogu näitel

Seejärel prooviti MONAIga läbi mudeli treenimise töövoog ChestXRay2017 andmekogu põhjal. Valmis mudel pakendati MAP rakenduseks, konfigureeriti MONAI Deploy Expressis ja kasutati testkujutistel järelduste tegemiseks.

4.5.1 Mudeli treenimine ChestXRay2017 andmekogu põhjal

Kõigepealt treeniti analoogselt juhendile¹ kopsupõletikku tuvastav mudel. Treenimiseks kasutati kopsuröntgenitel põhinevat andmekogu ChestXRay2017 [25], mis sisaldas 5863 tervet ja põletikuga kopsude kujutist. Kolme või nelja kanaliga kujutised teisendati ühele, kujutistel skaleeriti intensiivsus ning töödeldi ühte suurusele 256x256 pikslit. Andmete rikastamiseks kasutati suvalisi pööramisi ja keeramisi. DenseNet121 mudelit treeniti 100 epohhi, saavutades parima ROC-kõvera aluse pindala väärtusega 0.9955 64. epohhil.



Joonis 24. ChestXRay2017 kopsupõletikku klassifitseeriva mudeli treenimise meetrika

Mudeli treenimise vihk on kättesaadaval lisas 5 viidatud lingil. Tabel 3 kajastab mudeli täpsust andmekogu testhulga peal.

¹ https://docs.monai.io/projects/monai-deploy-app-sdk/en/latest/notebooks/tutorials/02_mednist_app.html

Tabel 3. ChestXRay2017 kopsupõletiku klassifitseerimise mudeli täpsuse klassifikatsiooniraport

	precision	recall	f1-score	support
none	0.9767	0.5385	0.6942	234
pneumonia	0.7818	0.9923	0.8746	390
accuracy			0.8221	624
macro avg	0.8793	0.7654	0.7844	624
weighted avg	0.8549	0.8221	0.8069	624

4.5.2 Pakendamine rakenduseks

Mudeli kasutamiseks loodi rakendus `ai_pneumonia2_classifier_app`. Sarnaselt eelnevate näidetega väljastas klassifitseerimise rakendus hinnangu, kas kujutisel leidis põletik või mitte. Mudel pakiti koos rakendusega MAP Docker Image'iks nimega `pneumonia2_app` ja konfigureeriti eralditöövoona failis `pneumonia2-wf.json` (vt lisa 5). Rakendus konfigureeriti otsima CR-modaalsuseid ehk röntgeneid ning töövoogu konfiguratsiooni `task_destinations` alla täpsustati tingimus, et DICOM kujutisel peab olema uuritud kehaosa silt (0018,0015) [5] väärtusega CHEST (rindkere).

4.6 Kopsupõletiku ja -haiguste klassifitseerimise mudelite testimine

Eestis tehtud röntgenite peal

Seejärel katsetati eelnevaid kopsuröntgenite mudeleid Eestis tehtud uuringute peal. Selleks hankis käesoleva töö kaasjuhendaja 20 kujutist ilma leiuta ning 19 kujutist kopsupõletikuga, mis läbisid anonümiseerimise. Anonümiseeritud kujutised laeti üles serverisse ning saadeti järelduste tegemisele. Tabel 4 kajastab MedNMIST andmekogudel treenitud mudelite hinnanguid. Esimene kopsupõletiku klassifitseerimise mudel PneumoniaMNIST kogu põhjal andis igale kujutisele sama hinnangu. Kopsuhaiguste klassifitseerimise mudel ChestMNIST põhjal ei tuvastanud kordagi kopsupõletiku olemasolu seda tegelikult sisaldavatel kujutistel. Õigete hinnangute näitaja $10/39 = 0.256$ ei ületanud juhuslikkust.

Tabel 4. MedNMIST andmekogudel treenitud mudelite järeldused Eesti röntgenkujutistel

Leiuta kujutised			Kopsupõletikuga kujutised		
Nimi	Kopsupõletiku mudeli hinnang	Kopsuhaiguste mudeli hinnang	Nimi	Kopsupõletiku mudeli hinnang	Kopsuhaiguste mudeli hinnang
N1	Pneumonia	None	P1	Pneumonia	Infiltration
N2	Pneumonia	None	P2	Pneumonia	None
N3	Pneumonia	None	P3	Pneumonia	Emphysema
N4	Pneumonia	None	P4	Pneumonia	None
N5	Pneumonia	None	P5	Pneumonia	None
N6	Pneumonia	None	P6	Pneumonia	None
N7	Pneumonia	Emphysema	P7	Pneumonia	None
N8	Pneumonia	Emphysema	P8	Pneumonia	None
N9	Pneumonia	None	P9	Pneumonia	None
N10	Pneumonia	Infiltration	P10	Pneumonia	None
N11	Pneumonia	None	P11	Pneumonia	None
N12	Pneumonia	Infiltration	P12	Pneumonia	Infiltration
N13	Pneumonia	None	P13	Pneumonia	Pneumothorax
N14	Pneumonia	Emphysema	P14	Pneumonia	None
N15	Pneumonia	None	P15	Pneumonia	None
N16	Pneumonia	Infiltration	P16	Pneumonia	Emphysema
N17	Pneumonia	Infiltration	P17	Pneumonia	None
N18	Pneumonia	Emphysema	P18	Pneumonia	None
N19	Pneumonia	Emphysema	P19	Pneumonia	None
N20	Pneumonia	Pneumothorax			

4.6.1 Järeldamised ChestXRy2017 andmekogul põhineval mudelil enne ja pärast ekslike hinnangute parandamist

ChestXRy2017 andmekogul treenitud mudel hindas leiuta kujutisi õigesti 10 korral 20st, see-eest kopsupõletikuga kujutisi 6 korral 19st. Seega õigete hinnangute suhtarv $16/39 = 0.410$ jäi antud testandmete koguga siiski alla juhuslikkuse piiri 0.5.

Seejärel sooviti näha, kas valesti hinnatud röntgenkujutiste kasutamine mudeli treenimisel suudab ülejäänud kujutiste hindamise täpsust parandada. Selleks võeti valede hinnangutega 5 leiuta ja 5 põletikuga kujutist, lisati treeningandmete hulka õige sildiga ja treeniti mudel uuesti (vt lisa 5). Võrdlusena kasutatav täiendamata andmekogu suhtarv oli $16/29 = 0.552$.

