

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond

Kristjan Luik 211809IAPM

Metsaraie tuvastamine rakendades arvutinägemise alusmudeleid

Magistritöö

Juhendaja: Juhan-Peep Ernits

PhD

Tallinn 2025

Autorideklaratsioon

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

Autor: Kristjan Luik

26.05.2025

Annotatsioon

Metsavarade jätkusuutlikuks majandamiseks on hädavajalik metsaraiete tuvastamise täpsus ja ajakohasus. Traditsioonilised statistilised meetodid Sentinel-2 multispektraalsete piltide töötlemisel on osutunud tõhusaks, kuid kannatavad sageli piiratud ruumilise eraldusvõime tõttu ja vajavad lisaks palju manuaalset tööd. Käesolev uurimus käsitleb isejuhitud Vision Transformer'i peaahelat DINOv2 ja SAM v2 väheste õppeandmete põhiseid semantilisi segmentatsioonisid lageraie sündmuste tuvastamiseks.

Esmalt loodi andmestik, mis koosneb viieteistkümnest lageraie polügoonist ja nende ümbritsevatest metsapiirkondadest. Nende põhjal ehitati programm, mis töötleb ja analüüsib metsateatisi, et luua geomeetrlised maskid, mis eristavad okas- ja lehtpuudega alasid. Lisaks sellele laeb programm alla Sentinel-2 taseme 2A satelliidi pildid, et luua andmestik, mis sisaldab nii lageraie maske kui ka nende ümbritsevaid metsapiirkondi.

Tulemuste hindamiseks võrdleme DINOv2-põhist raamistikku SAM v2 ja U-Neti segmentatsionimudelitega, kasutades peamiste kvaliteedimõõdikutena IoU-d ja F1 skoori. Avatud lõpptulemite analüüs võimaldab hinnata tuvastustäpsuse, märgendamistõhususe ja mudeli robustsuse paranemist erinevates metsakoostistes. Edasistes uuringutes käsitletakse ajarealise analüüsi ja täiustatud pilvekatte maskimise (*masking*) strateegiaid.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 45 leheküljel, 5 peatükki, 23 joonist, 8 tabelit.

Abstract

Logging detection using vision foundation models

Monitoring and accurately detecting forest logging activities is essential for sustainable forest management and environmental conservation. Traditional statistical approaches, such as Random Forest models applied to Sentinel-2 imagery, have shown promise but still suffer from limited spatial precision and require extensive manual post-processing. In this thesis, we explore the efficacy of a self-supervised Vision Transformer backbone, DINOv2, for few-shot semantic segmentation of logging events in Sentinel-2 multispectral images.

A smaller dataset comprising 15 clear-cut polygons and their adjacent forest environs was constructed by integrating publicly available metsateatis records from the Estonian Forest Registry with Sentinel-2 Level-2A surface reflectance tiles. Initial geometric masks were refined through manual delineation and K-Means clustering to differentiate coniferous and deciduous strata. The pretrained DINOv2 model was subsequently fine-tuned on this dataset, utilizing the 10 m spatial resolution spectral bands alongside derived vegetation indices to enable pixel-level discrimination of logging areas.

To evaluate performance, we compare the DINOv2-based framework against SAMv2 and U-Net segmentation models, using Intersection over Union (IoU) and F1 score as primary metrics. This open-ended analysis will assess relative improvements in detection accuracy, annotation efficiency, and robustness to varying forest compositions. Further investigation will address temporal sequence incorporation and advanced cloud-masking strategies.

The thesis is in Estonian and contains 45 pages of text, 5 chapters, 23 figures, 8 tables.

Lühendite ja mõistete sõnastik

Ülelennu sagedus	(<i>Revisit time</i>) ajavahemik, mis jäääb mingi kindla piirkonna satelliidi divaatluste vahele
Laine pikkus peaahel	(<i>Band</i>) laine pikkuste vahemik elektromagnetkiirguse spektris (<i>Backbone</i>) mudeli peamine arhitektuur, mis on eelnevalt treenitud ja millele on lisatud täiendavad kihid, et saavutada soovitud ülesanne
R2	R-ruut (<i>R-squared</i>) regressioonimudeli täpsuse mõõdik, mis näitab mudeli selgitusvõimet andmete variatsioonis
IoU	Ühenduse indeks (<i>Intersection over Union</i>) mõõdik, mis hindab mudeli täpsust, võrreldes ennustatud ja tegelikke tulemusi
F1 skoor	F1 skoor (<i>F1 score</i>) mõõdik, mis ühendab täpsuse ja meeldetuletuse ja annab tasakaalustatud hinnangu mudeli jõudlusele
NDVI	Taimede indeksi (<i>Normalized Difference Vegetation Index</i>) mõõdik, mis hindab taimekatte tihedust ja elujõudlust, arvutatakse punase ja lähedase infrapunase laine pikkuse vahekorra põhjal
Jääkühik	Jääkühikud (<i>residual blocks</i>) on närvivõrgu arhitektuurimuster, kus konvolutsioonikihide jadale lisatakse sisendi ja väljundi vahe (residuaal) skip-ühenduse kaudu, et hõlbustada sügavate võrkude treenimist ja vältida gradientide kadu.
DICE	DICE (<i>Dice coefficient</i>) on mõõdik, mis hindab segmentatsiooni
CrossEntropyLoss	Rist-entsentratsiooni kadu (<i>Cross Entropy Loss</i>) on tüüpiline kadufunksioon klassifitseerimisülesannetesse, mis mõõdab erinevust mudeli ennustatud töenäosuste ja tegelike klasside vahel.
Ülesobitamine	Ülesobitamine (<i>overfitting</i>) on nähtus, kus mudel õpib treeningandmeid liiga hästi, sealhulgas müra ja ebaolulisi mustreid, mis viib halvenenud üldistamisvõimele uute andmete puhul. See võib põhjustada kõrgeid treeningtulemusi, kuid madalaid valideerimistulemusi.

Sisukord

1	Sissejuhatus.....	9
2	Valdkonna ülevaade	11
2.1	Metsandus.....	11
2.2	Copernicus ja EstHub.....	11
2.2.1	Sentinel.....	12
2.2.2	Lainepikkuste spekter	13
2.2.3	Koordinaatsüsteemid ja CRS	14
2.3	Masinõppe meetodite kasutus kaugseires	14
3	Lahendus	16
3.1	Töövahendid.....	16
3.2	Andmestiku loomine	18
3.2.1	Raie piirkonna andmete kogumine.....	18
3.2.2	Raie piirkonna maskide loomine	25
3.3	Alusmudeli ülevaade	27
3.3.1	DINO v2 ja selle kasutus.....	28
3.3.2	SAM v2 ja selle kasutus	29
3.4	Seoste leidmine	31
3.5	Treenimis protsetuurid	36
4	Tulemuste analüüs	46
4.1	Tulemuste võrdlus	46
4.2	Edasiarendus ja täiustamine.....	52
5	Kokkuvõte.....	53
	Kasutatud kirjandus	54
	Lisa 1 – Lihtlitsents lõputöö reproduutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks.....	57
	Lisa 2 – Vektor maskide loomine kasutades QGISi	58

Jooniste loetelu

Joonis 1. Sateliidi andmete liikumine andmekeskuste vahel	12
Joonis 2. Näidis ühe lageraie päringust saadud ümbrus ja satelliitpilt	20
Joonis 3. KMeans klasterdamise tulemus, 100 raie ümbrust leht- või okaspuudega ..	21
Joonis 4. Andmestiku loomise töövoog.....	24
Joonis 5. Sentinel-2 RGB näidis	26
Joonis 6. RGB ortofoto näidis.....	27
Joonis 7. CIR-NGR ortofoto näidis	27
Joonis 8. Näidis sateliidi pilt	32
Joonis 9. Raie piirkonna mask sateliidi pildil.....	33
Joonis 10. Raie piirkonna mask eksperdi poolt korrigeeritud.....	34
Joonis 11. Segmenteeritud näidis sateliidi pildist.....	35
Joonis 12. T-SNE kluster analüüs DinoV2 mudeli väljunditest	36
Joonis 13. ResNet34 jätkühikust skeem	37
Joonis 14. ResNet50 jätkühikust skeem	38
Joonis 15. MobileNetV2 jätkühikust skeem.....	38
Joonis 16. DeepLabv3 ASPP moodul	40
Joonis 17. FPN segmentatsioonipea	41
Joonis 18. Mudeli arhitektuuri skeem.....	42
Joonis 19. Parimad DICE tulemused erinevate mudelite vahel.....	47
Joonis 20. Tagatausta ja noore metsa segmentatsiooni Dice tulemused üle katsete	48
Joonis 21. SAMv2 eksperimendi tulemused klasside lõikes.	50
Joonis 22. SAMv2 pareima mudeli segmentatsioonitulemused.....	51
Joonis 23. Metsateatistel põhinev segmentatsiooni tulemus.....	51

Tabelite loetelu

Tabel 1. Sentinel-2 MSI spektriribad ja nende kasutusvaldkonnad.....	13
Tabel 2. NGR indeksi värvि kasutusjuhend.....	14
Tabel 3. Kasutatud riistvara	18
Tabel 4. Hüperparameetrite kombinatsioonid	42
Tabel 5. DINO v2 eksperimenti seadistused.....	44
Tabel 6. SAM v2 eksperimenti seadistused.....	45
Tabel 7. DINOV2 eksperimenti tulemused.....	49
Tabel 8. Samv2 eksperimenti tulemused	50

1 Sissejuhatus

Mets katab 51,5%¹ Eesti pindalast ning metsaraiete järelvalve on üks Keskkonnaagentuuri seireülesannetest. Metsa raiumiseks peab Eestis kehtiva korra järgi metsaomanik esitama Keskkonnaametile vastava metsateatise, mille heakskiitmisel on metsaomanikul vastava raie tegemiseks 2 aastat aega. Keskkonnaagentuuri ülesannete hulka kuulub raiete tuvastamine ja kontroll, kas need raied on tehtud kehtiva metsateatise alusel.

Praegusel hetkel kasutatakse lisaks raiete ortofotodelt ja satelliidipiltidel inimeksperdi abil tuvastamisele ka Eestis 2020.-ndal aastal Keskkonnaagentuuri ja Tartu Ülikooli koostöös väljatöötatud masinõppe mudelit, mis raie tuvastamiseks satelliidipiltidel kasutab otsustusmetsa (Random forest) algoritmi [1]. Selle mudeli esmased tulemused olid paljulubavad, aga jooksvas kasutamise pole see ikkagi rahulda vaid tulemusi andnud ja Keskkonnaagentuuri töötajad on sunnitud siiski pigem tuginema ekspertide arvamusele.

Antud magistritöö põhieesmärgiks on võrrelda kaasaegseid masinõppe meetodeid sarnases rakenduses ja tuua välja täpseim mudel, mis suudaks rahulda täpsusega tuvastada metsaraiet satelliidipiltidel. Mida aeg edasi seda rohkem on riigid hakanud mõistma kui tähtis on metsamajandus, metsade säilitamine ja hoidmine. Tehnoloogia pideva arenguga on hakatud otsima viise kuidas riik või kogukond saaksid paremat ülevaadet suurtest metsaga kaetud aladest. Metsa seireks kasutatakse peamiselt mehitamata õhusõidukeid (Unmanned Aerial Vehicles), maapealseid sensoreid, satelliidipildi töötlust ja vabatahtlike kaasavaid rakendusi (Crowdsourcing Applications) [2].

Euroopa Liidu kaugseireprogramm Copernicus võimaldab Eesti riigil koguda satelliidipilte andmekeskusesse Esthub [3]. Lisaks muule infole, mida hallatakse Copernicus-es ja seeläbi ka Esthub-is, on kasutusel informatsioon, mis tuleb erinevatelt Sentinel nime kandvatelt satelliitidel [4]. Kuna Sentinel-2 on juba 2015. aastast töös olnud, sisaldab laia valikut valgusspektreid ning on tiheda korduskülastussagedusega [5], siis keskendub käesolev

¹<https://www.keskkonnaagentuur.ee/keskkonnaseire-ja-analuusid/mets>

magistritöö peamiselt sellele satelliidi tüübile.

Tööd tehes kujunes välja mõistmine, et automaatselt ei ole võimalik saada alusandmeid. Seetõttu tuli võtta kasutusele täielikult käsitsi märgendatud andmed, mille loomine on väga ajamahukas (ca $3\text{km}^2/\text{h}$). Seoses väikse andmekogumiga, on magistritöös ka keskendutud andmekesksele tehisintellektile (*data-centric AI*). Andmekesksete tehisintellektide on röhk süsteemsel tööl andmetega mudeli keerukuse kasvatamise asemel: hinnatakse andmekogumi kvaliteeti, mitmekesisust ja usaldusväärust.

Andmekogumi loomisel on üheks alam eesmärgiks luua Python programm, mis hõlbus-taks satelliidipiltide allalaadimist ja töötlemist. Peale andmete kogumist on viime läbi tänapäevaste masinõppe mudelite võrdluse raiete tuvastamiseks. Raiet hinnatakse piksli põhise täpsusega üle pildi. Hiljuti on tehtud mitmeid uuringuid selles valdkonnas, kus kasutatakse ka otsustusmetsi, XGBoost ja U-Net'il põhinevaid mudeli arhitektuure [6], [7]. Mõlemas uurimistöös on ka mudelite võrdlus välja toodud, aga need keskenduvad erinevatele suundadele. Esimese puhul ehitatakse mudelid kasutades rohkem pilte läbi aja, et mudel saaks paremini tuvastada muutust. Teise puhul keskendutakse erinevate lainepekkuste kombineerimisele, et tabada muutusi.

Peale mudelite treenimist samadel lähteandmetelt on antud magistritöös välja toodud tulemuste mõõtmise. Piksli tasemel täpsuse mõõtmiseks kasutatakse Intersection over Union - kattuvuse hinnang, Dice Coefficient - meetrika mis on põhimõtteliselt segmenteerimise F1 Score [8], [9]. Nende tulemuste abil saab teha võrdluse erinevate tuvastusmudelite vahel, et leida neist täpseim.

Sellest tulenevalt sai antud magistritöö uurimisküsimusteks:

- Kuidas on võimalik luua automatiseerida satelliidipiltide allalaadimist ja töötlemist ning treeningandmete ettevalmistamist?
- Kui täpselt on võimalik tuvastada lageraie sündmusi satelliidi piltidelt kasutades selleks arvutinägemise alusmudeleid?
- Milline andmekogum on vajalik, et saavutada piisav täpsus lageraie sündmuste tuvastamiseks?

2 Valdkonna ülevaade

2.1 Metsandus

Metsad omavad olulist rolli nii ühiskonna igapäevaelus kui ka planeedi heaolus. Alates mööblis kasutatavast puidust kuni paberini, millele kirjutame. Lisaks neile nähtavatele toodetele sisaldavad paljud ravimid, kosmeetika ja pesuvahendid metsadest saadud kõrvalsaadusi. Rohkem kui 1,6 miljardit inimest sõltub metsadest toidu ja kütuse saamisest ning umbes 70 miljonit, sealhulgas paljud põlisrahvad, peavad metsi oma koduks [10]. Metsad varustavad meid hapnikuga, pakuvad varjualust, töökohti, puhist vett ja toitu, olles seega inimkonna ellujäämiseks hädavajalikud. Kuna nii paljude inimeste elu sõltub metsadest, on metsade saatus otseselt seotud ka meie endi tulevikuga. [11]

2.2 Copernicus ja EstHub

Eesti metsade kaugseires on oluline roll Copernicuse programmil ja EstHubi keskusel. Copernicus on üks osa Euroopa kosmoseprogrammist (EUS), mis tegeleb planeedi jälgimisega. Copernicus programmi raames, lisaks maa pealse info kogumisele, on loodud mitmeid satelliite, mis koguvad informatsiooni kosmosest. See info on kõigile kättesaadav tasuta. Selle programmiga seotud satelliite kutsutakse **Sentineliks**. [12]

EstHub on Eesti riiklik satelliitandmete keskus, mis kogub ja integreerib mitmekesiseid georuumilisi andmeid automatiseritud protsesside kaudu. Andmekogumine hõlmab kõrge resolutsiooniga satelliitkaadrite allalaadimist ja standardiseerimist erinevatest allikatest. EstHubi eesmärk on koguda kokku satelliidi andmed mis katavad Eesti territooriumi. [13]



Joonis 1. Sateliidi andmete liikumine andmekeskuste vahel

2.2.1 Sentinel

Kaugseire valdkonnas on vastavalt vajadusele kasutusel erinevatelt satelliitidelt pärinevad andmed. Sentinel-1 on radaripõhine satelliit, mis võimaldab jälgida maapinna vajumist, struktuuride kahjustusi ning looduskatastroofe nagu maavärinad ja maalihked. Samuti on see ideaalne mere- ja Arktika seireks, sealhulgas laevade jälgimiseks ning naftareostuse tuvastamiseks. [14]

Sentinel-2 missioon koosneb kahest identsest satelliidist, Sentinel-2B (käivitatud 2017) ja Sentinel-2C (käivitatud 2024), mis töötavad koos, et pakkuda kõrge eraldusvõimega multispektraalseid pilte Maa pindadest, rannikualadest ja siseveekogudest iga viie päeva järel. Need andmed toetavad rakendusi põllumajanduses, metsanduses ja maakatte klassifitseerimisel. [15]

Sentinel-3 on Euroopa Maa seire satelliitmissioon, mille eesmärk on mõõta merepinna topograafiat, mere ja maa pinnatemperatuure ning ookeani ja maa pinnavärvi suure täpsusega. Neid andmeid kasutatakse ookeani prognoosisüsteemides, keskkonnaseires ja kliimaseires. [16]

Sentinel-5P on esimene Copernicuse missioon, mis on pühendatud atmosfääri seirele. See kannab tipptasemel **Tropomi** instrumenti, mis kaardistab mitmeid gaase nagu lämmastikdioksiid, osoon, formaldehüüd, väaveldioksiid, metaan, vingugaas ja aerosoolid - kõik need mõjutavad meie hingatavat õhku, tervist ja kliimat. [17]

2.2.2 Lainepekkuste spekter

Spektriribad on satelliitandmete analüüsimeisel üliolulised, sest need võimaldavad eristada maapinna erinevaid omadusi, lähtudes elektromagnetilise spektri konkreetsetest lainepekkustest. Näiteks Sentinel-2 MSI instrumendi 13 spektririba hõlmavad nähtavat valgust, lähedast infrapunat ja lühilaine infrapunat, võimaldades detailset maaстиku klassifitseerimist, sealhulgas metsade, veekogude ja muu loodusliku keskkonna eristamist. Iga ribaga seondub kindel lainepekkuse vahemik, mida spetsiifiliste filtrite abil eraldatakse. [18]

Riba	Resolutsioon	Kasutus
B01	$60m px^{-1}$	Aerosool
B02	$10m px^{-1}$	Sinine
B03	$10m px^{-1}$	Roheline
B04	$10m px^{-1}$	Punane
B05	$20m px^{-1}$	Vegetatsiooni klassifitseerimine
B06	$20m px^{-1}$	Vegetatsiooni klassifitseerimine
B07	$20m px^{-1}$	Vegetatsiooni klassifitseerimine
B08	$10m px^{-1}$	Lähiinfrapunariba on hea rannajoonte ja biomassisalduse kaardistamiseks
B8A	$20m px^{-1}$	Kitsam lähedane infrapunane
B09	$60m px^{-1}$	Veeaur tuvastus
B10	$60m px^{-1}$	Pilvede tuvastus
B11	$20m px^{-1}$	Lühilaine infrapunane 1
B12	$20m px^{-1}$	Lühilaine infrapunane 2

Tabel 1. Sentinel-2 MSI spektriribad ja nende kasutusvaldkonnad

Ribade kombinerimine võimaldab luua erinevaid indeksid, mis aitavad eristada erinevaid maapinna omadusi. Näiteks kasutades lähiinfrapunariba saab luua **NGRi** indeksi, mis põhineb lähi-infrapuna (LIF), rohelise ja punase kanali järjestusel (tähistatakse LRP), kuvab LIF-signaali punases kanalis, rohelise kanal jäab roheliseks ning punane kanal esitatakse sinisena. Selle tulemusel on tervislik taimestik piltidel eredalt punane, kuna rohelised lehed peegeldavad tugevalt lähi-infrapuna spektri osa. Veekogud, mis neelavad nii nähtavat valgust kui ka lähi-infrapuna, ilmuvalt tumedana, mullad ja ehituspinnad, mille peegeldus on mõõdukas kõigis kanalites, võtavad omale hallika või kõikuva tooni. Selline komposiit võimaldab metsanduses efektiivselt hinnata taimede seisundit ja kaardistada metsapiire.