Täiendatud mudel andis rohkem valepositiivseid hinnanguid kujutistel, mida treenimisel ei kasutatud – kui enne oli hinnati leiuta kujutisi õigesti 10/15 korral, siis täiendatud mudelil ainult 6/15 korra. See-eest kopsupõletiku tuvastamine leiuga kujutistel paranes

mõnevõrra – enne täiendamist hinnati täiendtreenimisel mittekasutatud kujutisi õigesti 6/14 kordadest; pärast täiendamist hinnati õigesti 9/14 kordadest. Üldine täpsuse suhtarv 15/29 jäi siiski alla enne täiend-treenimist saavutatud täpsusele 16/29.

Tabel 5 kajastab esialgse ja täiendatud mudeli antud hinnanguid igale kasutatud kujutisele.

Tabel 5. ChestXRay2017 andmekogul treenitud mudelite järelused Eesti röntgenkujutistel

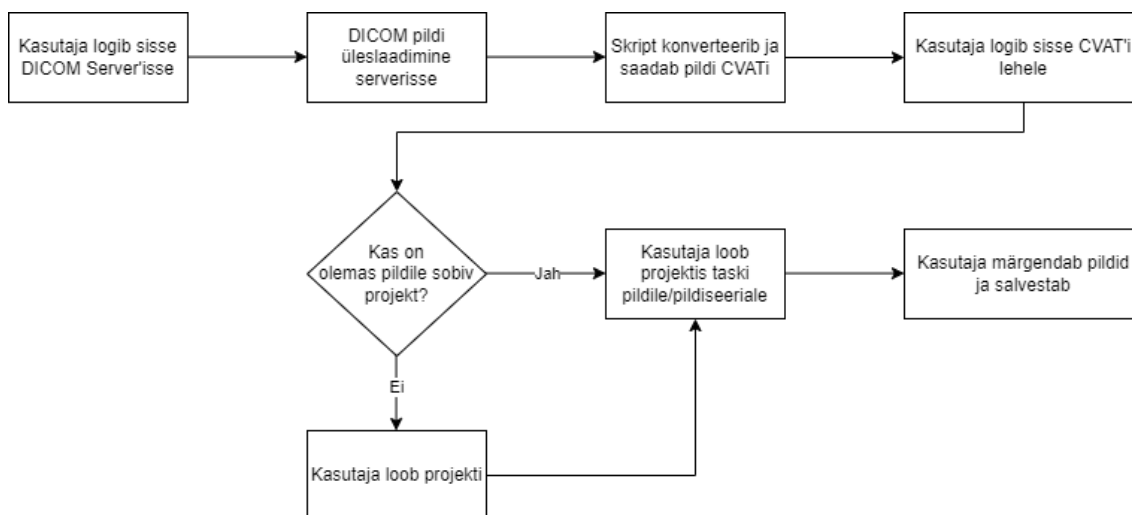
Leiuta kujutised			Kopsupõletikuga kujutised		
Nimi	Esialgse mudeli hinnang	Täiendatud mudeli hinnang	Nimi	Esialgse mudeli hinnang	Täiendatud mudeli hinnang
N1	None	Pneumonia	P1	Pneumonia	Pneumonia
N2	None	None	P2	None	-
N3	None	None	P3	Pneumonia	Pneumonia
N4	None	Pneumonia	P4	None	-
N5	None	None	P5	None	-
N6	None	Pneumonia	P6	None	-
N7	Pneumonia	-	P7	None	-
N8	Pneumonia	-	P8	None	None
N9	Pneumonia	-	P9	None	None
N10	Pneumonia	-	P10	None	None
N11	None	None	P11	None	None
N12	None	Pneumonia	P12	None	None
N13	None	None	P13	Pneumonia	Pneumonia
N14	Pneumonia	-	P14	None	Pneumonia
N15	Pneumonia	None	P15	None	Pneumonia
N16	None	Pneumonia	P16	Pneumonia	Pneumonia
N17	Pneumonia	Pneumonia	P17	None	Pneumonia
N18	Pneumonia	Pneumonia	P18	Pneumonia	Pneumonia
N19	Pneumonia	Pneumonia	P19	Pneumonia	Pneumonia
N20	Pneumonia	Pneumonia			

5 Alternatiivne prototüüp

Käesoleva lõputöö algfaasis prooviti kõigepealt ise valida välja erinevad komponendid, tööriistad ja teegid märgendamise- ja järeldamisprotsesside jaoks. Seejärel avastati alternatiivne tarkvara (kirjeldatud peatükis 3), mis töötas märkimisväärselt lihtsustada protsesside ülesseadmist, ning esimene prototüüp jäi kõrvale. Järgnev peatükk kaardistab esimese prototüübi realiseeritud sammud ja nende käigus saadud teadmised.

5.1 Töövoog

Esimest katsetatud variandi märgendamise töövoogu kujutab Joonis 25. Andmete salvestamine toetus Microsoft DICOM Serverile. Kasutaja logis veebilehele, kust sai üles laadida ja ning vaadelda DICOM pilte. Seejärel päris konverteriteenus DICOMid, mis konverteeris need NiFTi failideks, et saata need mudelile, ning png piltideks, mis saadeti CVATi (*Computer Vision Annotation Tool*, interaktiivne video- ja pilditöötlemise tööriist tehisenägemises [26]). Seal tuli määratleda või luua sobiv projekt märgendamisteks ning see järel sai alustada piltide segmenteerimisega.



Joonis 25. Esialgselt realiseeritud märgendamise töövoog

Realiseeritud märgendamise protsessist oli puudu veel CVATi märgenduste eksportimine NIFTI maskiks, mille abil koos NIFTI kujutisega treenida nnU-Netil põhinevat mudelit. Seejärel oleks olnud järgmine samm üles seada GPU (*Graphics processing unit*, graafikaprotsessor) võimekusega server, mille peal treenida nnU-Neti mudel.

Pärast märgenduste tegemist ja mudeli treenimist oleks järgnenud järeldamise protsessi ülesseadmine. See oleks endast kujutanud sama DICOM serveri ja nnU-Neti serveri omavahelist liidestamist pseudonümiseeritud kujutiste hindamiseks.