Objekt	Peegelduse omadused	Kuvatav värv
Tervislik taimestik	Kõrge LIF-peegeldus, mõõdukas roheline, madal punane	Ere punane
Kahjustatud või hõredam taimestik	Madalam LIF-peegeldus	Pruunikas või roosakas
Veekogud	Nii nähtav kui lähi-infrapuna valgus neelatakse	Tume sinine kuni must
Paljas muld / linnapinnad	Mõõdukas peegeldus kõigis kvaliteetides	Hallikas, tan või tsüaan

Tabel 2. NGR indeksi värvide kasutusjuhend

2.2.3 Koordinaatsüsteemid ja CRS

Satelliidilt saadud andmete maapinnaga sidumiseks on vaja kasutatada koordinaatsüsteeme ja koordinaatide viite süsteeme (CRS).

Koordinaatsüsteem on meetod, mille abil määratletakse ja kirjeldatakse punktide asukohti maaistikul, kasutades koordinaate. Selles kontekstis eristatakse kahte tüüpi: geograafilised koordinaatsüsteemid, mis kasutavad laiuse ja pikkuse väärtsusi, ning projekteeritud koordinaatsüsteemid, mis teisendavad geograafilised koordinaadid lameda kaardi koordinaatideks, kasutades matemaatilisi projektsioone. CRS ehk koordinaatide viite süsteem määratleb reeglid ja parameetrid, mille alusel need koordinaadid seonduvad realsele maaistikuga. [19]

2.3 Masinõppe metodite kasutus kaugseires

Käsitletava teemaga seotud, kuid teiste piirkondade põhjal loodud, uurimistööde analüüsimal prooviti välja selgitada, milliseid meetodeid on üldiselt kaugseires kasutatud, sealjuures kas on ehitatud mudeliteid nullist või kasutatud valmis mudeliteid. Samuti oli oluliseks eesmärgiks, välja selgitada, kas nendel juhtudel on kasutatud süvaõpet või mitte. Lisaks sooviti teada saada, mis satelliidi andmeid on varasemalt kasutatud ja kas eri lainepekkuste sidumine on andnud paremaid tulemusi.

Ukraina teadlaste poolt koostatud 2021. aastal välja antud artiklis „Deep Learning for Regular Change Detection in Ukrainian Forest Ecosystem With Sentinel-2“ kasutati Copernicus Sentinel-2 satelliidi pilte, mis sisaldasid kõrge resolutsiooniga (10 m) värv- ja spektrikanaleid, sealhulgas NDVI ja NDMI indekseid, võimaldades jälgida metsamuutusi

kuni 5-päevaste intervallidega. Andmekogum loodi käsitsi Kharkivi piirkonnas, kasutades mitut järjestikust pilti ja põhjalikku märgistust, et tagada täpne deforestsatsioonipiirkondade kaardistamine. Urijad rakendasid süvaõppe meetodeid, kasutades mitut U-Neti varianti (näiteks UNet-diff, UNet-CH, UNet2D, UNet3D, Siamese U-Netid ja UNet-LSTM), et hinnata nii ajast sõltuvaid kui ka ühekordseid lähenemisviise. Eraldi rõhutati piltidevahelise erinevuse kasutamise eeliseid, mis parandas segmentatsioonitulemusi ning tõstis Dice ja F1 skoore. Lisaks ilmnesid uuringus olulised nüansid, nagu pilvekatte, hooajaliste muutuste ning geograafiliste lahknevuste mõju, mis nõudsid täiendavat andmete eeltöötlust. Huvitav on, et kuigi kõik mudelid näitasid potentsiaali, saavutavad UNet-diff ja UNet-CH kõige kõrgema täpsuse, pakkudes sealäbi töenduspõhiseid lahendusi metsakatteta ala muutuste regulaarseks jälgimiseks. [6]

Uus-Meremaal läbi viidud ja 2024. aastal välja antud uurimistöös „Developing a forest description from remote sensing: Insights from New Zealand“ kasutati kõrglahutusega lennufotosid ning regionaalseid ALS-andmeid radiata männi metsade täpseks kaardistamiseks Uus-Meremaal. Analüüs tugines sügavõppepõhisel semantilise segmentatsiooni mudelil, mis kasutab DeepLabv3+ arhitektuuri koos ResNext-101 peaelanaga (backbone), saavutades IoU väärtsused 0,94, täpsuse 0,96 ja meeldetuletuse 0,98. Keeruliseks osutus aga noorte istikute tuvastamine, mille puhul *juvenile* klass (noored istutatud metsapiirkonnad) liideti *radiata* (küpsemad männi alad) klassiga. Lisaks sügavõppemudelile kasutati mitmemuutujalisi regressioonimudeleid metsade keskmise kõrguse, kogumahte ja vanuse hindamiseks, saavutades kõrged R2 väärtsused. [20]

3 Lahendus

3.1 Töövahendid

Python

Python on üldotstarbeline programmeerimiskeel, mida kasutatakse laialdaselt andmeteaduse, masinõppe ja ruumiandmete analüüsiga ülesanneteks oma lihtsuse ja mitmekülgsuse tõttu. Käesolevas töös kasutatakse Pythonit andmete töötlemiseks, mudelite treenimiseks ja tulemuste analüüsimiseks.

Jupyter Notebookid

Jupyter Notebookid pakuvad interaktiivset keskkonda, kus saab koodi kirjutada, käivitada ja dokumenteerida ühes kohas. Need võimaldavad dünaamilist andmeanalüüsia ja tulemuste visuaalset esitlust, muutes uurimisprotsessi läbipaistvaks ja korduvaks. Käesolevas töös kasutatakse Jupyter Notebooke peamiselt katsetuste tegemiseks ja tulemuste visualiseerimiseks.

Pandas

Pandas on andmetööluse teek, mis pakub paindlikke ja efektiivseid andmestruktuure tabelipõhise andmetööluse jaoks. See lihtsustab andmete puastamist, analüüsia ja manipuleerimist. Käesolevas töös kasutatakse Pandast andmete lugemiseks, töötlemiseks ja analüüsimiseks.

GeoPandas

GeoPandas laiendab Pandase võimalusi, lisades tuge georuumilistele andmetele. See võimaldab lugeda, analüüsida ja visualiseerida ruumiandmeid ning teostada geomeetrisi operatsioone nagu lõikumine ja ühendamine.

Rasterio

Rasterio on Pythoni teek, mis keskendub rasterandmete lugemisele ja töötlemisele tuginedes GDAL-ile. See võimaldab ruumiandmete analüüsni ning laseb rasterfailidega töötada efektiivselt ja intuitiivselt. Käesolevas töös oli Rasterio peamiseks tööriist rasterandmete lugemiseks ja töötlemiseks, sealhulgas satelliidipiltide koostamiseks ja maskide genereerimiseks.

QGIS

QGIS on tasuta ja avatud lähtekoodiga töölaua GIS-tarkvara, mis võimaldab kasutajatel andmeid visuaalselt analüüsida, redigeerida ja kaardistada. See toetab mitmeid andmeformaate ja pakub laialdasi geoprotsessimise võimalusi, olles populaarne nii akadeemilises kui ka professionaalses keskkonnas. Käesolevas töös kasutatakse QGIS-i peamiselt andmete visualiseerimiseks ja analüüsimiseks, et mõista ruumiandmete struktuuri ja omadusi.

PostGIS

PostGIS on PostgreSQL andmebaasi laiendus, mis lisab ruumiandmete töötlemise funktsionaalsuse. See võimaldab keerukaid ruumi operatsioone ja on oluline tööriist suurte ruumiandmete kogude haldamisel ning analüüsил. Käesolevas töös autor ise PostGIS-i ei kasutanud, kuid Keskkonnaagentuuri andmebaas, kust andmed saadi, on PostGIS-i baasil üles ehitatud. Seega on PostGIS selles uurimistöös oluline komponent andmete haldamisel ja töötlemisel.

Riistvara

Uurimistöö läbiviimisel kasutati ülikooli AI-labori ressursse. Labor koosneb ühest peasõlmest, mis haldab teisi masinaid, ning alamsõlmedest, mis teostavad töid. Autor kasutas ai-lab-07 sõlme CPU-intensiivsete ülesannete, nagu andmekogumi koostamine, piltide töötlemine ja kompressioon, ning mitte sügav närvivõrkude mudelite treenimiseks nagu ResNet34, ResNet50 ning MobileNetV2. Samas süvaõppe eksperimentide teostamiseks kasutati ai-lab-04 sõlme, mille GPU ja mälumaht võimaldasid keerukamate mudelite treenimist. Selline ressursside jaotus aitas töövoogu optimeerida ja tagada tööde sujuva teostuse vastavalt konkreetsetele arvutusvajadustele.

Sõlm	Protsessor	Mälu	GPU	GPU mälu
ai-lab-07	3960X 24-cores/48-threads	128 GB	NVidia 2080Ti	11 GB

Tabel 3. Kasutatud riistvara

3.2 Andmestiku loomine

Selles peatükis käsitletakse andmestiku loomise protsessi, sealhulgas andmete kogumist, töötlemist ja maskimist. Andmestiku loomine on oluline samm igaüiguste andmete analüüsimalisel ja seda ka masinõppe projektide puhul. Andmete kvaliteet ja sobivus mõjutavad otseselt mudeli täpsust ja usaldusväärsust. Nagu muudes valdkondades kehtib ka informaatikas Pareto printsip, mille kohaselt 80% probleemidest tuleneb 20% põhjustest. Seega on andmestiku loomine ja töötlemine äärmiselt oluline etapp, mis võib määrata kogu projekti edasise käigu.

3.2.1 Raie piirkonna andmete kogumine

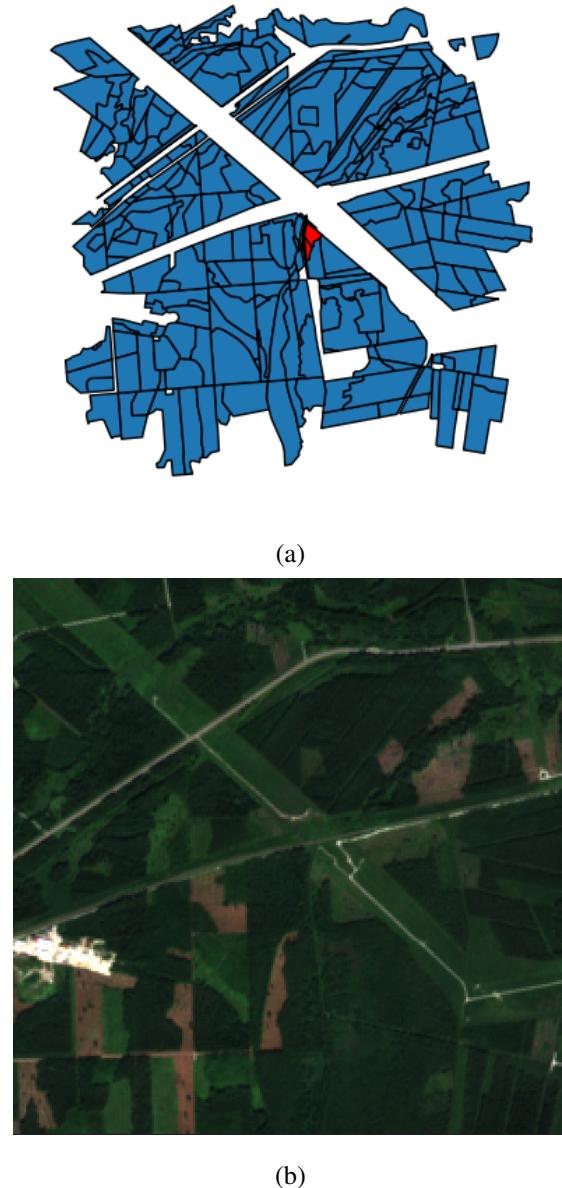
Metsateatis on dokument, mille kaudu metsaomanik esitab Keskkonnaametile kavandatavate raietööde või oluliste metsakahjustuste kohta teabe. Keskkonnaamet kontrollib esitatud teatiste nõuetekohasust ning veendub, et kavandatav raie vastab kehtivatele õigusaktidele. Metsateatised menetletakse ja säilitatakse riiklikus metsaregistris. Peale edukat menetlemist võib raietöödega alustada 10 päeva peale otsust ja kuni 24 kuu jooksul. [21] Metsateatised on avalikud ja neid saab vaadata riiklikus metsaregistris.

Metsade inventeerimise ja registrisse kandmise protsess algab metsaeraldiste täpse kaardistamisega, kasutades L-EST97 ristkoordinaatide süsteemi, Eesti põhiakaarti, katastriüksuse plaane ning vajadusel kaugseire andmeid eraldiste piiritlemiseks ja võimalike situatsioonielementide täpsustamiseks. Kaardistamise tulemusena koostatakse geoinfosüsteemi metsaeraldiste kiht, kus iga eraldis on nummerdatud ning selle pindala, arvutatuna piiripunktide koordinaatide alusel, esitatakse hektarites vähemalt kümnenkohani ning täpsusega 10 meetrit — see loob aluse usaldusväärsele pindalaarvestusele ja edaspidistele takseerimistoimingutele. [22]

Koostöös Keskkonnaametiga (Envir) saadi andmed metsateatistest, mis sisaldavad teavet nii metsateatise esitamise kuupäeva, metsateatise menetlemise kuupäeva, metsateatise keh-

tivuse alguskuupäeva kui ka metsateatise kehtivuse lõppkuupäeva kohta. Kuna riigimetsade teatised on täpsemas seisukorras, siis võeti need raieteatised selle uurimustöö aluseks. Seoses sellega et ühe lõigu peal võib olla väga väike kogus metsa, sai teatiste pärimine ümber ehitatud sedasi, et ühe metsa raie ümber kogutakse peale raie toimumist kokku ka kõik teiste raiete raadiuses asuvad piirkonnad, millel on teada, kas on mets või raielala. Piirkonniti pärimine sai teostatud kasutades PostGISi liidest Postgresi andmebaasiga. Iga raie sisaldab ka endas geomeetria veergu, mis esitab polügooni kujul selle asukohta.