5.2 Microsoft DICOM Server andmete hoiustamiseks

DICOM serverina katsetati Microsofti rakendust *Medical Imaging Server for Dicom* (meditsiiniline pildiserver DICOMile, alampeatükis edaspidi server). See on avatud lähtekoodiga DICOM server, mida on lihtne kasutusele võtta Azure'is (Microsofti pilveplatvorm) [27]. Server on .Net Core implementatsioon DICOMweb standardist.

Server on eelkõige loodud Azure'il töötamiseks, aga käesoleva töö raames kasutati võimekust kohalikuks kasutuselevõtuks [28]. Selleks kasutati Docker Compose'i, mis võimaldas kõik vajalikud sõltuvused automaatselt käivitada, vajamata arendusmasinas muid paigaldamisi. Eraldi konteiner käivitatakse SQL Serverile ja Azurite'ile, mis emuleerib Azure Storage'it. Rõhutatakse et lokaalne Azure Storage emuleerimine on arendus- ja test-stsenaariumites kasutamiseks ning mittesobilik tootmis-stsenaariumites. Seda eelkõige, kuna paroolid on arendusfailides, SQL server pole krüpteeritud, autentimine on maha võetud ning andmed ei säili konteinereid taaskäivitades.

Serveril on lisaks veel võimekus päringule *changefeed*, mis pakub logisid kõikidest muudatustest DICOM teenuses [29]. Logi on järjestatud ja kirjutuskaitstud ning võimaldab käia teenuses läbi loomiste ja kustutamiste ajaloo.

Serverit oli võrdlemisi lihtne üles seada, kui kasutajal oli eelnev kokkupuude Dockeriga. Microsoft DICOM Server on kindlasti kaalumist väärt alternatiiv, kui näiteks kasutada seda koos Azure'i pilveteenustega, kuid see paistis olevat eelkõige API-na kasutamiseks mõeldud. Lokaalsel kujul sai küll sisse lülitada OHIF-vaaturi, kuid see jäi mugavuse poolest alla MONAI pakutud Orthancile, mis pakkus võimalust otse liidesest üles laadida faile, anonümiseerida neid ja välja saata analüüsile.

5.3 CVAT piltide märgendamiseks

Märgendamiskeskonnana katsetati CVATi [26], mis on tasuta ja avatud lähtekoodiga videote ja piltide annoteerimise tööriist. CVAT valiti käesoleva töö raames katsetamiseks oma kasutusmugavuse tõttu. CVAT võtab sisse tavalisemaid pildiformaate, DICOM kujutisi ta vaikumisi ei toeta.

Käesoleva tööga seati üles konverter, mis põhines CVAT lähtekoodis olevale DICOM failide konverteril¹. Modifitseeritud skript kasutas ära Microsoft DICOM Serveri *changefeed* päringut. Lühikeste intervallidega päriti uusi muudatusi, sisemiselt arvel pidades töödeldud muudatuste arvu. Kui muudatuse kirjel oli tegevuseks „loomine“ (*Create*) ja staatus polnud „kustutatud“ (*Deleted*), päriti fail identifikaatoreid kasutades ning teostati konverteerimine png'ks. Kuna DICOM failid võivad sisaldada mitmeid pilte, struktureeriti pildid uuringute (*studies*), seeriade (*series*) ja instantside (*instances*) põhiselt kaustadesse.

Skript toetas 8- ja 16-bitiste DICOM piltide konverteerimist, lisati juurde tugi ka 12-bitistele piltidele. Tulevikku ettevaatavalt lisati samale konverterile ka tugi NiFTi-deks teisendamiseks, mida vajab nnU-Net mudeli treenimisel. Konverter seati püsi Dockeri konteinerina ja regulaarseid päringuid tehti *cron*'i abil. Konverteri lähtekood on kättesaadav lisa 6 oleva lingiga.

CVAT pakkus samuti valmis Dockeri konteinerit, mida ka rakendati. Seda aga modifitseeriti, et anda ligipääs konverteri piltide kaustale. Selleks tuli compose failis spetsifitseerida *volume* nimega *cvat_share*, et kaust ilmuks CVATis ülesannet luues *Connected file share*'i all (Joonis 26).

¹ Kättesaadav https://github.com/opencv/cvat/tree/develop/utils/dicom_converter, katsetati juhiseiga lingil <https://community.intel.com/t5/Blogs/Tech-Innovation/Artificial-Intelligence-AI/Medical-Data-DICOM-Annotation-in-Computer-Vision-Tool/post/1335754>

Basic configuration

* Name

Kopsupatsient nr 1

Project

Kopsud

Subset

Train

Labels

Project labels will be used

* Select files

My computer

Connected file share

Remote sources

Cloud Storage

root

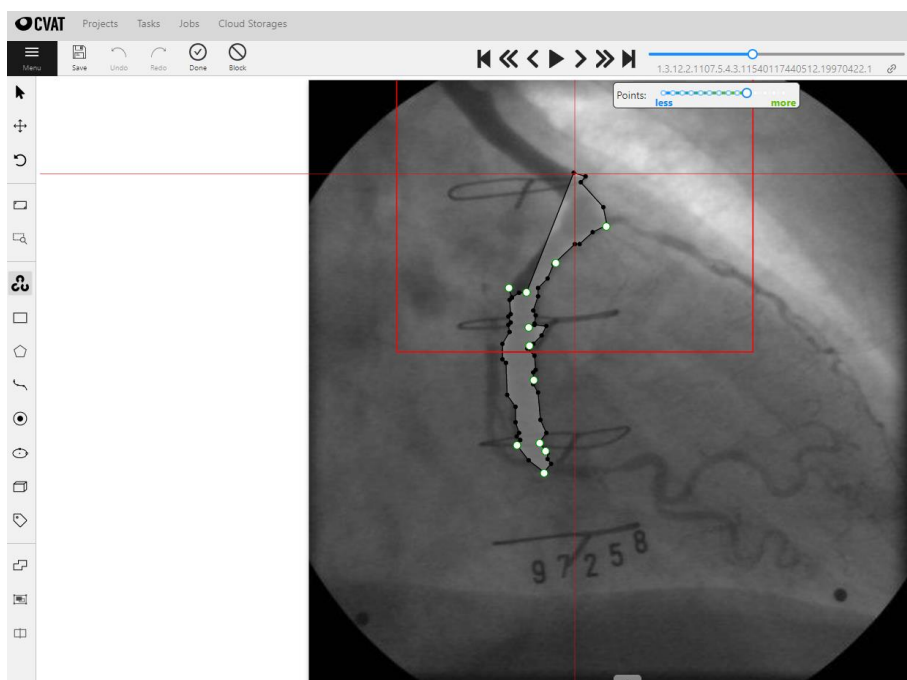
1.3.12.2.1107.5.4.3.11540117440512.19970422.140030.45