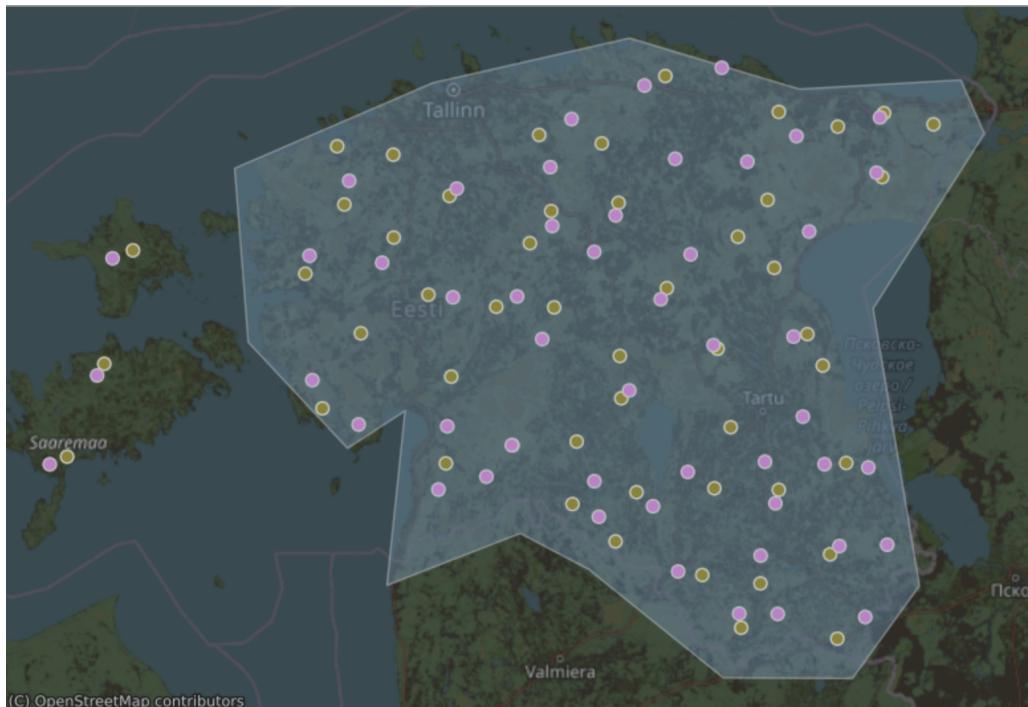
Polügoon ehk hulknurk on geomeetriline kujund, mis määratleb kindla ala, ühendades üksteisega punktid, et moodustada suletud piirjoon. Andmetööluse ja ruumiandmete analüüsni kontekstis kasutatakse polügoone, et täpselt määratleda geograafilisi alasid. [23]



Joonis 2. Näidis ühe lageraie pärингust saadud ümbrus ja satelliitpilt

Magistritöö peamiseks uurimisküsimuseks on, kas ja kuidas on võimalik kasutada vähestest näidete (*Few-Shot*) põhist alusmudelit. Väikese valimi puhul on aga vaja eriti täpseid andmeid millelt õppida. Nagu eelnevalt mainitud siis metsaregistrist saadud andmed ei ole piisavalt täpsed, et neid otse kasutada alusmudeli treenimisel metsaraie tuvastamiseks. Seetõttu lisandusid uurimistöösse andmete täiendava töötlemise ja maskimise etapid, kus andmed käiakse käsitsi läbi, kasutades registri andmeid maski põhjana. Ajalise piirangu tõttu pidi tegema alamvalimi. Et saada Eesti metsade kohta üldisemaid näiteid kasutati selleks KMeans klasterdamise meetodit, et jagada metsad omakorda kahte erinevasse klassi

okas- ja lehtpuud. Eesmärgiks oli koguda kokku 100 raiet ja nende ümbrust, et luua piisavalt andmeid, mille pealt mudelit treenida. Kui peaks juhtuma et mõni pilt pole kõlblik, siis valitakse samas klustrist uus lageraie piirkond.



Joonis 3. KMeans klasterdamise tulemus, 100 raie ümbrust leht- või okaspudega

Peale selle etapi lõppu, kus on loodud 100 raie ümbrust polügoonidena. Edasi kirjeldatakse lahti sammud, välja toodud joonisel 4, et saada nendega seonduvad satelliitpildid.

Kuna andmete päring ja töötlemine nõudis suurt arvutus- ja salvestusressurssi, algas töö ühe üksiku lageraie piirkonna ja selle vahetu ümbruse salvestamisega ühte .shp-vormingus vektorfaili. Edasiste testide käigus selgus, et optimaalne lähenemine hõlmab 10–15 lageraie ala ning nende ümbruste sõltumatut päringut ja salvestamist eraldi shape-failidena. Selle põhjal viidi läbi järgnev töölökude jadastik:

1. Shape-failide loomine ja kokkuliitmine

Iga lageraie piirkonna ja selle vahetu ümbruse (buffer) geomeetria päriti riiklikust puistu andmebaasist ja salvestati iseseisvatesse shapefile'idesse. Selle lähenemise eeliseks oli võimalus vajalikke alasid hõlpsalt lisada või eemaldada ilma terve andmestiku ümberlaadimiseta. Paremaks haldamiseks ja analüütiliseks töötlusteks liideti kõik eraldi genereeritud shapefile'id tervikusse, mis võimaldas vektorandmete

ühtlustatud töötlemist järgmistes etappides.

2. Piltide päring Sentinel-2 andmebaasist

Iga shape-failiga määratletud lageraie ala serva genereeriti puhvrina ristikülikukujuline “kast” (bounding box), mis täielikult kataks huvipakkova piirkonna. Selle kasti geometria alusel loodi päring Sentinel-2 metadatabaasist, eesmärgiga leida piirkonnale vastav optiline satelliitpilt. Päringule lisati järgmised piirangud:

- (a) **Ajapiirang:** pildistamise kuupäev peab jääma lageraie teadmisest alates kuni 40 päeva jooksul, tagamaks muutuste jälgitavuse tuvastamise võimalikult värskele materjalil.
- (b) **Pilvekatte piirang:** pilt valiti vastavalt väikseimale keskmisele pilvekattele, et maksimeerida nähtavust ja andmete kvaliteeti.

Kui sobiliku satelliitpilti nimetatud tingimustel ei leitud, laiendati ajavahemikku ning tõsteti lubatavat pilvkatte protsendi kuni eelnevalt määratud lävendini. Sel juhul, kui sobivat pildistust ikkagi ei tuvastatud, alustati protsess uuesti järgmise lageraie piirkonnaga samas klastris, et tagada piisav andmepunktide hulk analüüsile teostamiseks.

3. Andmete allalaadimine ja eeltöötlus

Sobiva Sentinel-2 toote puhul laeti vastav .SAFE kataloog alla. Sentineli toode sisaldab üheskoos nii optilisi riba- (band) kui ka metadatastruktuure, mis on vajalikud edasiseks detailseks analüüsiks. Eriti olulised on:

- (a) **Spektraalsed ribad** (sh paindliku dünaamikaga B02–B12 ja veespetsiifilised B8, B8A), mis võimaldavad eristada metsaobjekte ja maapinnanähtusi.
- (b) **Pilve- ja lumeindeksiriba (B10)**, mis lihtsustab pilvisuse automaatset tuvastamist ning tagab eeltöötlemisel täpsema piirkondliku pilvete kaardistamise.

4. Atmosfääriline korrektsioon ja lõikamine

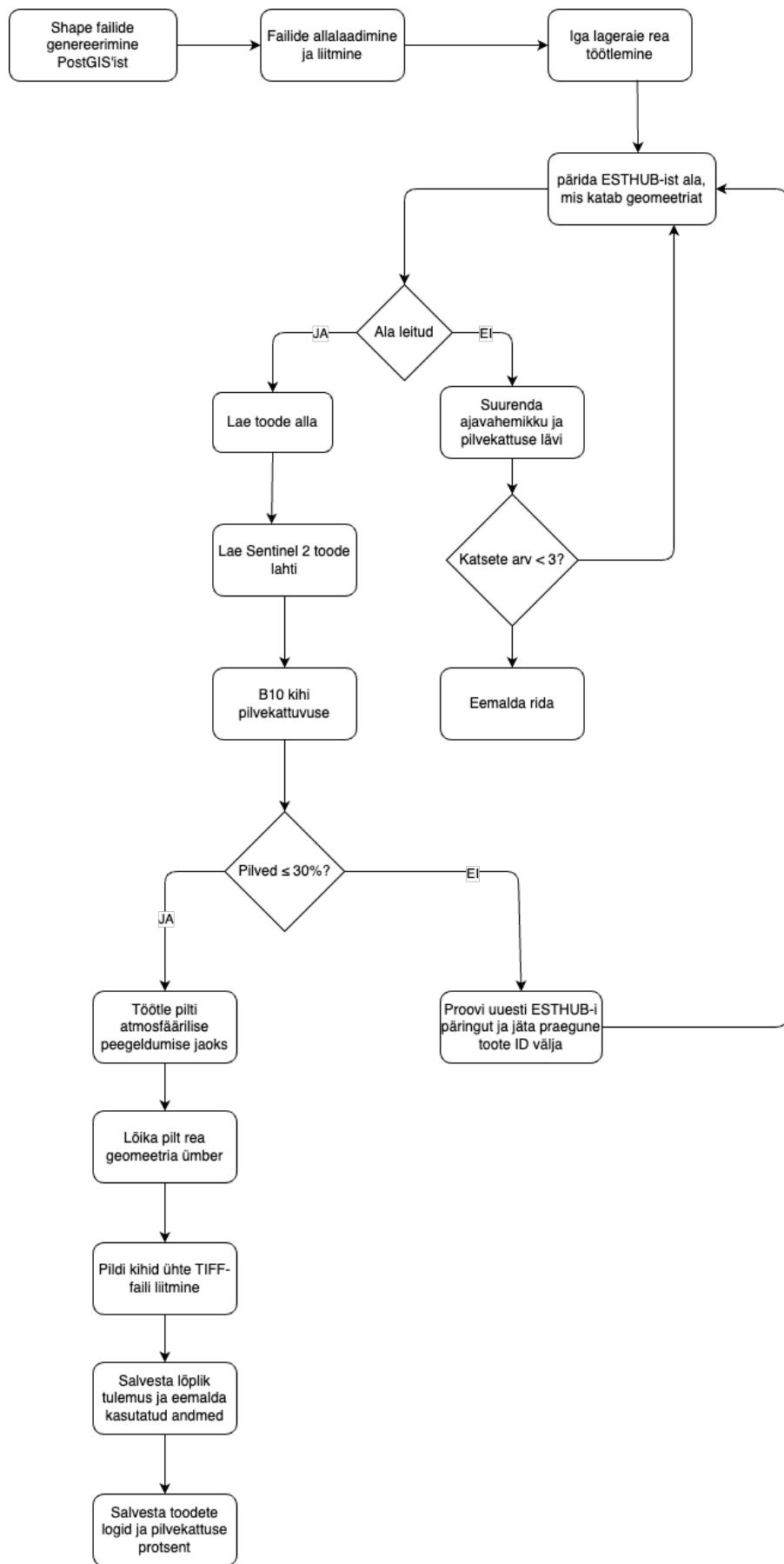
Kõigilt valitud optilistelt ribadelt viidi läbi atmosfääriline korrektsioon, kasutades Sentinel-2 töötlemise standardalgoritmi (Sen2Cor või QGIS’i lahtise lähtekoodiga moodul), et eemaldada atmosfäärist tulenevad häired (nt aerosoolid, gaasiline

neeldumine). Pärast korrektsiooni genereeriti rasterfailide (level-2A) baasil täpsed TIFF-vormingus pildid, kus iga kanal (band) on salvestatud eraldi kihina, säilitades kooskõlastatud ruumilist ja spektraalse täpsust.

5. Huvipiirkondade täpne väljavõtmine

Lõpuks lõigati korrektsiooni läbinud piltidel välja eelnevalt genereeritud kastide ulatuses täpsed rasterfragmendid, et töödelda vaid huvipakkuvaid alasid. Selline meetod tagab, et iga lageraie objekt ja selle ümbris on rasterandmes täpselt isoleeritud, mis lihtsustab edasist masinõpp- ja analüüsitoötlust. Lõppfailid salvestati georeferentseeritud GeoTIFF-idenä, mis toetavad mitmetasandilisi kanalivaateid ja on integreeritavad GIS-keskkondadesse.

Käesolev töövoog võimaldas optimeerida andmekätluse ressursikasutust, tagada piltide kõrge kvaliteet ning kindlustada, et iga lageraie piirkonnast on saadud piisav hulk informatsiooni edaspidiseks statistiliseks ja mudelipõhisteks analüüsideks.



Joonis 4. Andmestiku loomise töövoog

3.2.2 Raie piirkonna maskide loomine

Selles peatükis käsitletakse andmete maskimist, kuna metsaregistrist saadud andmed pole piisavalt täpsed, et neid otse masinõppes kasutada, siis sai loodud sammud, et andmeid täpsemaks muuta. Terviklik dokument on töö lisade hulgas.

Eesmärk oli luua dokument, mis aitaks kõigil kasutajatel luua vektormaske kasutades selleks enne mainitud QGISi tarkvara. Selleks et anda parem hinnang pildil olevatele aladele on kasutusel võetud kihid. Kaks peamist kihti on:

- **Metsateatised** — andmed, mis on saadud riiklikust metsaregistrist ja sisaldavad teavet metsateatise esitamise kuupäeva, metsateatise menetlemise kuupäeva, metsateatise kehtivuse alguskuupäeva ja metsateatise kehtivuse lõppkuupäeva kohta. Selle info abil saab kindlaks teha, kas üldse on vaja märgendada antud piirkonda.
- **Ortofotod** — on kõrge eraldusvõimega õhust tehtud fotod, mis on georeferentseeritud ja millel on täpselt määratletud koordinaadid. Ortofotod võimaldavad täpset analüüsni ja hindamist, et tuvastada metsade seisundit ja muudatusi. Ortofotosid tehakse harvemini kui näiteks Sentinel-2 satelliidipilte ja mitte alati üle terve Eesti, kuid need on vajalikud, et saada täpsemat teavet metsade seisundi kohta. Ortofotod on saadaval erinevates spektraalsetes ribades, sealhulgas nähtavas ja infrapunases spektris, mis võimaldab analüüsida erinevaid maapinna omadusi. Käesolevas töös kasutati nii RGB kui ka CIR-NIR ortofotosid.

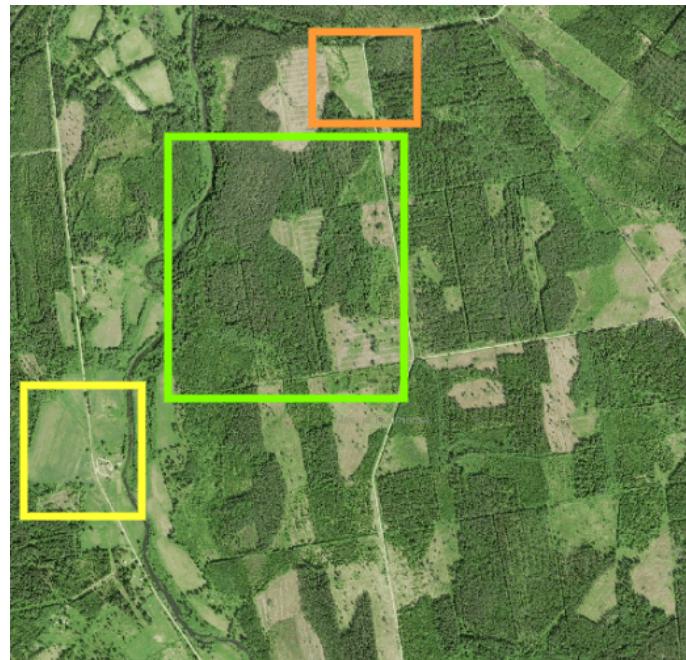
Teiseks on dokumendis toodud välja juhtnöörid ja soovitused, mida peaks järgima maskide loomisel. Järgnevalt on välja toodud mõned näidised, mis illustreerivad ortofotode ja Sentinel-2 piltide erinevusi, seega on tähtis, et orototosid kautatakse juhtnööridena aga mitte põhitõena. Masinõppes kasutatavate andmete kvaliteet antud uurimistöös mängib suurt rolli, seega kui maskide loomise käigus tekkis kahtlus klassi tuvastamises, siis jäeti need alad alati välja.

Joonisel 5, joonisel 6 ja joonisel 7 on välja toodud näidised, mis illustreerivad Sentinel-2 saadud pilti ja selle sama piirkonna ortofotosid. NGR värvitölgendused on toodud välja varasemas peatükis 2 tabelis 2. Oranžiga on välja toodud piirkond, mille Sentinel-2 satelliidipildi järgi peaks märkima lageraieks, RGB ortofoto samuti viitab sellele. NGR

pildil on ka näha tsüaan tooni, mis viitab sellele, et seal peaks olema mulda, ehk suure tõenäosusega võib seal olla lageraie. Siinjuhul peab meeles pidama ka, et urimistöös kasutatud ortofotode ja Sentinel-2 piltide vahel on ajaline erinevus ja olgugi, et ortofotol on selgesti oranžis kassis kaks lageraiet, siis Sentinel-2 pildi järgi saab märkida ainult ühe neist lageraieks. Metsa näidis on toodud välja rohelise kastiga, millel on kõigil kolmel pildil suuresti ühtlane kattuvus. Eraldi on piltidel kollase kastiga välja toodud heinamaa piirkond. Heinamaad ja põllumaad on alad, mis satelliitpildilt ja RGB ortofoto võivad põgusal vaatamisel tunduda lageraie alana, kuna seal puid ei kasva, aga kui vaadata lähemalt NGR pilti, siis on näha, et seal on vähem mullale viitavat tsüaan tooni, millest võib järeldada, et seal võiks hoopis põld asuda.



Joonis 5. Sentinel-2 RGB näidis



Joonis 6. RGB ortofoto näidis



Joonis 7. CIR-NGR ortofoto näidis

3.3 Alusmudeli ülevaade

Alusmudelid (*Foundation models*) on suuremahulistel andmekogudel ennastjuhendavalt treenitud sügavad närvivõrgud, mis toimivad üldotstarbelise baasina mitmesuguste masinõppe ülesannete lahendamiseks. Erinevalt traditsioonilistest mudelitest, mis on välja töötatud konkreetse ülesande jaoks ja nõuavad eraldi treeningut, on alusmudelid eelnevalt

ettevalmistatud laia valiku ülesannete sooritamiseks — alates loomlikust keele töölemissest ja tekstigeneratsioonist kuni pildiklassifitseerimise ja vastuste genereerimiseni — ilma täiendava märgendatud õppematerjalita. Nende mudelite kohanemisvõime tuleneb nii suurest parameetrite hulgast kui ka enesekontrollil põhinevast õppestrategiast, mis võimaldab neid hõlpsasti peenähälestada konkreetsete rakenduste jaoks. Võimalus keskenduda mudeli peenähälestusele ja mitte nullist treenimisele vähendab omakorda oluliselt arendusaega ja arvutiressursside vajadust. [24]

3.3.1 DINO v2 ja selle kasutus

Käesolev peatükk analüüsib DINO v2 mudeli võrdlust teiste tipptasemel mudelitega, eriti semantilise segmenteerimise kontekstis, ning põhjendab selle valikut metsaalade ja metsaraie tuvastamiseks satelliidipiltidel.

DINO v2 on Meta AI poolt loodud isejuhendatud (self-supervised) mudelite kogum, mille eesmärk on õppida üldotstarbelisi visuaalseid omadusi ilma märgendatud andmeteta. Mudel põhineb Vision Transformer (ViT) arhitektuuril, mille erinevad variandid (nt ViT-S/14, ViT-B/14, ViT-L/14 ja ViT-g/14) on eelõpetatud suurel, mitmekesise sisuga ja kureeritud pildikogumil. Mudeli struktuuri põhjaks on õpetaja–õpilase skeem, kus õpilasmudeli parameetreid koheldakse tavalise ViT-võrguna, aga õpetajamudeli kaale uuendatakse õpilase kaalude eksponentsialse libiseva keskmise kaudu. Treeningprotsessi stabiliseerimiseks ja tunnusruumi hajutamiseks on lisatud Kozachenko–Leonenko (KoLeo) regulaarija, mis soodustab tunnuste ühtlast jaotust. Õppetöö lõpus suurendatakse sisendpiltide resolutsiooni ajutiselt 518×518 pikslile, et parandada piksli tasemel ülesannete, näiteks semantilise segmenteerimise ja objektivastuse ennustuse täpsust. Praktikas saavutab DINO v2 tänu optimeeritud FlashAttention-i ja PyTorch Full-Sharded Data Parallel (FSDP) meetodile kuni kahekordse kiiruse ning vajab kuni kolm korda vähem mälumahtu, võrreldes varasemate SSL-mudelitega. [25]

Ühes uuringus võrreldi DINO v2 mudelit mitmete teiste segmenteerimismeetoditega geoloogiliste kompuutertomograafia (KT) skaneeringute analüüsил. Hinnati Otsu lävendamist, klastritehnikaid (K-keskmised, fuzzy C-means), juhendatud masinõppe meetodit (Random Forest) ja süvaõppe mudeliteid (UNet, ResNet152 ja DINOV2). Tulemused näitasid, et eelnevalt treenitud DINO v2 demonstreeris tugevat jõndlust kivimipiltide klassifitseerimisel

isegi siis, kui KT-skaneeringud ei kuulunud selle esialgsesse treeningandmestikku. Segmenteerimise ülesannetes ületas eriti LoRA abil peenhäälestatud DINO v2 teisi meetodeid mitmeklassiliste ülesannete puhul, isegi piiratud andmetega. Visuaalne kontroll näitas, et DINO v2 poolt genereeritud segmenteerimismärgid olid sageli täpsemad kui algsed sihtmärgid. [26]

Teises uuringus hinnati DINO v2 jõudlust vasaku koja segmenteerimisel MRI piltidel. Eelnevalt treenitud DINO v2 saavutas Dice'i sarnasuse koefitsiendi (DSC) 87.1 %. Vöndluses nullist treenitud U-Net mudeliga näitasid tulemused, et iseseisev õpe DINO v2 abil võib oluliselt parandada meditsiinilise pildisegmenteerimine tõhusust ja täpsust, vähendades vajadust suurte märgistatud andmestike järele. [27]

Lisaks on Panopticon, DINO v2 raamistikule ehitatud mudel, saavutanud tiptasemel tulemusi GEO-Bench vöndlustestis maakatte klassifitseerimisel, ületades teisi sensoragnostilisi mudeliteid ning näidates konkurentsivõimet ka kindlatele sensoritele kohandatud mudelitega. See demonstreerib DINO v2 arhitektuuri tugevust ja kohanemisvõimet erinevate kaugseire andmetüüpidega. [28]

Ühes teises töös vörreldi DINO (DINO v2 eelkäija) laiendust DINO-MC teiste iseseisvalt õppivate meetoditega kaugseire ülesannetes. Lineaarse testimise käigus saavutas DINO-MC 2.56% kõrgema täpsuse väiksema eelnevalt treenitud andmestikuga kui SeCo. Peenhäälestamisel kahel kaugseire klassifitseerimise ja muutuste tuvastamise ülesandel ületas DINO-MC nii DINO kui ka SeCo mudeliteid. [29]

Ülaltoodud näited illustreerivad DINO v2 mudeli tugevust ja mitmekülgust erinevates segmenteerimise ülesannetes, sealhulgas geoloogiliste, meditsiiniliste ja kaugseire piltide analüüsил. Mudeli võime saavutada konkurentsivõimelisi tulemusi ilma spetsiifilise peenhäälestusetaga ning selle parem üldistusvõime muudavad selle sobivaks valikuks metsaalade ja metsaraie tuvastamiseks satelliidipiltidel.

3.3.2 SAM v2 ja selle kasutus

Segment Anything Model 2 (SAM 2) on pildi- ja videosegmentatsiooni ülesannete fundamentaalne mudel, mis laiendab esmalt välja töötatud SAM-i (Kirillov et al., 2023 [30]) võimekust staatelistel piltidel ka ajaliselt dünaamilisele videomaterjalile. SAM 2

arhitektuur põhineb lihtsal transformeril, millel on integreeritud vooguv mälumoodul — see võimaldab reaalajas töödelda järjestikku sisenevaid kaadriseid ja säilitada informatsiooni sihtrühma objektist. Mälumoodul hoiab kuni N varasemat kaadrit ning M kasutajapõhist sisendkaadrit, mille alusel ennustatakse masklet' e (masklet = objekti mask ruumilis-ajalises ulatuses). [31]

SAM v2 põhistruktuur koosneb neljast põhimoodulist, mis töötavad järjestikku, et ühendada nii pildist tulenev informatsioon kui ka kasutajapoolne juhis ning ajaliselt varasemad tulemused. Esiteks võtab hierarhilise pildikodeerija sisse sisendkaadri ja teisendab selle mitme skaalaga omadusvektoriteks. Teiseks kodeerib prompt-moodul kasutaja märgitud punkte, kaste või maski, et mudel teaks, mida täpselt otsida. Kolmandaks lisab mälumoodul iga kaadri puhul salvestatud maski- ja pildiomadused FIFO-põhimõttel mälupanga, võimaldades mudelil säilitada konteksti eelmistest kaadritest. Neljandaks rakendab transformer-põhine tähelepanumehhanism nii jooksva kaadri kui ka mälust saadud tunnustevahelise suhtluse, mille tulemuseks antakse maskidekooderile täpsed ruumilised vihjad. Lõpuks dekodeerib maskidekooder paralleelselt kõik saadud sisendid ning ennustab lõpliku objektimaski, mida mudel iteratiivselt uuendab ja talletab järgmiste kaadrite jaoks. Selline ülesehitus võimaldab SAM v2-1 kiiresti ja tõhusalt segmenteerida nii staatilisi pilte kui ka videoid, säilitades minimaalse kasutajainteraktsiooni juures detailse täpsuse. SAM 2 rakendused on kasutatav laial valikul valdkondades: meditsiinilises pilditöötlustes, kaugseires ja satelliitpildi analüüsides, liikumisskeeme segmenteerivates uuringutes ning maskeeritud objektide tuvastamisel. [31]

Artiklis „MPG-SAM 2: Adapting SAM 2 with Mask Priors and Global Context for Referring Video Object Segmentation“ kohandati SAM V2 mudelit viitematerjalidel põhinevaks videoobjektide segmentimiseks (RVOS). Teadlased lõid multimodaalse enkooderi, mis ühendas video- ja tekstipõhiseid tunnuseid, ning kasutasid maskide eelgeneratoreid globaalse konteksti kaardistamiseks. Uuenduslik hierarhiline ajalooline agregeerimismoodul võimaldas SAM 2-1 säilitada objekti kontuure läbi videokaadrite, parandades ajaliselt järjepidevat segmentatsiooni. Katsetused mitmetel RVOS andmestikel näitasid märgatavat täpsuse kasvu võrreldes varasemate meetoditega. [32]

Artiklis „SAM2 for abdomen: One-shot and no finetuning“ esitletakse SAM2 ja Emb-SAM meetodite kombinatsiooni, mis võimaldab kõhuelundite segmentatsiooni CT (*computed*

tomography) kujutistel, kasutades ainult ühte märgistatud pilti. SAM2 käsitleb järjestikuseid CT-viilusid kui videosarju ning propagatsioon toimub maskimälu kaudu. Emb-SAM täiendab protsessi, luues täpseid pseudo-märkeid iseenesliku õppimisega. Tulemused BTCV andmestikul näitavad tugevat segmentatsioonitäpsust, minimeerides käsitsi sekkumise vajadust. [33]

3.4 Seoste leidmine

Selleks et tõestada antud magistrityö eesmärki segmenteerida lageraie ja metsa alasi suurte visioonimodelitega, klusterdatakse DinoV2 väljundid, et leida, kas need alad on eristatavad. Selleks on vajalik eksperdigia valideerida ühe lageraie piirkonna alad, et kindel olla, et need vastavad tõele satelliidipildilt.

Joonisel 8 on näidis pilt eksperimentist, mille käigus püüti välja selgitada, kas mudelid suudavad eristada lageraie ja metsa alasid satelliidipiltidelt. Antud Sentinel-2 andmebaasist saadud pilt sai valitud esiteks sellepärast, et see on hea nähtavusega, võimalikult vähe pilvine ja soojemal hooajal, kui pole lund ja puud on lehteis. Teiseks on käesolev pilt heaks aluseks, sest sellel on sattunud raie- ja metsapiirkondi kõrvuti, mis võimaldab erinevatesse klassidesse kuuluvate alade üleminekukohti analüüsida.



Joonis 8. Näidis sateliidi pilt

Joonisel 9 on näha lageraie piirkond, mis on toodud välja tumedama roosana, ja metsa piirkond, mis on joonisel heleroosana. Tegemist on aladega, mis on teada Keskkonnaametile ja mille kohta onolemas ka metsateatis. Satelliidipilt on saadud Sentinel-2 andmebaasist ja sellel on kümnemeetrine resolutsioon. Antud pilt on saadud raie teostamise ajast kuni 40 päeva hiljem, seega mahub õigesse ajaraamistikku, et piisavalt täpselt hinnata metsade seisukorda.



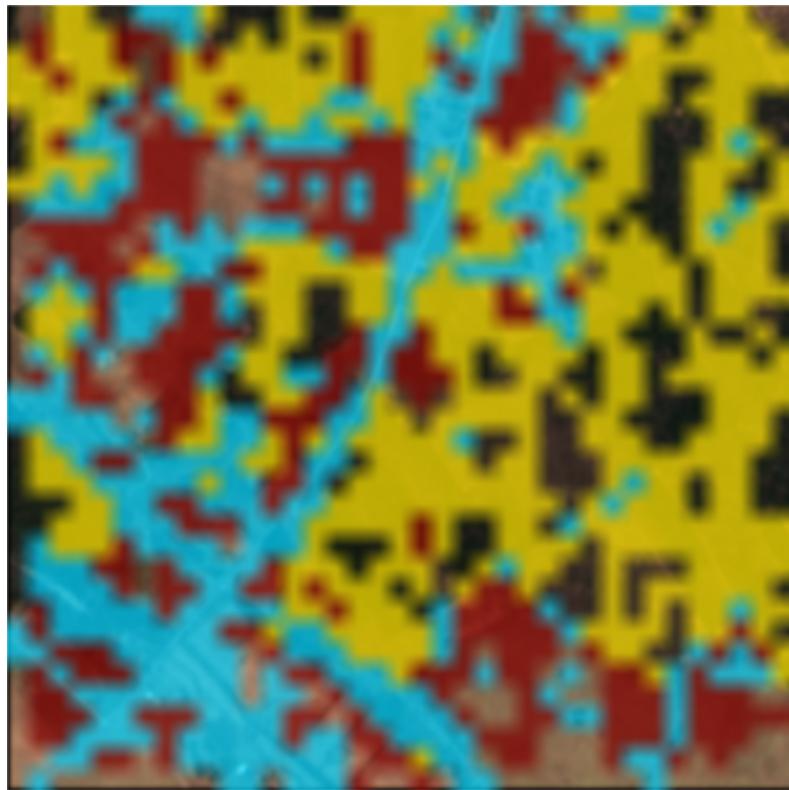
Joonis 9. Raie piirkonna mask sateliidi pildil

Joonisel 10 on näha lageraie piirkond, mis on toodud välja tumepunasena ja metsa piirkond, mis on märgitud rohelisega. Ekspert on võtnud metsateatistest saadud info alade tuvastamisel aluseks, aga sellele lisaks kasutas ta ka muud teavet nagu ortofotosid ja erinevate lainepekkuste pilte, et leida täpsemad piirded metsade ja raiealade vahel. Ekspert on leidnud, et antud piirkonnas leidub alasid, mis ei vasta metsateatistes märgendatud infoole. Näiteks on pildil alasid, mis teatistes on märgitud metsaks, aga silmaga vaadates kujutab hoopis raiet ja ka vastupidi. Seega on eksperdi poolt loodud mask palju täpsem ja seetõttu oli eksperdi kaasamine vajalik andmestiku loomise protsessis.



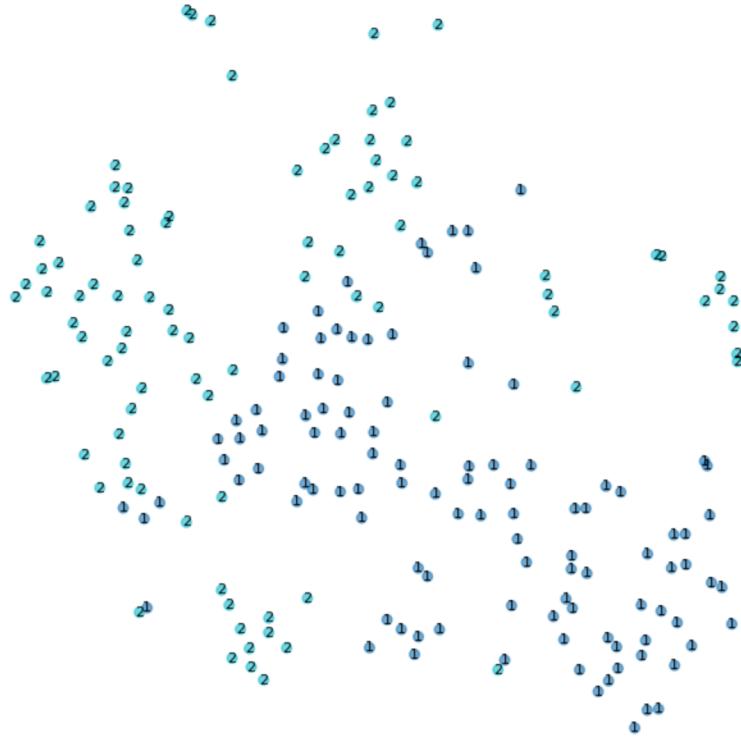
Joonis 10. Raie piirkonna mask eksperdi poolt korrigeeritud

Järgneval joonisel 11, rakendati klastrianalüüsü süvaõppemudelist saadud kõrgdimensionaalsetele tunnusevektoritele, mis esindavad sisendpildi diskreetseid paiku. Eelkõige kasutati k-keskmise algoritmi, et grupeerida need tunnusevektorid sarnaste semantiliste omaduste alusel. Igale pildilaigule vastav tunnusevektor määrati ühte eelnevalt defineeritud arvu klastritesse, mille tulemusena saadi diskreetne klastermärgis iga pildipaiga jaoks. Seejärel visualiseeriti saadud klastermärgis pildil värvilise maskina, kus iga klaster on esitatud unikaalse värviga, võimaldades seeläbi kvalitatiivselt hinnata mudeli õpitud representatsioonide kohalikku sarnasust sisendpildil.



Joonis 11. Segmenteeritud näidis sateliidi pildist

Joonisel 12 on rakendatud klassipõhist klasterdamise meetodit, mille eesmärk on luua pildimaterjalist semantiliselt sidusaid piirkondi, grupeerides pildi elemendid eelnevalt defineeritud klassifikatsioonikategooriate alusel. Lähtudes mudeli genereeritud klassifikatsiooniväljunditest, määrratakse iga pildielement kõige tõenäolisemasse klassi, mille alusel moodustatakse klastrid. Selle meetodi rakendamise tulemusena saadakse segmentatsioon, kus ühte klastrisse kuuluvad pildi osad on mudeli poolt klassifitseeritud sarnaselt. Lõppesmärk on seeläbi genereerida pildist arusaadavam representatsioon, mis võimaldab interpreteerida pildi sisu semantilisel tasemel, tuues esile objektide ja piirkondade klassipõhisid seosed. Pildil on väljatoodud klass 1 mis on mets ja klass 2 mis on lageraie, lisaks sai eemaldatud tausta klass, et oleks kergem jälgida uuritavaid klasse.



Joonis 12. T-SNE kluster analüüs DinoV2 mudeli väljunditest

klass 1 - mets, klass 2 - lageraie

Siiu võis näha, et lageraie ja mets on eristatavad ja annavad alust edasi uurida, millised alusmudelid suudavad paremini eristada lageraie ja metsa alasid.

3.5 Treenimis protsetuurid

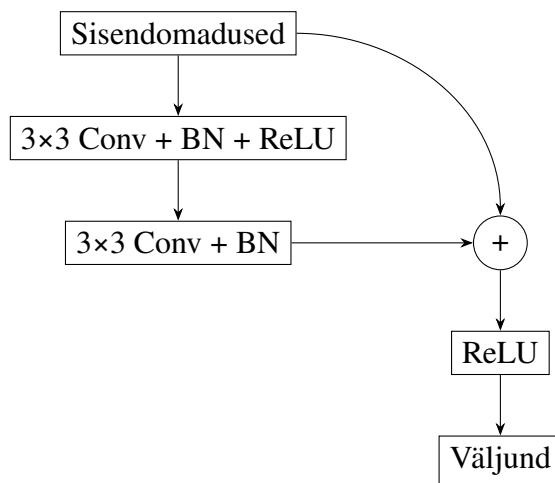
Selles lõigus tutvustame lühidalt kasutatud aluspõhimõtteid ja arhitektuure: ResNet34, ResNet50 ning MobileNetV2 kodeerijad, optimeerijad Adam ja AdamW ning segmентаatsioonipead U-Net, DeepLabv3 ja FPN. ResNet-põhised arhitektuurid rakendavad jäakühikuid, mis võimaldavad väga sügavate võrkude stabiilset treenimist. MobileNetV2 kasutab pööratud jäakühikuid ja lineaarseid pudelikaelu, et tagada mobiilsete seadmete jaoks kõrge efektiivsus. Adam-optimeerija pakub adaptatiivseid esimest ja teist momenti põhinevaid õppemäärade uuendusi , samas kui AdamW eraldab kaalude hääbumise graidentiuendusest . U-Net koosneb kokkutõmbuvast ja laienevast teest, DeepLabv3 kasutab atrooskonvolutsioone ning astrusruumilist pirnikogumit (ASPP), ning FPN loob mitmetsandilise omaduskaardi püramiidi, mis sobib hästi kõrge lahutusvõimega satelliitpiltidega.