1.3.12.2.1107.5.4.3.11540117440512.19970422.140030.46

1.3.12.2.1107.5.4.3.4975316777216.19951114.94101.16

Joonis 26. Ülesande loomine CVATis

CVAT võimaldab märgendamisel kasutada mitmeid mugavaid märgendamistööriistaid, lihtsamatest näiteks polügoone ja ristkülikuid. Võimekus on ka tehisintellekti rakendavaid vahendeid kasutada. Eriti kasulik on näiteks OpenCV, mis punktide põhjal suudab automaatselt täpsustada segmentatsiooni piire (Joonis 27).



Joonis 27. Märgendamine OpenCVga

CVATiga märgendamine oli väga mugav OpenCV abil, kuid see mugavus piirdus 2D kujutistega. 3D seeriatega läks protsess pikemaks ja tüütumaks. MONAI pakutud märgendamistööriistades seas oli muuhulgas CVAT, kuid seal oli ka alternatiive, mis pakkusid mugavamat ja kiiremat kogemust eriti 3D kujutiste annoteerimisel.

5.4 nnU-Net mudelite treenimiseks ja järelduste tegemiseks

Enne MONAI avastamist plaaniti kasutada nnU-Neti mudeleid. nnU-Net on esimene segmentatsioonimeetod, mis on disainitud hakkama saama valdkonnas leiduva andmekogu varieeruvusega [30]. Teek lubab standardiseeritud alustõde, et hõlbustada teadlastel algoritmide võrdlemist nnU-Netiga, karbist-välja automaatselt seadistuvat segmenteerimismeetodit eelkõige 3D kujutistele ning raamistikku segmenteerimismeetodite kiireks arendamiseks. Lisaks pakub sama organisatsioon ka teeki nnDetection [31]. Selle eesmärk on süstematiseerida ja automatiseerida meditsiiniliste objektide tuvastamise seadistamisprotsessi.

Üks tulevase uurimisküsimusi oleks, kuidas nnDetection mudeleid integreerida ka MONAI töövoogu.

6 Arutelu

Järgnevas peatükis analüüsitakse käesoleva tööga pakutud automaatanalüüsi protsessi ning uuritud võimalikke prototüüpe. Seejärel loetletakse võimalikud probleemid ning viimasena kaardistatakse ettepanekud edasisteks arendusteks.

6.1 Automaatanalüüsi tegemise protsess

Automaatanalüüsi protsesside kavandamisse oli kliendina kaasatud käesoleva töö kaasjuhendaja, kellel oli selleks ajaks aastaid kogemusi radioloogina ning alates 2022. aastast töötas Ida-Tallinna Keskhaiglas radioloogiakeskuse juhatajana. Koos kaasjuhendajaga kaardistati vajadused, mille põhjal plaaniti pakutavad protsessid ja nõuded. Need said omakorda valideeritud ja tagasisidestatud kaasjuhendaja poolt.

Mudeli täpsuse kohta tagasiside saamine võib osutuda keeruliseks. Tõenäoliselt on pakutud variantidest parim rakendada lihtsat tagasisidevormi, sundimata radiologe ise parandusmärgendusi tegema. Sisuliselt järgneks neli stsenaariumi:

1. Mudel tegi õige järelduse, kuid radioloog ei anna tagasisidet. See pole kriitilise tähtsusega.
2. Mudel tegi vale järelduse, kuid radioloog ei anna tagasisidet. See on problemaatilisem, sest jääb parandamata ebatäpsus mudelis.
3. Mudel tegi vale järelduse ning radioloog annab selle kohta tagasisidet. Sel juhul tuleb selgelt välja tuua, et vigaselt märgendatud kujutis saadetakse koos kommentaariga üle vaatamisele ning kasutatakse vajadusel anonümiseeritud kujul mudeli ületreenimiseks.
4. Mudel tegi õige järelduse ning radioloog annab vastava tagasiside. Sel juhul mudelit üle treenima ei pea ning tagasiside läheb statistikasse.

6.2 Realiseeritud prototüüp

Pakutud protsessi prototüübi realiseerimist alustati eri tükide üles seadmisega – vaja oli DICOM serverit, märgendamiskeskonda, treenitavaid ja järeltõlke pakkuvaid mudeleid koos oma raamistikuga. Lisaks oli vaja need kõik omavahel liidestada – DICOM failid oli vaja konverteerida märgendamiskeskonnale sobivasse formaati ja sinna saata; märgendused oli vaja konverteerida mudelitel kasutatavasse NiFTI formaati ning lõpuks tulemus visualiseerida ja seejärel konverteerida tagasi seeriaks DICOM kujutisele. Individuaalsete serverite ülesseamine ja nende omavaheline liidestamine osutus oodatust keerulisemaks.

6.2.1 MONAI

Jooksvalt käesoleva töö jaoks kirjanduspäringuid tehes avastati uudne MONAI, mille rakendused tundusid paljulubavad, võimaldades kiiret ja lihtsat ülesseadmist. Seejuures töötas raamistik vähendada vajadust eelnevas lõigus kirjeldatud eraldiseisvate serverite seadistamistele ja liidestamistele. Rakendused olid ka pidevas arenduses. 2022. aasta novembris ilmunud artikkel [32] ütles, et NVIDIA toetatud raamistikku oli alla laetud juba üle 650 tuhande korra ning oli kasutusel näiteks Cincinnati lastehaiglas alates sama aasta septembrist. Lisaks on MONAI-põhine tehisintellekti juurutusmootori platvorm AIDE (*AI Deployment Engine platform*) leidnud kasutust Ühendkuningriikide Riiklikus tervishoiuteenistuses, pakkuades tehisintellektil põhinevaid haiguse tuvastamise vahendeid neljale haiglale, teenindades 5 miljonit patsienti aastas.

MONAI ülesseadmine oli tõesti kohati oluliselt lihtsam esimesest lahenduskäigust. Dokumentatsioonides kirjeldatud näidised märgendamiskeskonnast¹, MAPi loomisest ja sisestamisest töövoogu² ning Deploy Expressi käivitamisest³ läksid tööle mõistliku ajaga. Küll aga andis tunda, et rakendused olid pidevas arenduses – nii mõnigi kord juhtus, et rakendusi seadistades kulutati üleliigset aega vigade parandamisele, mille kohta mõni

¹ <https://docs.monai.io/projects/label/en/latest/quickstart.html>

² https://docs.monai.io/projects/monai-deploy-app-sdk/en/latest/getting_started/tutorials/monai_bundle_app.html

³ <https://github.com/Project-MONAI/monai-deploy/tree/main/deploy/monai-deploy-express>

nädal hiljem märgati, et olid kas *git*-is kasutajate poolt vigadena üles loetletud või juhendis ära parandatud.

MONAI pakutav näidismudelite kogum Model Zoo oli küll järjest täienev, kuid siiski veel päris pisike, millest suurem osa oli loodud organite segmenteerimiseks. Näidistega analoogselt teiste mudelite liidestamine töövoogu osutus kohati päris keeruliseks, näiteks ei õnnestunud mõistliku ajaga tööle saada kopsusõlme tuvastamise mudelit (*Lung nodule ct detection*), kasutatav rakendus, mis standardsete segmentatsioonide asemel pakkus ümbritsevaid kaste (*bounding-box*) koos tõenäosustega sõlme olemasolust. Lisaks oli MONAI sihitud eelkõige ka 3D kujutiste mudelitele. MONAI rakendused ise olid kõrge potentsiaaliga, kui käesoleva töö tegemise käigus toored. Seda ilmestasid ka mitmes sammus märgatud hoiatused, et mudel või rakendus ei sobinud kliiniliseks kasutamiseks.

6.2.2 Mudelite sooritusvõime

Käesoleva tööga treenitud mudelid ei saavutanud erakordset täpsust ning ei läbinud piisavat valideerimist kliiniliseks kasutamiseks. Eriti ebatäpne oli treenitud 14 erineva kopsupõletikku tuvastav rakendus. Katsetades treenitud mudeleid Eestis tehtud röntgenite peal, osutus kõigi treenitud mudelite täpsus olemaks alla juhuslikkuse. Mudeli täpsust ei suudetud parandada ka pärast väikse hulga valesti järeldatud piltide lisamist andmekogule ületreenimiseks, tõenäoliselt liiga väikse koguse testkujutiste tõttu. Kuigi valim 39 kujutise näol oli liiga väike ammendava järelduse tegemiseks, vihjasid tehtud eksperimendid sellele, et treenitud mudelid vajavad täiendavat täpsustamist ja kohalike andmete abil kohandamist konkreetseesse konteksti.

Töö fookus oli eelkõige valideerida prototüübi töövoogude toimimist ja kasutajasõbralikkust, seega ei hakatud käsitsi treenitud mudeleid uuesti treenima teiste konfiguratsioonidega. Edaspidi peaks keskenduma konkreetsetele mudelitele ning neid kohandama kohalike andmete abil täpsuseni, mis on piisav lahenduse juurutamiseks haigla töövoogudesse.

6.3 Edasised arendused

Käesoleva tööga jäid uurimata mõned teemad, mida soovitatakse käsitleda tehisintellekti rakendamise protsessi edasises arendamises. Läbi jäi proovimata põhjalikumalt MONAI Label abil andmete segmenteerimine ja selle põhjal mudeli treenimine. See oleks

nõudnud hulganisti anonümiseeritud DICOM kujutisi, mida mitme eksperdi abil märgendada. Sadade või isegi tuhandete sobiva uuringu kujutiste saamine Eesti meditsiinisüsteemist ning nende märgendamine töötab olla üpris ajakulukas protsess, mida edasiste arenduste käigus katsetada. Peatükis 3.3 kirjeldatud MONAI Label võimekuste ja käesoleva töö autori poolt põgusalt tehtud katsetuste põhjal on põhjust arvata, et rakendus suudab ise lahendada ära probleemid erinevate märgendamisviiside kombineerimise näol.

6.3.1 Potentsiaalsed probleemid mudelite kasutamisel

Lisaks oli plaanis ka katsetada pideva õppimise töövoogu, treenides olemasolevaid mudeleid üle paarikümne märgendatud kujutise põhjal. Kuigi see võiks nõuda ekspertidelt aega märgendatud andmete kogumiseks, oleksid siinkohal uurimisküsimused, kuidas treenida mudeleid üle nii, et ta eelnevaid tähtsaid parameetreid ära ei unustaks. Näiteks kirjeldatakse ühes artiklis [33] kirjeldatakse lähenemist dünaamilist mälu kasutades, et pidevalt õppides adapteerida kujutiste loomise seadmete ja tarkvara arengust tuleneva kujutiste muutumisega ajas. Samal ajal üritatakse takistada mudeli täpsuse langemist varasematel kujutistel katastroofilise unustamise (*catastrophic forgetting*) näol. Kasutatud tehnika tasakaalustas mälu, tuvastades pseudo-domeene, mis esindasid andmevoos erinevaid stiili-klastreid.

Tehisintellektiga on kaasnenud mõned eetilised katsumused [34]. Algoritmidega on esile kerkinud probleem, et nad võivad otsuste tegemises peegeldada inimlikke eelarvamusi. Kuna tervishoiuteenuste osutamine varieerub etniliselt, võib juhtuda, et mõned etnilised eelarvamused ehitatakse algoritmidesse. Mittemeditsiinilised algoritmid on juba välja toonud probleeme eelarvamuslike hinnangutega. Näiteks kohtuotsuseid teha aitavad programmid, mis hindavad rikkuja retsidiivsuse ohtu, on näidanud kalduvust rassilisele diskrimineerimisele [35].

Sarnased rassilised eelarvamused võivad esile kerkida tervishoiu algoritmides, kui geneetilise uuringuid tehakse vaid kindlates rahvahulkades.