Väärtuste kombineerimine ja ulatuslik hüperparameetrite otsing erinevate kodeerijate, optimeerijate ja segmentatsioonipeade vahel annab võimaluse modelleerimisel täpsust ja efektiivsust optimeerida.

Kodeerijad (Encoders)

ResNet34

ResNet34 on 34-kihiline sügav närvivõrk, mis põhineb VGG-19 inspireeritud „plain” arhitektuuril, kuhu on lisatud jäækühikud (residual blocks). Iga jäækühik sisaldab kahe või enama konvolutsioonikihti, mille väljund liidetakse sisendiga „lünkühenduse” (skip connection) kaudu, mis aitab vältida gradientide kadu väga sügavate võrkude puhul. Spetsiifiliselt ResNet34 puhul on kasutusel järjestus 3×3 konvolutsioone ning iga lünkühik korrigeerib sisendi ja väljundi vahelist erinevust, võimaldades algset teavet säilitada.

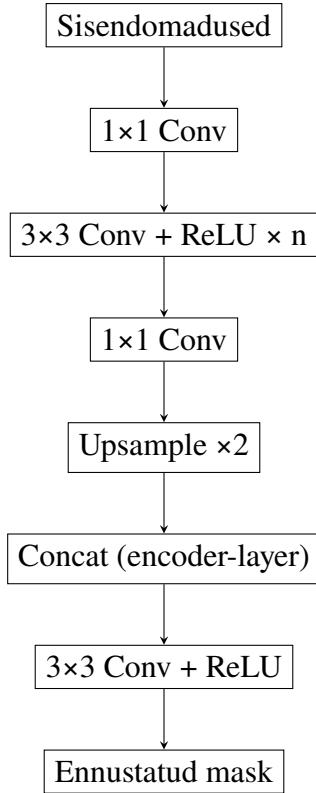


Joonis 13. ResNet34 jäækühikust skeem

ResNet50

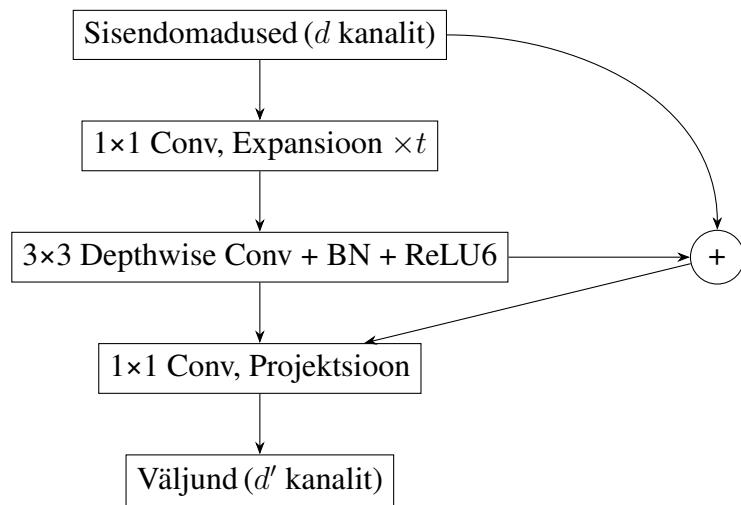
ResNet50 laiendab jäækühikute ideed kolmik-kihiliste pudelikaelablokkidega: 1×1 konvolutsioon vähendab dimensiooni, 3×3 konvolutsioon töötlev omadusi ja 1×1 konvolutsioon taastab dimensiooni. See arhitektuur võimaldab sügavust tõsta 50 kihini, hoides samal ajal arvutuskoormust kontrolli all. Põhimõte on sama – lünkühendus ühendab sisendi väljundiga, et säilitada gradientide voo kogu võrgu ulatuses.

MobileNetV2



Joonis 14. ResNet50 jäälühikust skeem

MobileNetV2 on kavandatud eelkõige mobiilseadmete jaoks, kasutades pööratud jäälühikuid ja lineaarseid pudelikaelu. Pööratud jäälühikud algavad kitsast sisendkanalist, laiendavad omaduste ruumi läbi 1×1 konvolutsiooni, töötlevad seda 3×3 süvendkonvolutsiooniga (depthwise convolution) ja vähendavad seejärel lineaarse pudelikaelaga dimensiooni. Lineaarne aktivatsioon pudelikaelal aitab vältida teabekadu kitsas kanaliruumis.



Joonis 15. MobileNetV2 jäälühikust skeem

Optimeerijad

Adam

Adam (Adaptive Moment Estimation) on gradiendipõhine optimeerija, mis kombineerib RMSProp-i ja inertsiteguri eelised, kasutades kohandatud momentide hinnanguid. Esimene moment (keskmise) ja teine moment (ruutkeskmise) hinnatakse eksponentsialse libiseva keskmisenä, mis võimaldab õppemäära dünaamilist kohandamist iga kaalu tasemele. Adam on efektiivne suurte ja/või mürarohkete gradientide tingimustes ning vajab tavaliselt vähe hüperparameetrite häällestust.

AdamW

AdamW täiustab Adam-i, eraldades kaalude kahandamise (weight decay) gradiendiuuendust, mis parandab regulaarimist ja üldist jõudlust. Kaalude kahandamine lisatakse otse update -sammule, mitte gradientide skaleerimise kaudu, tagades korrektse L2-regularisatsiooni teostuse. See lahendab Adam-i algupärase L2 regulaarimise ühilduvusprobleemi.

Segmentatsioonipead

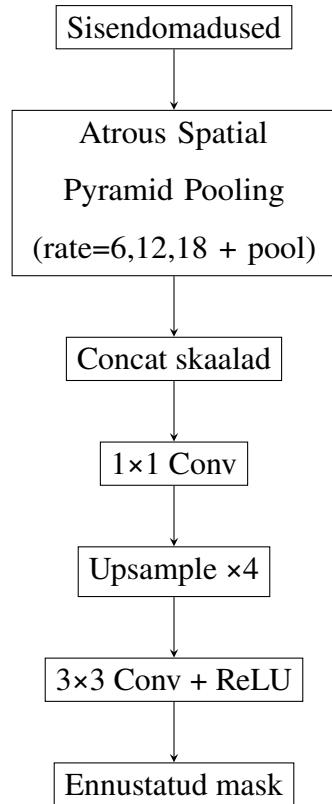
U-Net

U-Net koosneb sümmeetrilisest kokkutõmbuva ja laieneva radade (encoder-decoder) arhitektuurist, kus iga kokkutõmbuv tasand vähendab ruumilist lahutusvõimet ja suurendab tunnuste arvu, ning vastupidises laienevas osas ruumiline lahutusvõime taastatakse. Lünnühendused ühendavad kokkutõmbuva ja laieneva teekonna vastavad kihid, võimaldades detailse teabe säilimist.

DeepLabv3

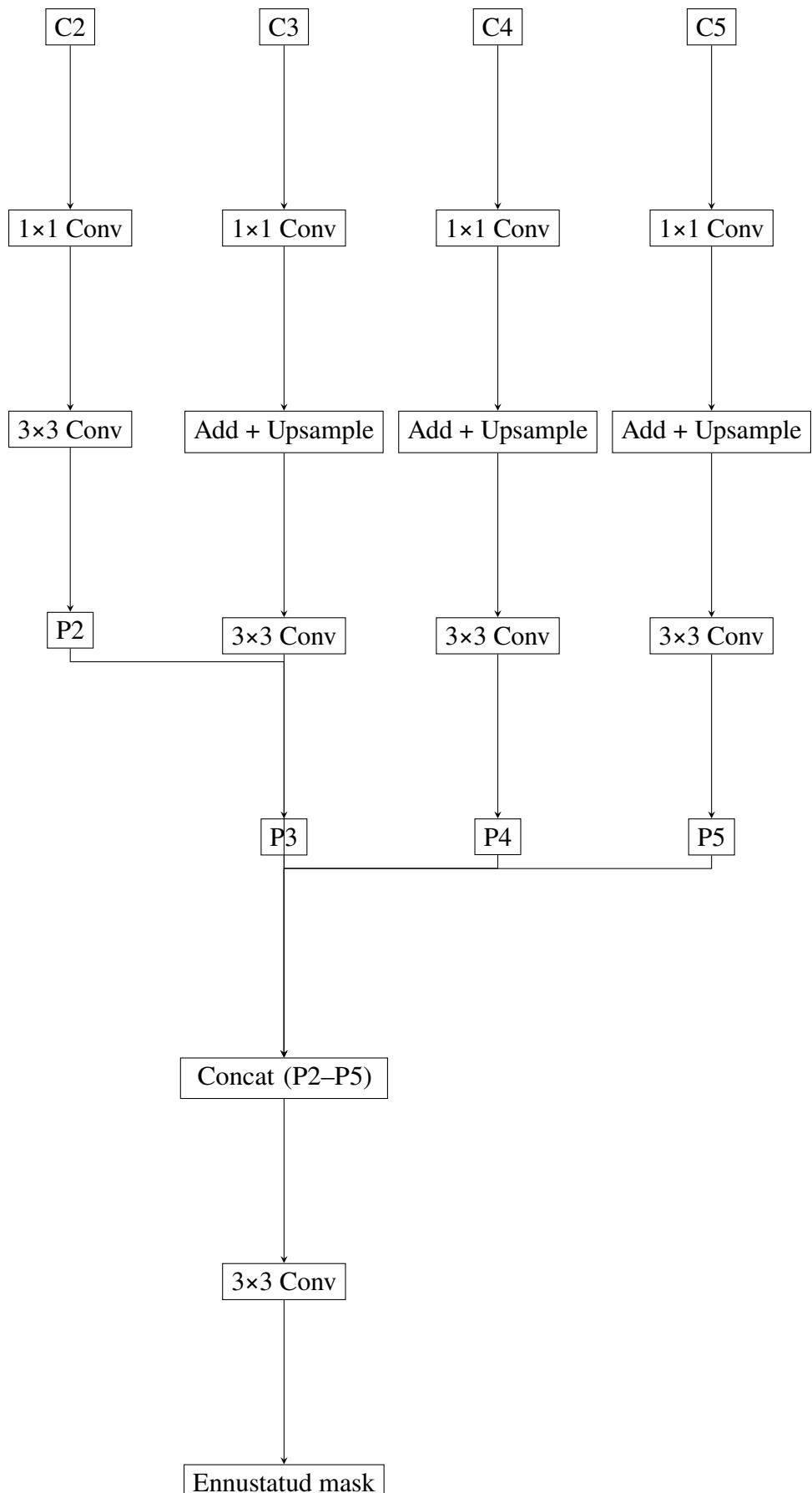
DeepLabv3 kasutab atrooskonvolutsioone (dilated convolutions) erinevatel kihtidel, et laiendada vastuvõtupiirkonda ilma lahutusvõimet kahjustamata, ning Atrous Spatial Pyramid Pooling (ASPP) moodulit, mis kogub omadusi eri skaala tasanditel. See lähenemine võimaldab segmentatsioonil tuvastada objektide piirjooni täpsemalt, eriti keerukates satelliitpiltides.

Feature Pyramid Network (FPN)

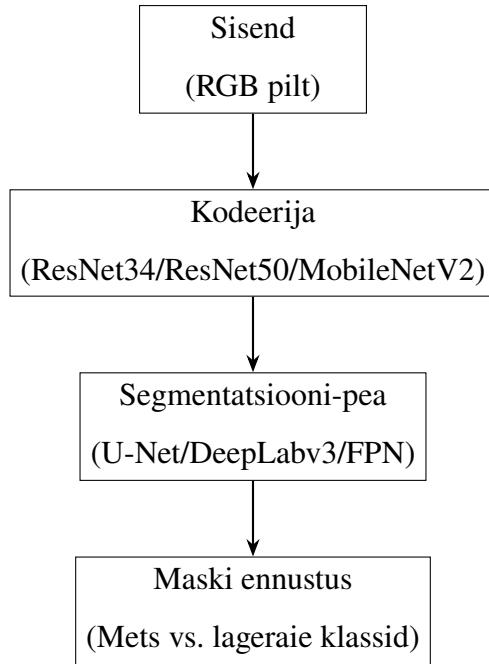


Joonis 16. DeepLabv3 ASPP moodul

FPN loob hierarhilise omaduskaardi püramiidi, kombineerides madalama lahutusvõimega sügavamate kihtide tugevaid semantilisi omadusi kõrgema lahutusvõimega pinnakihtide detailidega. See sobib eriti hästi kõrge lahutusvõimega satelliitpiltide segmentatsiooniks, kus on oluline korraga töödelda nii suuri objekte kui ka peeneid detaile.



Joonis 17. FPN segmentatsioonipea



Joonis 18. Mudeli arhitektuuri skeem

Mudelite kombineerimine ja hüperparameetrite häälestus

Mudelite kombinatsioon sõltub kolmest peamisest komponentist: kodeerijast, optimeerijast ja segmentatsioonipeast. Hüperparameetrite ruumiline uurimine hõlmab kõigi võimalike kombinatsioonide treenimist ning tulemuste võrdlust, mis tagab parima täpsuse ja üldistamisvõime. Hüperparameetrite otsing sisaldas järgmisi seadeid:

Hüperparameeter	Võimalikud väärtsused
encoder_name	{resnet34, resnet50, mobilenet_v2}
encoder_weights	{imagenet, None}
lr	[1×10^{-5} , 1×10^{-3}]
batch_size	{2}
optimizer	{adam, adamw}
head	{unet, deeplabv3, fpn}

Tabel 4. Hüperparameetrite kombinatsioonid

Sellise $3 \times 2 \times 3$ kombinatsiooni korral tuleb kokku 18 eraldi mudelit, mida hinnata näiteks IoU (Intersection over Union) ja täpsuse alusel. Treening seisneb esmajärjekorras kodeerija eel-treenitud kaalu lahtipakkimises (transfer learning), seejärel kogu võrgu peenhäälestuses, kasutades sobivat optimeerijat ning õppemäära skeemi.

Järeldused

Ülaltoodud arhitektuuride ja optimeerijate kombineerimine võimaldab luua segmentatsioonimudeleid, mis suudavad satelliitpilte tõhusalt töödelda ja tuvastada metsasid ning lageraie alasid. ResNet-baasilised kodeerijad pakuvad tugevat omadusloome võimekust, MobileNetV2 lisab efektiivsust piiratud ressursiga kasutuskeskkondades. Adam ja AdamW optimeerijad tagavad stabiilse suuremaks üksuseks liitumise, samas kui U-Net, DeepLabv3 ja FPN pakuvad erinevaid lähenemisi segmentatsioonipeade ülesehitamiseks. Laialdane hüperparameetrite võreotsing (gridsearch) ning kombineeritud mudelite analüüs annab objektiivse aluse valida kõige sobivam konfiguratsioon metsade ja lageraielõikude tuvastamiseks.

Kodeerijad, Optimeerijad ja Segmentatsioonipead

Uurimisöö põhiosa moodustavad DINO v2 ja SAM v2 segmentatsioonimodelite eksperimendid erinevad üksteisest nii mudeli arhitektuuri, hüperparameetrite otsingu kui ka pea-(head-)komponentide poolt, võimaldades võrrelda segmentatsiooni-võimekust erinevates seadistustes.

DINO v2 katse käivitamisel imporditakse esmalt eeltreenitud ViT-tuumvõrgu (backbone) DINO v2 mudelist ja initsialiseeritakse uus DinoV2Segmenter, mille dinov2_backbone_name-parameetriga valitakse ViT-versioon (näiteks “small” või “base”). Tuumvõrgu kaalusid on võimalik kas täielikult külmutada või lasta neil treeningus uuesti kohanduda (freeze_backbone=True/False), mis võimaldab hinnata nii puhtalt eelõpetatud omaduste kui ka peenhäälestamisel põhineva lähenemise mõju. Peade valikus on kolm erinevat varianti: lihtsat LinearClassifier-pead, hierarhilise UPerNetHead-lahendus ning FPN-pea (FPNHead), mis kombineerib eri-tasandilisi omaduskaarte püramiidi kujul. Igal peal on omapärane viis, kuidas manuspaigalt (patch-embedding) kogutud omadused ümber vormida lõplikuks klassi-logititeks või täpsemaks semantilise ruumi rekonstruerimiseks. Treeningprotsessis kasutatakse optimeerijana AdamW-i ja kadufunktsionina mudeli oma sisend-väljastuse kaotusmõõdikut (outputs.loss), samas kui Optuna hüperotsing kontrollib erinevaid õppemäära, backbone'i külmutamise ja pea tüibi kombinatsioone. Kõik eksperimentid salvestavad iga tsükli järel nii IoU kui ka Dice tulemused ning parimad kontrollpunktid kopeeritakse lõppanalüüsiks.