6.3.2 Seletatav mudel

Tehisintellekti võimekus on märkimisväärselt kasvanud, tõstatades samaaegselt küsimusi miljonite parameetritega sügavate võrkude mudelite hinnangute usaldusväärsuse suhtes [36]. Vajadus on kasvanud mudelitele, mille puhul on võimalik otsustusi selgitada.

Süvavõrgud on viimastel aastatel näidanud vähkkasvaja tuvastamisel radioloogidega konkureerivat täpsust, kuid nende suurimaks probleemiks on selgitamatus, s.t. otsuse puhul ei saa kasutaja teada, miks võrk otsuse tegi. See tekitab usalduse ja läbipaistvuse puudumist sellistesse musta-kasti-algoritmidesse.

Seletatavus on meditsiinis lisaks sisulistele põhjustele vajalik ka juriidilistel ja eetilistel põhjustel, näiteks Isikuandmete kaitse üldmäärus sätestab, et patsientide andmeid klassifikatsioonideks ja soovitusteks kasutavad organisatsioonid peavad pakkuma nendele soovi korral selgitusi [37].

Käesoleva töö raames seletavat mudelit ei uuritud, kuid see pakutakse välja tulevaseks uuringupunktiks. Näiteks pakub käesoleva töö ajal värskest ilmunud raamat¹ [38] lahendusi masinõppe mudelite otsuste selgituste saamiseks.

6.3.3 Homomorfne krüpteerimine

Masinõppe ja mudelite treenimine nõuab suurtes kogustes meditsiinilisi andmeid. Eetilised nõuded ja privaatsusregulatsioonid aga piiravad nende vaba jagamist. Krüpteerimistehnikad nagu täielik homomorfne krüpteerimine (FHE, *fully homomorphic encryption*) võimaldavad kasutada süvaõppemudeleid krüpteeritud andmetel [39]. Andmetel on võimalik teha operatsioone hoolimata sellest, et nad on krüpteeritud; samuti pole vaja ka dekrüpteerimisvõtit.

Privaatust säilitavad masinõppe mudelid on näiteks [39]:

- Privaatne väljastpoolt tellitud arvustus (*private outsourced computation*) – Meditsiiniuasutus kasutab homomorfset krüpteerimist privaatset väliste arvutuste tegemiseks pilveserveris.
- Privaatne prognoos (*private prediction*) – klient klassifitseerib oma andmeid treenitud mudelil ilma otsese ligipääsuta mudelile, andmata mudeli omanikule olulist teavet oma andmete kohta. Selle jaoks võib kasutada vahendavat salvestust ja arvutusi pilves

¹ Kättesaadav ka e-raamatuna <https://christophm.github.io/interpretable-ml-book/>

- Privaatne treenimine (*private training*) – pilveserver treenib mudelit kliendi krüpteeritud andmetel.

Krüpteerimise rakendamiseks on olemas mitmeid teke eri raamistikele [39]. C++ keelele on näiteks SEAL (*Microsoft's Simple Encrypted Arithmetic Library*), HELib (*IBM's Homomorphic Encryption Library*), TFHE (*Fast Fully Homomorphic Encryption over the Torus*) ja PALISADE. Pythoni peal on olemas Pyfhel (*PYthon For Homomorphic Encryption Libraries*), mis toetud SEAL, HELib ja PALISADE teekidele.

Käesoleva tööga soovitatakse protsessi edasiste arendustel uurida ja rakendada homomorfset krüpteerimist mudeli treenimisel ja hinnangute saamisel. Samas küsimused, missugused mudelite otsuste seletatavuse lähenemised jäävad homomorfset krüpteeritud mudelite korral toimima jäävad samuti edasise uurimise teemaks.

7 Kokkuvõte

Järjest enam digitaliseeritud meditsiinisüsteem on avanud uusi võimalusi meditsiiniteenuste parendamiseks. Käesoleva tööga uuriti võimalusi rakendada kaasaegseid tehismõistuse lahendustele tuginevaid tehnoloogiaid radioloogiliste kujutiste analüüsiks. Koostöös radioloogia valdkonna eksperdiga kavandati ja valideeriti protsessid täiendava automaatanalüüsi rakendamiseks radioloogide töövoogu, seejuures hoides fookust sellel, kuidas minimaliseerida täiendavat ajakulu radioloogi tööle. Lisaks pakuti välja protsessid nende analüüsides parandamiseks. Protsessid tuli seejuures hoida kooskõlas Isikuandmete kaitse üldmäärusega.

Pakutud protsessidel põhinev prototüüp prooviti kõigepealt ise kokku panna erinevatest tarkvaradest, seahulgas Microsoft DICOM Server, CVAT ja nnU-Net mudelist, kuid seejärel leiti sobivam alternatiiv. Tarkvara MONAI lubas pakkuda kogu lahendust paketina, kus oli võimalik pakendada mudeleid lihtsasti kasutatavaks rakendusteks. MONAI raamistikku kasutades prooviti trenida mõned mudelid kopsuröntgeni baasil ning neid kasutada rakendustena järeldamise protsessis koos mõne eeltreenitud mudeliga MONAI enda poolt. Kopsupõletiku klassifitseerimise mudeleid katsetati anonümiseeritud Eestis tehtud röntgenikujutiste peal ja ühte mudelit prooviti parandada väikse hulga valesi hinnatud kujutiste lisamise läbi treeningandmete hulka. 39 Eestis tehtud röntgenuurinuga läbi viidud eksperimendi käigus ei suudetud mudeli täpsust parandada, kuid edasised eksperimendid oleksid vajalikud olulised suurema andmekoguga.

Järeldamise ja trenimise protsessid osutusid sobivaks ja piisavalt mugavaks, kuigi mudelid ise vajasisid täpsuse parandamist ja täiendavat valideerimist enne kliinilisse kasutusse sobivust. Rakenduste loomise protseduur osutus algul keeruliseks piiratud dokumentatsiooni ja rakenduse uudsuse tõttu – nii mõnelgi korral juhtus, et juhendite ja tarkvara parandusi teostati paar nädalat pärast nende katsetamist käesoleva töö autori poolt. Samas osutus iga järgneva eksperimendi tegemine lihtsamaks võrreldes eelmisega.

Viimaks kaardistati ära potentsiaalsed probleemid ja edasised arendused. Käesoleva tööga ei jõutud põhjalikult proovida märgendamisprotsessi, kuid välja pakutud MONAI Label näitas potentsiaali nõuetele sobivuseks. Lisaks käsitleti potentsiaalset probleemi mudelite täpsuse langemist ületreenimisega katastroofilise unustamise näol. Rõhutati ka vajadust seletatavatele mudelite, mis peaksid radioloogile ja patsiendile põhjendama oma järelduste tagamaid. Viimaks käsitleti probleemi seoses suures koguses meditsiiniliste andmete hoiustamisega ja sellest lähtuvat vajadust homomorfsele krüpteerimisele.

Kasutatud kirjandus

- [1] Riigikogu, „Isikuandmete kaitse seadus,“ Riigi Teataja, 15 jaanuar 2019. [Võrgumaterjal]. Available: <https://www.riigiteataja.ee/akt/104012019011>. [Kasutatud 13 detsember 2022].
- [2] P. Ross ja R. Mill, „Digitaalsete dokumentide jagamise standardprofiilid tervishoius,“ *Eesti Arst*, p. 8, 31 10 2013.
- [3] „MONAI,“ MONAI, [Võrgumaterjal]. Available: <https://monai.io/about.html>. [Kasutatud 9 12 2022].
- [4] Y. W. Kim ja L. T. Mansfield, „Fool Me Twice: Delayed Diagnoses in Radiology With Emphasis on Perpetuated Errors,“ *American Journal of Roentgenology*, nr 202, pp. 465-470, 2014.
- [5] Innolitics, „Body Part Examined Attribute,“ Innolitics, [Võrgumaterjal]. Available: <https://dicom.innolitics.com/ciods/cr-image/general-series/00180015>. [Kasutatud 15 dets 2022].
- [6] J. Rumbold ja B. Pierschionek, „The Effect of the General Data Protection Regulation on Medical Research,“ *Journal of Medical Internet Research*, kd. 19, nr 2, 2017.
- [7] L. Sweeney, „Simple Demographics Often Identify People Uniquely,“ 2000. [Võrgumaterjal]. Available: <https://dataprivacylab.org/projects/identifiability/paper1.pdf>. [Kasutatud 6 11 2022].
- [8] EUR-Lex, „Isikuandmete kaitse üldmäärus (GDPR),“ Euroopa Liidu Väljaannete Talitus, 7 1 2022. [Võrgumaterjal]. Available: https://eur-lex.europa.eu/legal-content/ET/TXT/HTML/?uri=LEGISSUM:310401_2. [Kasutatud 6 11 2022].
- [9] Euroopa Komisjon, „Mis on isikuandmed?,“ Euroopa Komisjoni veebisait, [Võrgumaterjal]. Available: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_et. [Kasutatud 6 11 2022].
- [10] EUR-Lex, „Euroopa Parlamendi ja nõukogu määrus (EL) 2017/745,“ 5 aprill 2017. [Võrgumaterjal]. Available: <https://eur-lex.europa.eu/eli/reg/2017/745/2020-04-24>. [Kasutatud 6 11 2022].
- [11] DICOM, „DICOMweb™,“ The Medical Imaging Technology Association (MITA), [Võrgumaterjal]. Available: <https://www.dicomstandard.org/using/dicomweb>. [Kasutatud 8 11 2022].
- [12] L. Gupta, „What is REST,“ 7 aprill 2022. [Võrgumaterjal]. Available: <https://restfulapi.net/>. [Kasutatud 9 11 2022].
- [13] DICOM, „Search (QIDO-RS),“ The Medical Imaging Technology Association (MITA), [Võrgumaterjal]. Available: <https://www.dicomstandard.org/using/dicomweb/query-qido-rs>. [Kasutatud 9 11 2022].
- [14] DICOM, „Worklist (UPS-RS),“ The Medical Imaging Technology Association (MITA), [Võrgumaterjal]. Available:

- <https://www.dicomstandard.org/using/dicomweb/workflow-ups-rs>. [Kasutatud 8 11 2022].
- [15] DICOM, „Capabilities,“ The Medical Imaging Technology Association (MITA), [Võrgumaterjal]. Available: <https://www.dicomstandard.org/using/dicomweb/capabilities>. [Kasutatud 9 11 2022].
- [16] S. Jodogne, „The Orthanc Ecosystem for Medical Imaging,“ *J Digit Imaging*, nr 31, pp. 341-352, 2018.
- [17] A. Diaz-Pinto, S. Alle, A. Ihsani, M. Asad, V. Nath, F. Pérez-García, P. Mehta, W. Li, H. R. Roth, T. Vercauteren, D. Xu, P. Dogra ja S. Ourselin, „MONAI Label: A framework for AI-assisted Interactive Labeling of 3D Medical Images,“ 23 Märts 2022. [Võrgumaterjal]. Available: <https://arxiv.org/abs/2203.12362>. [Kasutatud 4 11 2022].
- [18] MONAI Consortium, „MONAI,“ 2022. [Võrgumaterjal]. Available: <https://docs.monai.io/projects/label/en/latest/index.html>. [Kasutatud 6 11 2022].
- [19] NEMA, „DICOM,“ NEMA, 7 4 2022. [Võrgumaterjal]. Available: <https://dicom.nema.org/dicom/dicomwsi/>. [Kasutatud 6 11 2022].
- [20] M. J. Cardoso, W. Li, R. Brown, N. Ma, E. Kerfoot, Y. Wang, B. Murrey, A. Myronenko, C. Zhao, D. Yang, V. Nath, Y. He, Z. Xu, A. Hatamizadeh, A. Myronenko, W. Zhu, Y. Liu ja M. Zheng, „MONAI: An open-source framework for deep learning in healthcare,“ 4 november 2022. [Võrgumaterjal]. Available: <https://arxiv.org/abs/2211.02701>. [Kasutatud 9 detsember 2022].
- [21] J. Yang, R. Shi, D. Wei, Z. Liu, L. Zhao, B. Ke, H. Pfister ja B. Ni, „MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical Image Classification,“ *arXiv preprint arXiv:2110.14795*, 2021.
- [22] J. Yang, R. Shi ja B. Ni, „MedMNIST Classification Decathlon: A Lightweight AutoML Benchmark for Medical Image Analysis,“ *IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 191-195, 2021.
- [23] D. S. Kermany, M. Goldbaum ja e. al., „Identifying medical diagnoses and treatable diseases by image-based deep learning,“ *Cell*, kd. 172, nr 5, pp. 1122-1131, 2018.
- [24] X. Wang, Y. Peng ja e. al., „Chest x-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases,“ *CVPR*, pp. 3462-3471, 2017.
- [25] D. Kermany, K. Zhang ja M. Goldbaum, „Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification,“ 2018. [Võrgumaterjal]. Available: <https://data.mendeley.com/datasets/rscbjbr9sj/2>. [Kasutatud 8 jaanuar 2023].
- [26] CVAT, „CVAT,“ 2022. [Võrgumaterjal]. Available: <https://opencv.github.io/cvat/docs/>. [Kasutatud 6 11 2022].
- [27] Microsoft Corporation, „Medical Imaging Server for DICOM,“ Microsoft Corporation, [Võrgumaterjal]. Available: <https://github.com/microsoft/dicom-server#readme>. [Kasutatud 9 11 2022].
- [28] Microsoft Corporation, „Deploy the Medical Imaging Server for DICOM locally using Docker,“ Microsoft Corporation, [Võrgumaterjal]. Available: <https://github.com/microsoft/dicom-server/blob/main/docs/quickstarts/deploy-via-docker.md>. [Kasutatud 9 11 2022].