Tunnus	Backbone	Freeze Backbone	Head Type	Optimeerija
Trial 1	ViT-small	Yes	LinearClassifier	AdamW
Trial 2	ViT-small	Yes	UPerNetHead	AdamW
Trial 3	ViT-small	Yes	FPNHead	AdamW
Trial 4	ViT-small	No	LinearClassifier	AdamW
Trial 5	ViT-small	No	UPerNetHead	AdamW
Trial 6	ViT-small	No	FPNHead	AdamW
Trial 7	ViT-base	Yes	LinearClassifier	AdamW
Trial 8	ViT-base	Yes	UPerNetHead	AdamW
Trial 9	ViT-base	Yes	FPNHead	AdamW
Trial 10	ViT-base	No	LinearClassifier	AdamW
Trial 11	ViT-base	No	UPerNetHead	AdamW
Trial 12	ViT-base	No	FPNHead	AdamW

Tabel 5. DINO v2 eksperimendi seadistused

SAM v2 katsed on üles ehitatud Meta “Segment Anything Model” teise põlvkonna raamistikku põhjal ja kasutavad YAML-konfiguratsiooni, et deklareerida nii SamConfig parameetrid kui ka SamModel tüüp. Eksperimendi konfiguratsioonis saab samuti määrata tuumvõrgu külmutamise taseme, et võrrelda, kas pildipõhised tunnused vajavad täiendavat peenhäälestust. Andmestiku laadimine põhineb .pt-failidest kokku kogutud Sentinel-andmetel, mida töödeldakse CustomForestryDataset klassi kaudu, et rakendada sobivad pilditransformatsioonid (skaleerimine, normaliseerimine jne). Treening- ja valideerimislaadurid töötavad batch_size ning varisemise (drop_last) parameetrite alusel, seejärel viakse läbi epohhipõhine treening ja valideerimine, kasutades train_one_epoch ja validate_one_epoch funktsioone. Kadufunktsioniks on klassikaline CrossEntropyLoss ning optimeerijaks samuti AdamW. FPN-pea lisamine SAM eksperimendis võimaldab sügavamaid võrdlusi, sest nii DINO v2 kui ka SAM v2 mudelid kasutavad nüüd sama püramiidipõhist pea-arhitektuuri, mis toetab peenemat lõpliku segmentatsiooni juhtimist. Lõplik analüüs koondab tulemused mudeli konfiguratsioonide, pea-valiku ja tuumvõrgu külmutamise kombinatsioonides, et tuvastada kõige tõhusam lähenemine metsa alade ja lageraielõikude tuvastamiseks satelliitpiltidelt.

Tunnus	Model Config	Freeze Backbone	Head Type	Optimeerija
Trial 1	default	Yes	MaskDecoder (orig.)	AdamW
Trial 2	default	Yes	FPNHead	AdamW
Trial 3	default	No	MaskDecoder (orig.)	AdamW

Trial 4	default	No	FPNHead	AdamW
Trial 5	large	Yes	MaskDecoder (orig.)	AdamW
Trial 6	large	Yes	FPNHead	AdamW
Trial 7	large	No	MaskDecoder (orig.)	AdamW
Trial 8	large	No	FPNHead	AdamW

Tabel 6. SAM v2 eksperimendi seadistused

4 Tulemuste analüüs

Selles peatükis analüüsime eelmises peatükis kirjeldatud katsete tulemusi, et mõista, kuidas erinevad mudelikonfiguratsioonid ja treeningstrateegiad mõjutavad metsastunud ja lageraelõikude tuvastamise täpsust satelliitpiltidel. Arutatakse ka tulemuste statistilist olulisust ning võrreldakse erinevate lähenemiste efektiivsust.

4.1 Tulemuste võrdlus

Baasjoon

Eksperimendi käigus testiti süstemaatiliselt erinevaid hüperparameetreid, et leida optimaalne konfiguratsioon antud ülesande jaoks. Esitatud tulemuste analüüs näitas, et parimad mudelid saavutasid märkimisväärselt kõrgeid Dice'i skoore, ulatudes kuni ~ 0.984 . See on eriti üllatav, arvestades andmestiku koostamise väikest valimit.

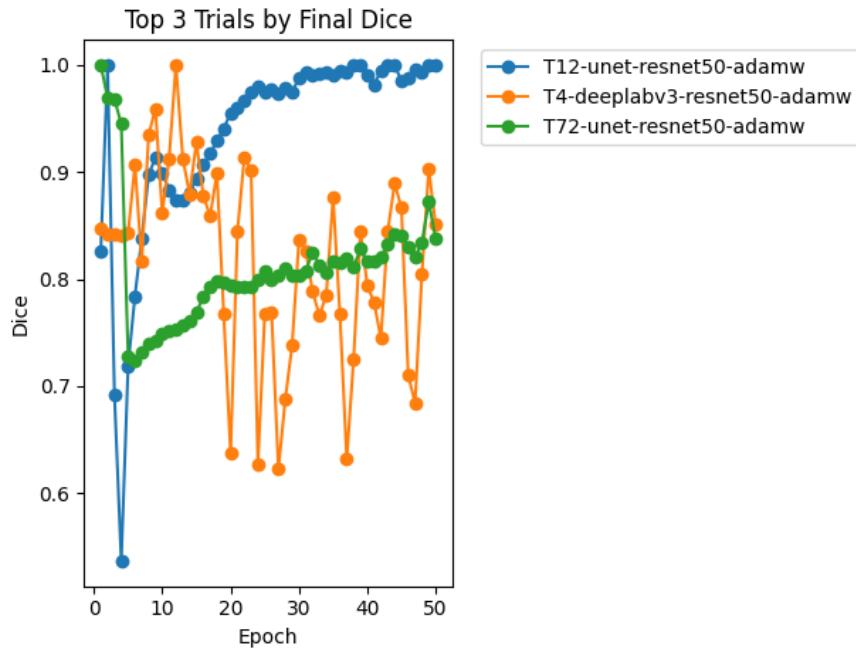
Parimaks osutunud mudeli konfiguratsioon oli järgmine:

- **Arhitektuur:** Unet
- **Enkooder:** ResNet50
- **Enkooderi kaalud:** ImageNet (eelkoolitatud)
- **Optimeerija:** AdamW
- **Õpisamm:** $\sim 1.09\text{e-}4$

Ka teised kombinatsioonid, näiteks DeepLabV3 koos ResNet50-ga, saavutasid kõrgeid tulemusi ($\text{Dice} > 0.97$). Osaliselt on tipptulemused visualiseeritud ka lisatud joonisel 19.

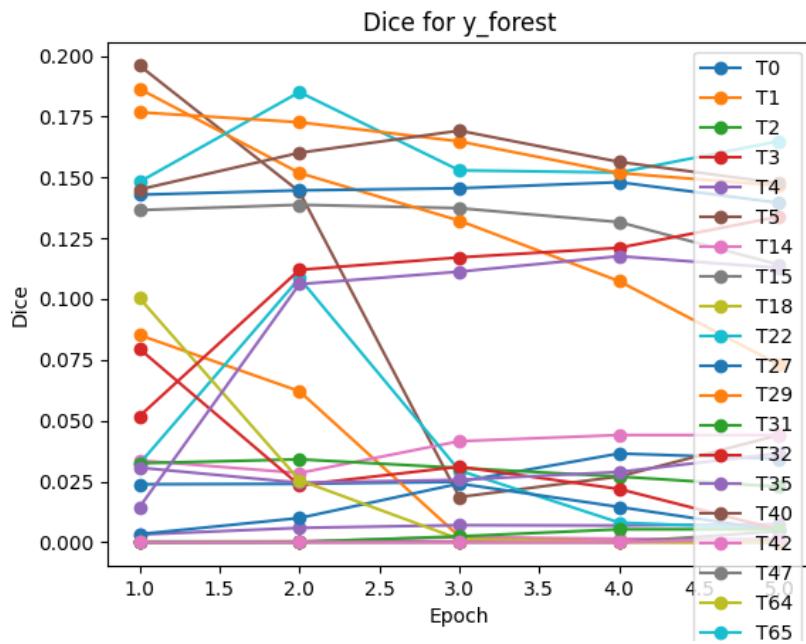
Kõrgete skooride põhjused Kõrgeid tulemusi väikesel andmestikul võib seletada mitme teguriga. Siirdeõpe (*Transfer Learning*): ImageNet andmestikul eelkoolitatud enkooderite kasutamine on tõenäoliselt peamine edu võti. Eelnevalt treenitud mudelid omavad juba võimekust tuvastada üldiseid visuaalseid mustreid (nt servad, tekstuurid), mida saab efektiivselt kohandada spetsiifilisele metsanduse segmenteerimise ülesandele. See vähendab

oluliselt vajamineva treeningandmestiku mahtu. Peaaegu kõik parimad tulemused saavutati just imangenet kaaludega. Kuigi valideerimistulemused on kõrged, on väikse andmestiku kasutamine riskantne, sest alati peab arvestama ülesobitamise(*overfitting*) ohuga. Samas näitab analüüs, et paljudel juhtudel valideerimis- ja treeningkahju (val_loss, train_loss) vähenesid sünkroonis, mis viitab sellele, et mudelid suutsid siiski valideerimisandmetele edukalt üldistuda ega õppinud treeningandmeid lihtsalt pähe.

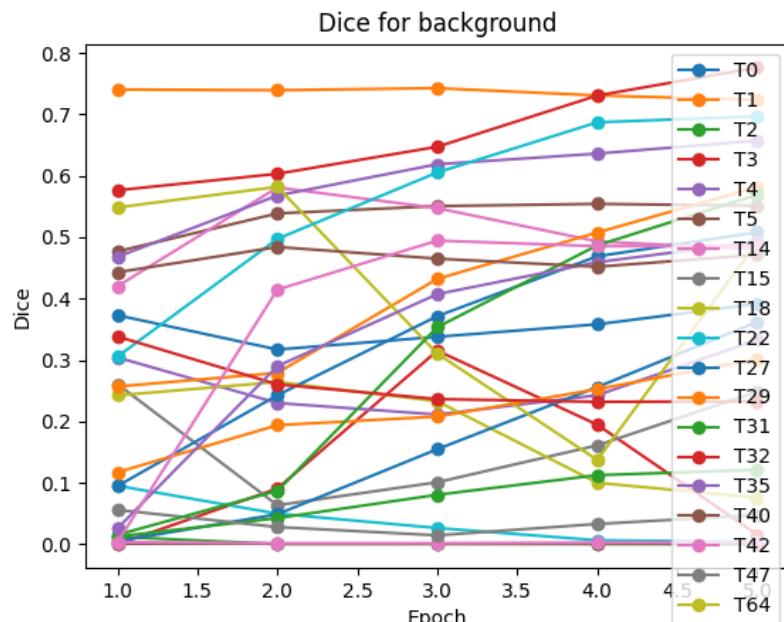


Joonis 19. Parimad DICE tulemused erinevate mudelite vahel.

Samas individuaalsed klassi tulemused (nt metsastumine vs lageraie) Dice skoorid näitasid suuresti, et mudel ei suuda täpselt eristada metsastumist ja lageraielõike. Kõrge skoor tuleneb suuresti tagatausta suurest osakaalust pildil. Joonisel 20 on näha et ülejäänud tausta keskmise Dice skoor on 0.98, mis on väga kõrge. Aga klassipõhised tulemused näitavad, et metsastumise ja lageraielõike eristamine on keeruline. Mudelite katsed ka näitavad, et ka parimate mudelite puhul ei saa nad metsastumist ja lageraielõike eristada.



(a)



(b)

Joonis 20. Tagatausta ja noore metsa segmentatsiooni Dice tulemused üle katsete

DINOv2 eksperimentid

Käesolevas uurimuses viidi läbi eksperimente, mille eesmärk oli hinnata DINOv2 raamistikul põhinevate masinõppemudelite efektiivsust satelliidipiltide segmenteerimisel

metsanduslikus kontekstis. Katsetati erinevaid Vision Transformer (ViT) arhitektuure (nt ViT-B/14, ViT-G/14, ViT-L/14, ViT-S/14) koos erinevate segmenteerimispeadeega (nt LinearHead, SimpleHead, FPNHead, UPerNetHead). Modelite jõudlust hinnati keskmise Dice'i koefitsiendi, keskmise IoU (Intersection over Union) ja keskmise täpsuse (Mean Accuracy) alusel, jälgides neid mõõdikuid kuni 800 epohhi vältel. Treeningutingimustes varieeriti hüperparameetreid, nagu õppimismäär (nt 1e-5, 5e-5, 1e-4), partii suurus (1, 2, 4) ning rakendati nii külmutamist (Freeze) kui ka peenhäälestamist (FineTune).

Konfiguratsioon	Mean Dice	Mean IoU	Mean Täpsus
ViTB14 LinearHead LR5e-5 BS2 E5 Freeze	0.18	0.16	0.26
ViTB14 SimpleHead LR1e-5 BS4 E5 FineTune	0.200	0.26	0.35
ViTG14 FPNHead LR1e-5 BS4 E5 FineTune	0.258	0.28	0.33
ViTG14 SimpleHead LR1e-4 BS4 E5 Freeze	0.22	0.275	0.350
ViTL14 UPerNetHead LR1e-5 BS1 E3 Freeze	0.160	0.270	0.34
ViTS14 FPNHead LR1e-5 BS4 E5 FineTune	0.23	0.25	0.38

Tabel 7. DINOv2 eksperimendi tulemused

Eksperimentide tulemused näitavad, et DINOv2-põhiste modelite jõudlus satelliidipiltide segmenteerimisel sõltub oluliselt tuumvõrgu suurusest, segmenteerimispea keerukusest ja treeningustrategiast. Suuremad tuumvõrgud (nt ViTG14) ja mitme skaala tunnuseid töötlevad pead (nt FPNHead) andsid parimaid tulemusi, saavutades keskmise IoU kuni 0,28 ja täpsuse kuni 0,38. Pehnhäälestamine parandas jõudlust võrreldes külmutatud mudelitega, kuid pikajalistes treeningutes (500–800 epohhi) ilmnes ebastiabiilsus, mis viitab ületreenimisele või ebapiisavale regulatsioonile.

Soovitused edasiseks uurimistööks hõlmavad varajase peatamise ja õppimismäära ajakava rakendamist stabiilsuse suurendamiseks, samuti täiendavate arhitektuuride (nt hübriidmodelite) testimist jõudluse lae ületamiseks. Praegused tulemused pakuvad väärthuslikku alust metsanduslike satelliidipiltide segmenteerimise optimeerimiseks.

SAMv2 eksperimendid

Sarnaselt DINOv2 katsetele viidi läbi ka eksperimendid Segment Anything Model (SAMv2) raamistikku kasutamisel. SAMv2 on tuntud oma võime poolest tegeleda erinevate segmensiooniülesannetega, sealhulgas satelliidipiltide analüüsiga. Katsetes uuriti erinevaid

hüperparameetreid, sealhulgas õppimismäära, partii suurust ja treeninguetappe, et leida parim konfiguratsioon metsanduse kontekstis.

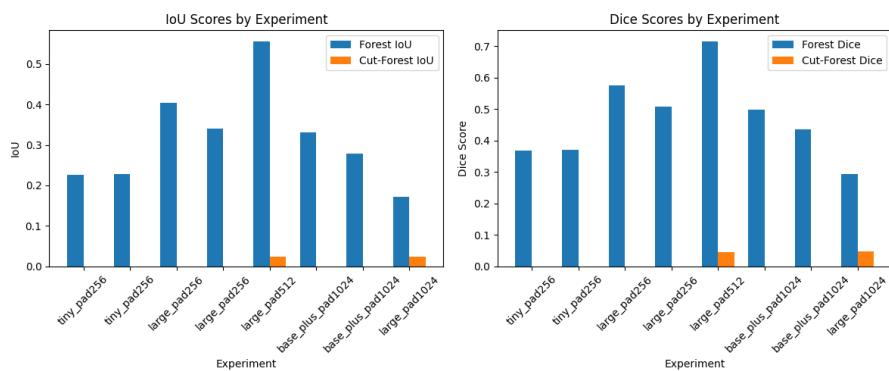
Tabelis on seitse eksperimenti, mis erinevad mudeli variandi, epohhide arvu, õppimiskiiruse ja segmenteerimispea tüübi poolest. Kõik eksperimendid kasutavad 501 epohhi. Parima konfiguratsiooni leidmiseks võrdleme IoU ja Dice väärtsusi, kuna need on peamised näitajad mudeli jõudluse hindamisel. Tabelist selgub, et kõrgeimad tulemused on järgmised:

- Kõrgeim IoU: 0.350205 (konfiguratsioon: large, õppimiskiirus 1×10^{-5} , Multi-Layer pea).
- Kõrgeim Dice: 0.262514 (sama konfiguratsioon: large, õppimiskiirus 1×10^{-5} , Multi-Layer pea).

Variant	Epochs	Learning Rate	Head Type	IoU	Dice
tiny	501	1×10^{-4}	Simple	0.229138	0.150820
tiny	501	1×10^{-4}	Multi-Layer	0.256366	0.178527
large	501	1×10^{-4}	ASPP	0.327914	0.246620
large	501	1×10^{-4}	U-Net Style	0.289425	0.205382
large	501	1×10^{-5}	Multi-Layer	0.350205	0.262514
base_plus	501	1×10^{-4}	Multi-Layer	0.279433	0.195139
large	501	1×10^{-4}	Simple	0.242327	0.163633

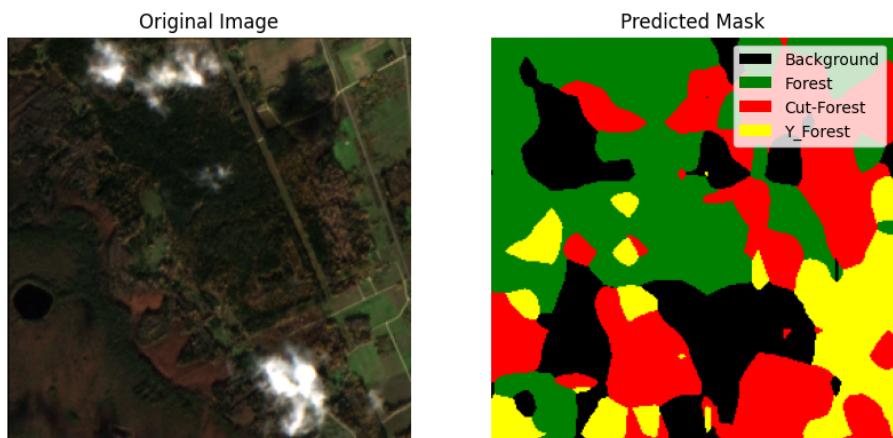
Tabel 8. Samv2 eksperimendi tulemused

Järgnevas joonises on välja toodud SAMv2 eksperimendi tulemused eksperimendi lõikes ja jaotadud klassideks.