- [29] M. Mitrik, J. Degnan, M. Weaver, S. Wohl, A. Soy ja G. Lee, „Change Feed Overview,“ Microsoft, 15 10 2022. [Võrgumaterjal]. Available: <https://learn.microsoft.com/en-us/azure/healthcare-apis/dicom/dicom-change-feed-overview>. [Kasutatud 16 november 2022].
- [30] F. Isensee, P. Jaeger ja S. e. a. Kohl, „nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation,“ *Nat Methods*, kd. 18, pp. 203-211, 2021.
- [31] M. Baumgartner, P. F. Jäger, F. Isensee ja K. H. Maier-Hein, „nnDetection: A Self-configuring Method for Medical Object Detection,“ *Medical Image Computing and Computer Assisted Intervention*, Springer International Publishing, 2021, pp. 530-539.
- [32] D. Niewolny, „MAP Once, Run Anywhere: MONAI Introduces Framework for Deploying Medical Imaging AI Apps,“ nvidia, 28 11 2022. [Võrgumaterjal]. Available: <https://blogs.nvidia.com/blog/2022/11/28/monai-deploy-framework-medical-imaging-ai-apps/>. [Kasutatud 28 12 2022].
- [33] M. Perkonigg, J. Hofmanninger, C. J. Herold, J. A. Brink, O. Pianykh, H. Prosch ja G. Langs, „Dynamic memory to alleviate catastrophic forgetting in continual learning with medical imaging,“ *Nature Communications*, kd. 12, nr 5678 , 2021.
- [34] F. Pesapane, C. Volonté ja M. e. a. Codari, „Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States,“ *Insights Imaging*, pp. 745-753, 15 august 2018.
- [35] J. Angwin, J. Larson, S. Mattu ja L. Kirchner, „Machine bias,“ ProPublica, 23 mai 2016. [Võrgumaterjal]. Available: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. [Kasutatud 6 11 2022].
- [36] C. Schorr, P. Goodarzi, F. Chen ja T. Dahmen, „Neuroscope: An Explainable AI Toolbox for Semantic Segmentation and Image Classification of Convolutional Neural Nets,“ *applied sciences*, p. 16, 3 märts 2021.
- [37] M. Gulum, T. C.M. ja M. Kantardzic, „A Review of Explainable Deep Learning Cancer Detection Models in Medical Imaging,“ *Appl. Sci*, p. 11, 17 05 2021.
- [38] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, Independently published, 2022.
- [39] A. Wood, K. Najarian ja D. Kahrobaei, „Homomorphic Encryption for Machine Learning in Medicine and Bioinformatics,“ *ACM Comput. Surv.*, kd. 53, nr 4, p. Article No.: 70pp 1–35, 2020.
- [40] N. H. Strickland, „PACS (picture archiving and communication systems): filmless radiology,“ *Archives of disease in childhood*, kd. 83, nr 1, pp. 82-86, 2000.

Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹

Mina, Martin Joonas Pariis

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Radioloogiliste kujutiste masinõppele tugineva automaatanalüüsi juurutamine radioloogide töövoogu“, mille juhendaja on Juhan-Peep Ernits ja kaasjuhendaja Julius Juurmaa
 - 1.1. reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2. üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.
3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

08.01.2023

¹ Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtjaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.

Lisa 2 – Eesnäärme segmenteerimise rakenduse failid

Rakenduse kood ja töövoog konfiguratsioon asuvad GitLab repositooriumis.

https://gitlab.com/mart1nj/master-thesis/-/tree/main/prostate_app

Lisa 3 – PneumoniaMNIST kopsupõletiku mudeli treenimiseks ja kasutamiseks loodud failid

Vihik mudeli treenimiseks, kood rakenduse pakendamiseks ja töövoogu konfiguratsioon asuvad GitLab repositooriumis.

https://gitlab.com/mart1nj/master-thesis/-/tree/main/pneumonia_app

Lisa 4 – ChestMNIST mitme kopsuhaiguse klassifitseerimise mudeli treenimiseks ja kasutamiseks loodud failid

Vihik mudeli treenimiseks, kood rakenduse pakendamiseks ja töövoogi konfiguratsioon asuvad GitLab repositooriumis.

https://gitlab.com/mart1nj/master-thesis/-/tree/main/chest_app

Lisa 5 – ChestXRay2017 kopsupõletiku mudeli treenimiseks ja kasutamiseks loodud failid

Vihikud mudeli esimeseks treenimiseks ja treenimiseks koos täiendandmetega, kood rakenduse pakendamiseks ja töövoogu konfiguratsioon asuvad GitLab repositooriumis.

https://gitlab.com/mart1nj/master-thesis/-/tree/main/pneumonia2_app

Lisa 6 – DICOM konverteerija png ja NifTI failideks

DICOM konverteerija lähtekood ja juhised asuvad GitLab repositooriumis, ligipääsetav lingiga <https://gitlab.com/mart1nj/master-thesis/-/tree/main/dicom-converter>.