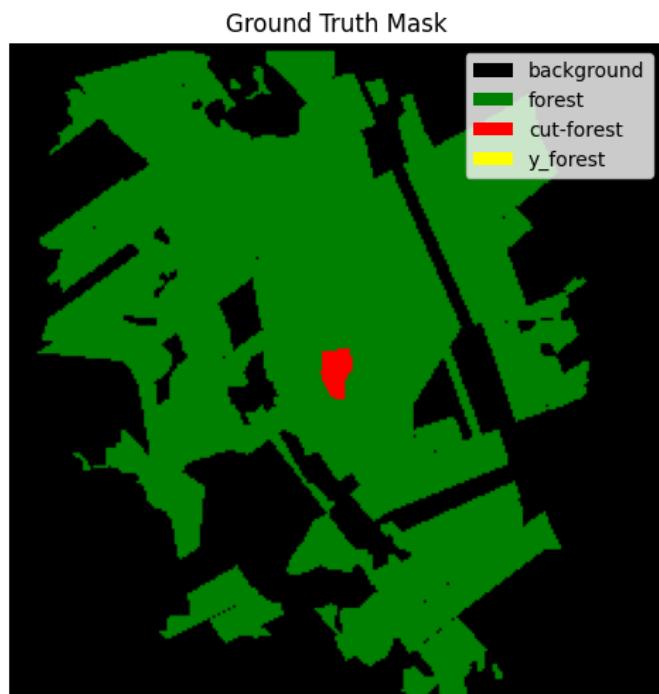
Joonis 21. SAMv2 eksperimendi tulemused klasside lõikes.

Kasutades eelmises osas treenitud SAMv2 mudelit, genereerisime segmentatsioonikaardid:



Joonis 22. SAMv2 pareima mudeli segmentatsioonitulemused

Metsateatistel põhinev segmentatsiooni tulemus:



Joonis 23. Metsateatistel põhinev segmentatsiooni tulemus

Analüüsikoikuvõte

Kokkuvõttes näitavad eksperimendid, et DINOv2 ja SAMv2 raamistike kasutamine satelliidipiltide segmenteerimisel metsanduse kontekstis võib anda häid tulemusi, kuid nõuab hoolikat hüperparameetrite valikut ja treeningstrateegiate rakendamist. Parimad tulemused saavutati suuremate tuumvõrkude ja keerukamate segmenteerimispeadeega, mis

suudavad efektiivselt töödelda erinevaid visuaalseid mustreid. Tulevikus võiks uurida ka hübriidmudelite kasutamist, mis kombineerivad erinevate arhitektuuride eeliseid.

4.2 Edasiarendus ja täiustamine

5 Kokkuvõte

kokkuvõte

Kasutatud kirjandus

- [1] *Tartu Ülikooli teadlased töötasid välja statistilised meetodid Eesti metsaressursi hindamiseks kaugseireandmete põhjal* | Tartu Ülikool. R, 11.09.2020 - 11:05. URL: <https://ut.ee/et/sisu/tartu-ulikooli-teadlased-tootasid-valja-statistikilised-meetodid-eesti-metsaressursi-hindamiseks> (vaadatud 11.01.2025).
- [2] Loretta Cheung, Jonathan Mason ja Meaghan Parker-Forney. „Perimeter Defense: 4 Technologies for Detecting and Preventing Illegal Logging“ (Fri, 11/06/2015 - 11:16). URL: <https://www.wri.org/insights/perimeter-defense-4-technologies-detecting-and-preventing-illegal-logging> (vaadatud 01.03.2025).
- [3] Maa-amet. *Riiklik satelliidiandmete keskus ESTHub*. URL: <https://geoportaal.maaamet.ee/est/ruumiandmed/riiklik-satelliidiandmete-keskus-esthub-p443.html> (vaadatud 26.02.2025).
- [4] *Infrastructure Overview* | Copernicus. URL: <https://www.copernicus.eu/en/about-copernicus/infrastructure-overview> (vaadatud 01.03.2025).
- [5] *Sentinel-2 - an Overview* | ScienceDirect Topics. URL: <https://www.sciencedirect.com/topics/earth-and-planetary-sciences/sentinel-2> (vaadatud 01.03.2025).
- [6] K. Isaienkov *et al.* „Deep Learning for Regular Change Detection in Ukrainian Forest Ecosystem with Sentinel-2“. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021), lk. 364–376. doi: 10.1109/JSTARS.2020.3034186.
- [7] N.S. Podoprigorova *et al.* „Recognition of Forest Damage from Sentinel-2 Satellite Images Using U-Net, RandomForest and XGBoost“. Teoses: Proceedings of the 2024 6th International Youth Conference on Radio Electronics, Electrical and Power Engineering, REEPE 2024. 2024. doi: 10.1109/REEPE60449.2024.10479810.
- [8] *Intersection over Union (IoU): Definition, Calculation, Code*. URL: <https://www.v7labs.com/blog/intersection-over-union-guide> (vaadatud 26.02.2025).
- [9] *Understanding DICE COEFFICIENT*. URL: <https://kaggle.com/code/yerramvarun/understanding-dice-coefficient> (vaadatud 26.02.2025).
- [10] A. Karsenty. „Underlying Causes of the Rapid Expansion of Illegal Exploitation of Tropical Timber“. *International Forestry Review* 5.3 (1. september 2003), lk. 236–239. ISSN: 1465-5489. doi: 10.1505/IFOR.5.3.236.19136. URL: <http://www.ingentaconnect.com/content/10.1505/IFOR.5.3.236.19136> (vaadatud 04.03.2025).
- [11] *WWF - The Importance of Forests*. URL: https://wwf.panda.org/discover/our-focus/forests_practice/importance_forests/ (vaadatud 04.03.2025).

- [12] *About Copernicus | Copernicus*. URL: <https://www.copernicus.eu/en/about-copernicus> (vaadatud 04.03.2025).
- [13] Maa-amet. *National Satellite Data Centre ESTHub*. URL: <https://geoportaal.maaamet.ee/eng/spatial-data/national-satellite-data-centre-esthub-p654.html> (vaadatud 10.03.2025).
- [14] *SI Applications*. URL: <https://sentiwiki.copernicus.eu/web/s1-applications> (vaadatud 04.03.2025).
- [15] *S2 Applications*. URL: <https://sentiwiki.copernicus.eu/web/s2-applications> (vaadatud 04.03.2025).
- [16] *S3 Mission*. URL: <https://sentiwiki.copernicus.eu/web/s3-mission> (vaadatud 04.03.2025).
- [17] *S5P Applications*. URL: <https://sentiwiki.copernicus.eu/web/s5p-applications> (vaadatud 04.03.2025).
- [18] *S2 Mission*. URL: <https://sentiwiki.copernicus.eu/web/s2-mission> (vaadatud 25.03.2025).
- [19] *8. Coordinate Reference Systems — QGIS Documentation Documentation*. URL: https://docs.qgis.org/3.40/en/docs/gentle_gis_introduction/coordinate_reference_systems.html (vaadatud 26.03.2025).
- [20] Grant D. Pearse *et al.*, „Developing a Forest Description from Remote Sensing: Insights from New Zealand“. *Science of Remote Sensing* 11 (1. juuni 2025), lk. 100183. ISSN: 2666-0172. doi: 10.1016/j.srs.2024.100183. URL: <https://www.sciencedirect.com/science/article/pii/S2666017224000671> (vaadatud 06.04.2025).
- [21] *Metsateatis ja metsaregister | Keskkonnaamet*. URL: <https://keskkonnaamet.ee/elusloodus-looduskaitse/metsandus/metsateatis-ja-metsaregister> (vaadatud 02.04.2025).
- [22] *Metsa Korraldamise Juhend–Riigi Teataja*. URL: <https://www.riigiteataja.ee/akt/13124148?leiaKehtiv> (vaadatud 20.04.2025).
- [23] *What Is a Location Polygon?* URL: <https://www.narrative.io/knowledge-base/concepts/data-terms/what-is-a-location-polygon> (vaadatud 02.04.2025).
- [24] *What Are Foundation Models? - Foundation Models in Generative AI Explained - AWS*. Amazon Web Services, Inc. URL: <https://aws.amazon.com/what-is/foundation-models/> (vaadatud 19.04.2025).
- [25] Maxime Oquab *et al.* *DINOv2: Learning Robust Visual Features without Supervision*. 2. veebruar 2024. doi: 10.48550/arXiv.2304.07193. arXiv: 2304.07193 [cs]. URL: <http://arxiv.org/abs/2304.07193> (vaadatud 19.04.2025). Eelnevalt avaldatud.
- [26] *DINOv2 Rocks Geological Image Analysis: Classification, Segmentation, and Interpretability*. URL: <https://arxiv.org/html/2407.18100v3> (vaadatud 07.05.2025).

- [27] Bipasha Kundu *et al.* *Assessing the Performance of the DINov2 Self-supervised Learning Vision Transformer Model for the Segmentation of the Left Atrium from MRI Images*. 14. november 2024. doi: 10.48550/arXiv.2411.09598. arXiv: 2411.09598 [eess]. URL: <http://arxiv.org/abs/2411.09598> (vaadatud 07.05.2025). Eelnevalt avaldatud.
- [28] Panopticon: Advancing Any-Sensor Foundation Models for Earth Observation. URL: <https://arxiv.org/html/2503.10845v1> (vaadatud 07.05.2025).
- [29] Extending Global-Local View Alignment for Self-Supervised Learning with Remote Sensing Imagery. URL: <https://arxiv.org/html/2303.06670v2> (vaadatud 07.05.2025).
- [30] Alexander Kirillov *et al.* Segment Anything. 5. aprill 2023. doi: 10.48550/arXiv.2304.02643. arXiv: 2304.02643 [cs]. URL: <http://arxiv.org/abs/2304.02643> (vaadatud 17.05.2025). Eelnevalt avaldatud.
- [31] Nikhila Ravi *et al.* SAM 2: Segment Anything in Images and Videos. 28. oktoober 2024. doi: 10.48550/arXiv.2408.00714. arXiv: 2408.00714 [cs]. URL: <http://arxiv.org/abs/2408.00714> (vaadatud 16.05.2025). Eelnevalt avaldatud.
- [32] Fu Rong *et al.* MPG-SAM 2: Adapting SAM 2 with Mask Priors and Global Context for Referring Video Object Segmentation. 10. märts 2025. doi: 10.48550/arXiv.2501.13667. arXiv: 2501.13667 [cs]. URL: <http://arxiv.org/abs/2501.13667> (vaadatud 16.05.2025). Eelnevalt avaldatud.
- [33] Jongyun Hwang *et al.*, „SAM2 for Abdomen: One-shot and No Finetuning“. Teoses: 2024 9th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS). 2024 9th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS). Köide 9. November 2024, lk. 551–555. doi: 10.1109/ICIIBMS62405.2024.10792815. URL: <https://ieeexplore.ieee.org/document/10792815/> (vaadatud 16.05.2025).

Lisa 1 – Lihtlitsents lõputöö reproduutseerimiseks ja lõputöö üldsusel kättesaadavaks tegemiseks¹

Mina, Kristjan Luik

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose ‘Met-saraie tuvastamine rakendades arvutinägemise alusmudeleid’, mille juhendaja on Juhan-Peep Ernits
 - 1.1. reproduutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaaja lõppemiseni;
 - 1.2. üldsusel kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaaja 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.

26.05.2025

¹Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reproduutseerida 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 tähtajaks nõusolekut lõputöö reproduutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaaja jooksul ei kehti.

Lisa 2 – Vektor maskide loomine kasutades QGISi

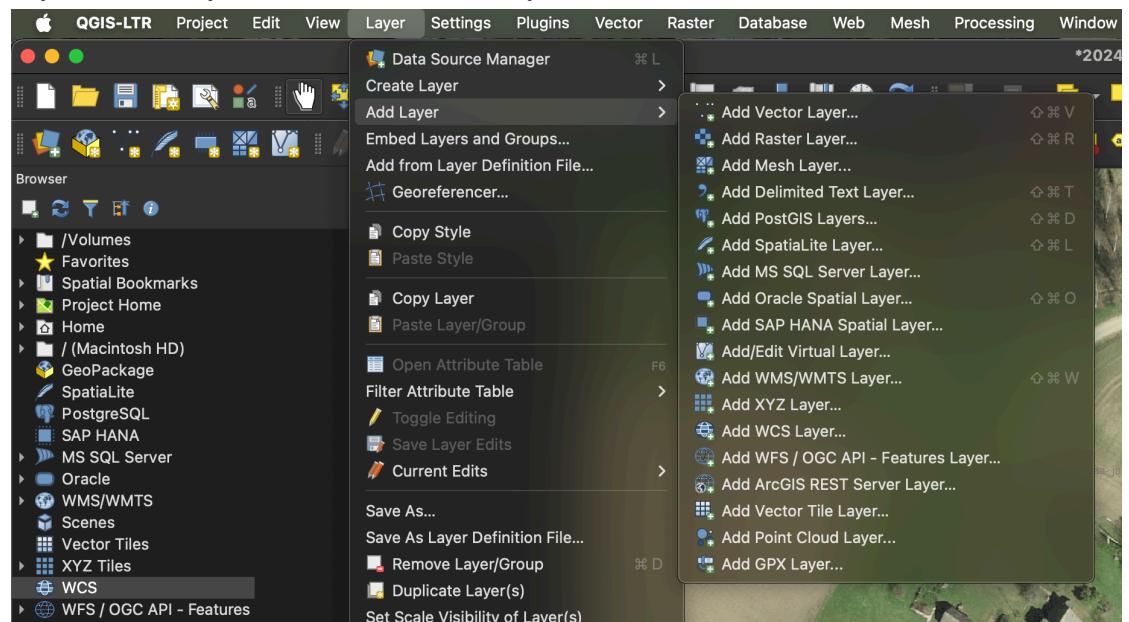
Vektor maskide loomine kasutades QGISi

Abi kihtide lisamine

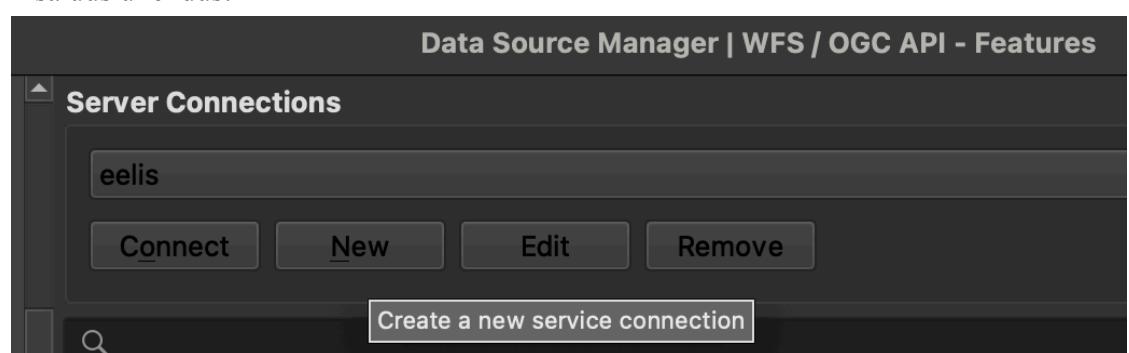
1) Lae alla QGIS: <https://www.qgis.org/>

2) Lisa metsaeraldiste kiht:

a) Layer > Add Layer > WFS API feature layer:

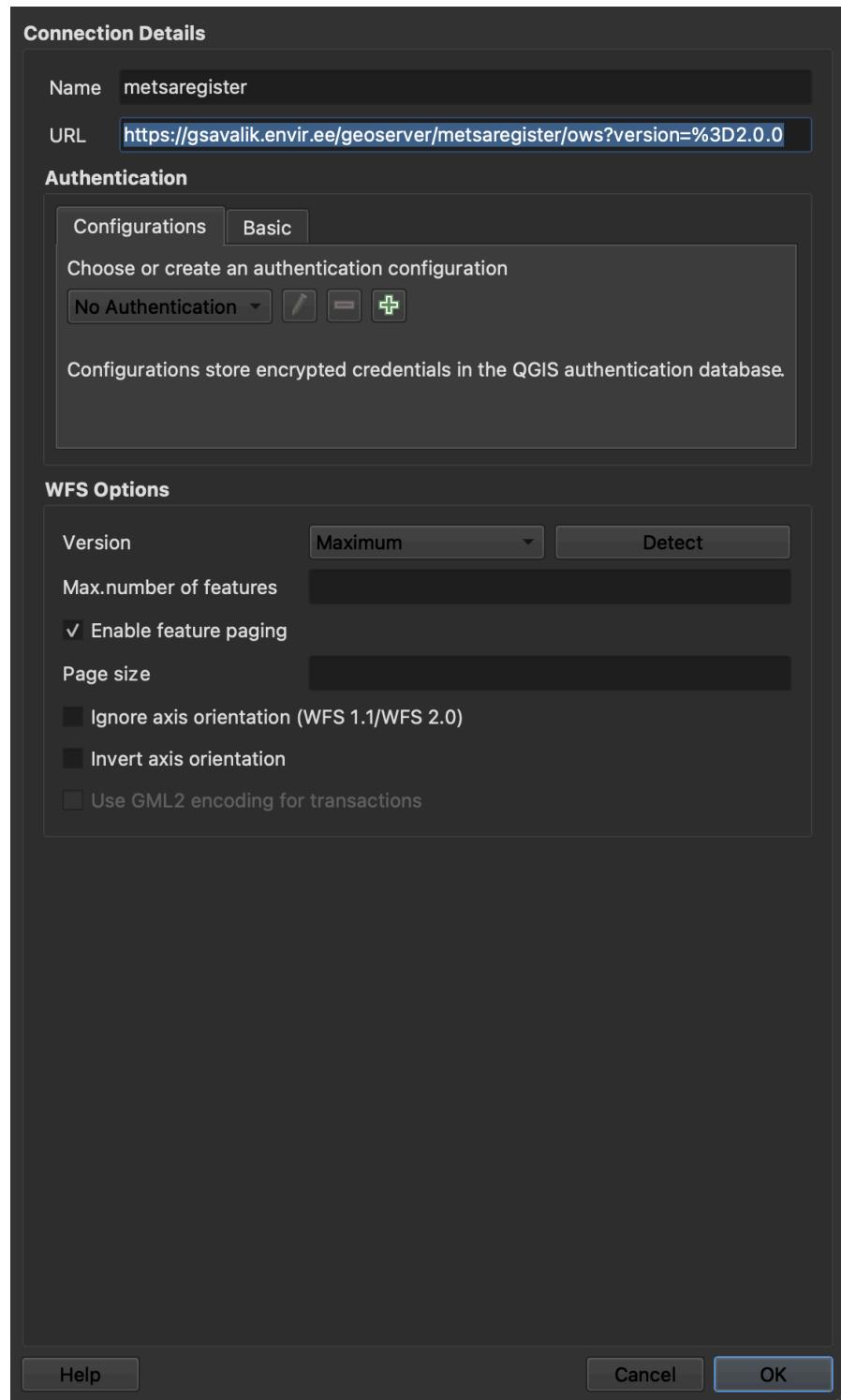


b) Lisa uus ühendus:



- c) Lisa serveri URL ja nimi:

URL: <https://gsavalik.envir.ee/geoserver/metsaregister/ows?version=%3D2.0.0>



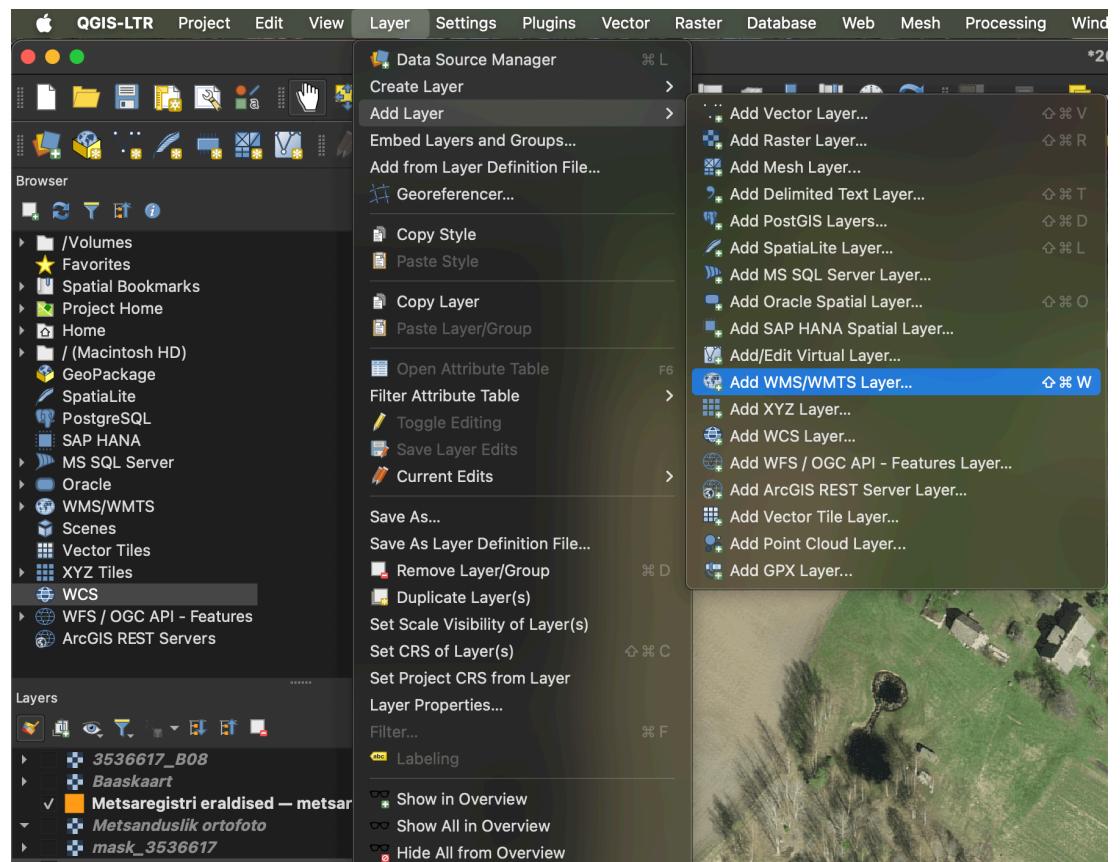
- d) Vajutage connect

e) Lisage aktiivsete metsaeraldiste kiht:

Title	Name	Abstract	Sql
Arhiveeritud metsakaitseekspertiisi (MKE) ...	metsaregister:mke_arhiiv	Avalikud metsateatist...	
Arhiveeritud avalikud metsateatise andmed	metsaregister:teatis_arhiiv	Metsaregistris kehtiva...	
Iseärasused	metsaregister:isearasused	Metsaregistris kehtiva...	
Kahjustused	metsaregister:kahjustused	Metsaregistris kehtiva...	
Katsealad	metsaregister:katsealad	Metsaregistri kehtiva...	
Kehtivad metsateatised	metsaregister:teatis	Metsaregistri kehtiva...	
Klassifikaator: Metsa kasvukohatüübidi	metsaregister:kl_kasvukoht	Raiesovitusega mets...	
Klassifikaator: Omandidorm	metsaregister:kl_omandidorm		
Klassifikaator: Puistu rindad	metsaregister:kl_rinne		
Klassifikaator: Puuligid	metsaregister:kl_puulik		
Metsa uuendamise eksperitiidid MUE	metsaregister:mue	Metsa uuendamise ek...	
Metsakaitseekspertiisid kuuse-kooreüraski ...	metsaregister:kuusekooreyrask_mke	Raiesovitusega mets...	
Metsaregisti eraldised	metsaregister:eraldis	Metsaregistri kehtiva ...	Metsaregistris kehtiva...
Metsaregisti eraldise puistu koodseis	metsaregister:eraldis_element	Metsaregistris kehtiva...	
mr __ teema_raie_diameteer	metsaregister:mr __ teema_raie_diametee..		
mr __ teema_raie_liik	metsaregister:mr __ teema_raie_liik		
mr __ teema_raie_talus	metsaregister:mr __ teema_raie_taius		
mr __ teema_raie_vanus	metsaregister:mr __ teema_raie_vanus		
Natura 2000 alad	metsaregister:natura_2000_alad	Koondkiht Natura 20...	
Raie soovitusega metsakaitseekspertiisid	metsaregister:mke	Metsaregistri raie so...	
ÜP Kaitsemetsad	metsaregister:yp_kaitsemets	Üldplaneeriguga keht...	

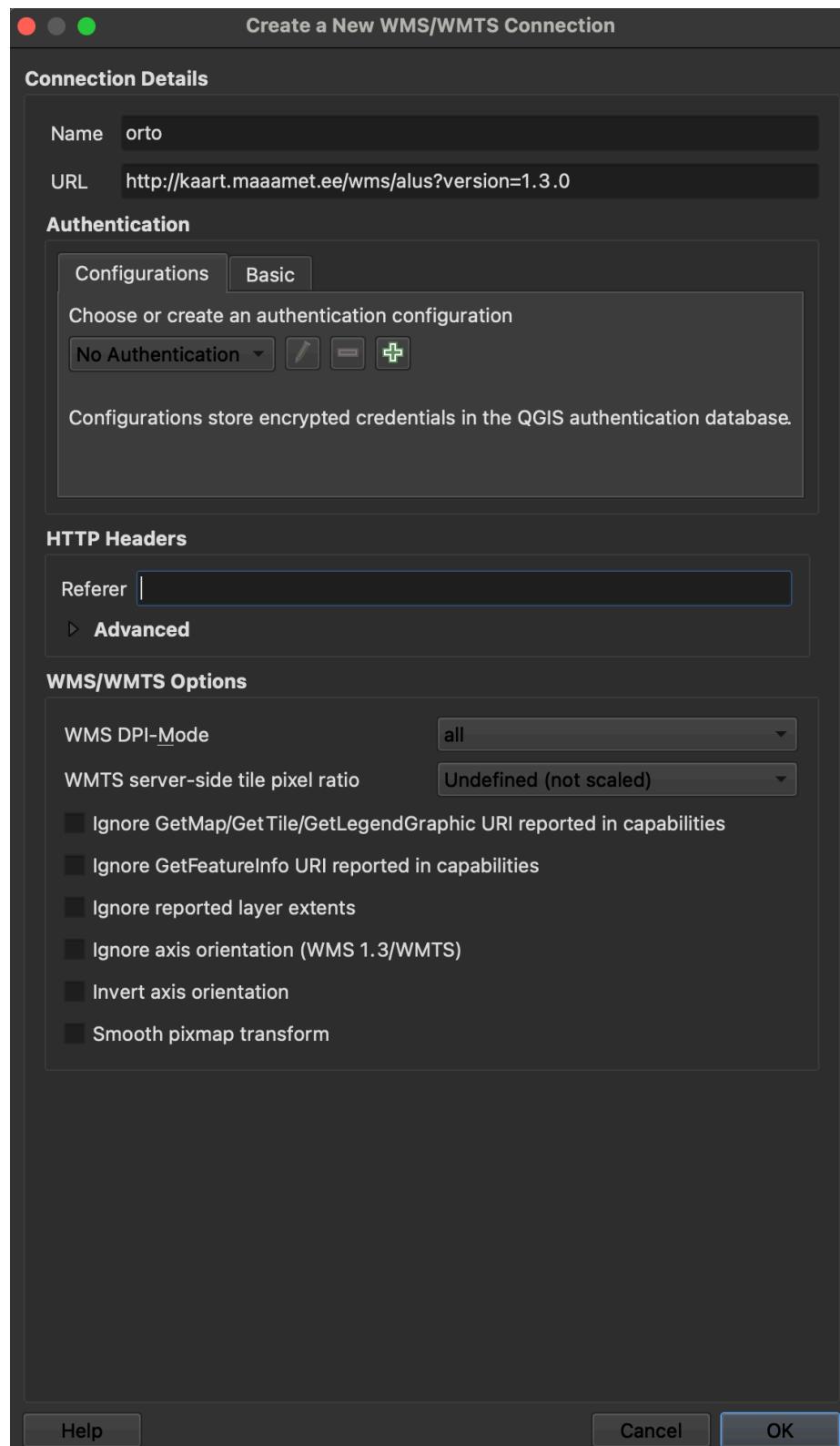
3) Lisa ortofotode kiht

a) Layer > Add layer > new WMS layer

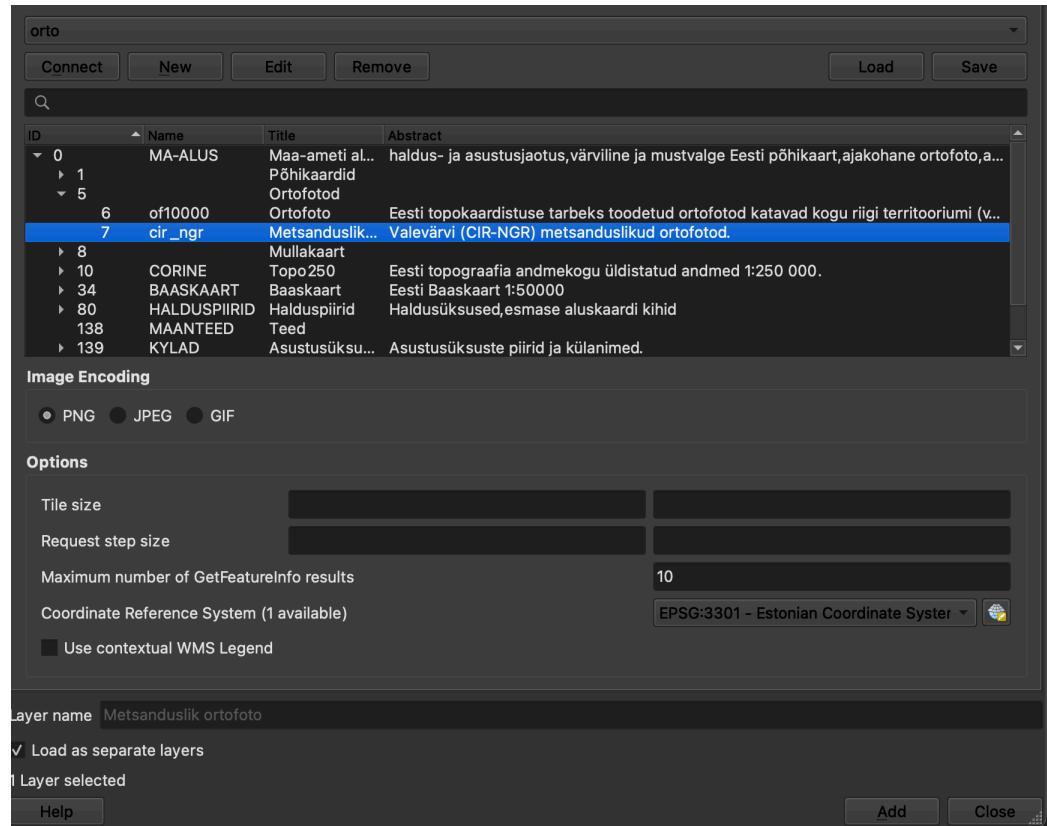


- b) Lisa serveri URL ja nimi:

URL: <http://kaart.maaamet.ee/wms/alus?version=1.3.0>

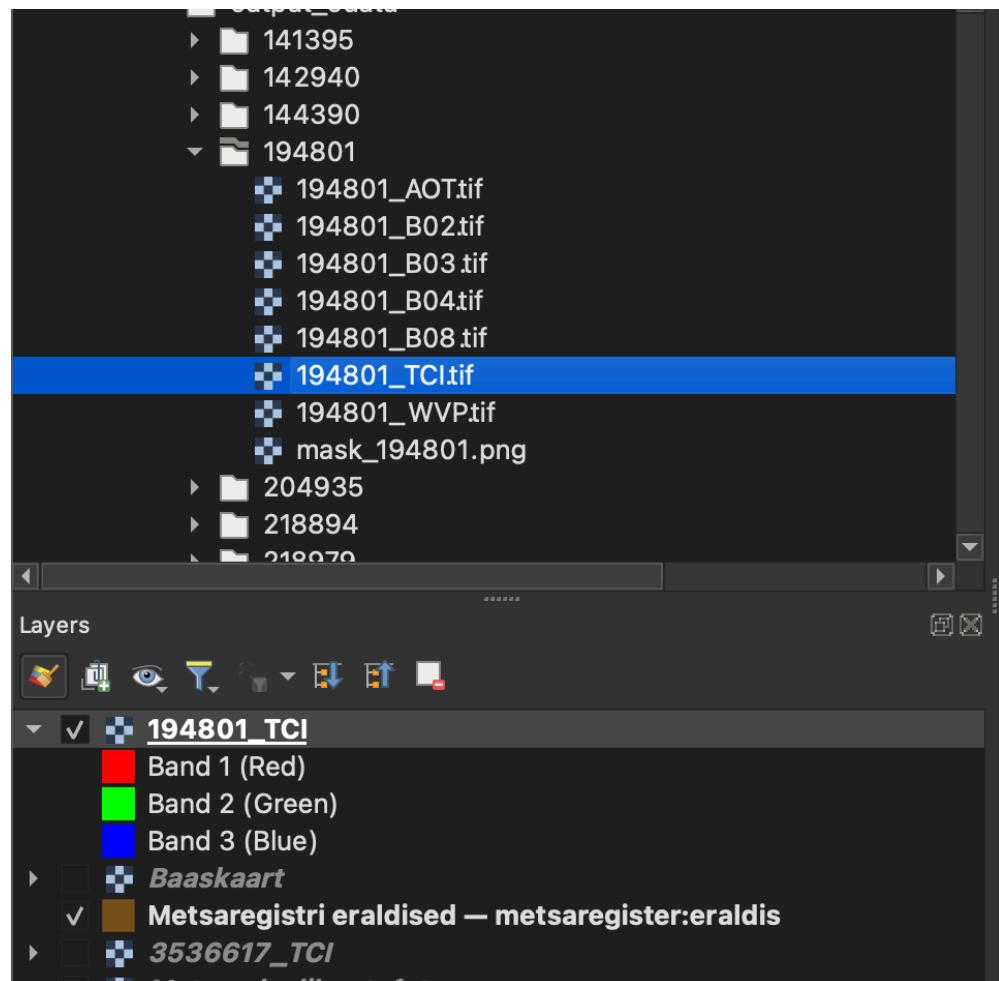


- c) Saadud nimekirjast valida Metsanduslik ortofoto:



Satelliidi piltide lisamine

- 1) Tõsta lõigatud satelliidipildid projektile



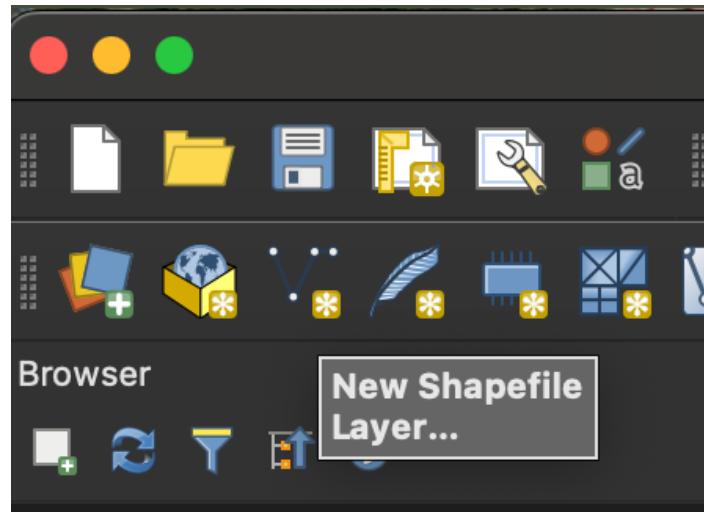
a)

- b) Et satelliidi pildile fokuseerida vajutage parem klõps pildi kihil ja valige zoom into layers:



Vektorite loomine (ehk maskide loomine)

- 1) Loome uue shapefile, mis hakkab hoidma kõiki märgendatud polügoone:



a)

b) See sisaldab kolme välja:

Field	Type
id	Integer64
type	String
satImage	Integer64

id = polügooni järjekorranumber (märkida igale polügoonile uus number)

type = see on maski klass ja peab kuuluma ühte (“forest”, “clearcut”, “y_forest”)

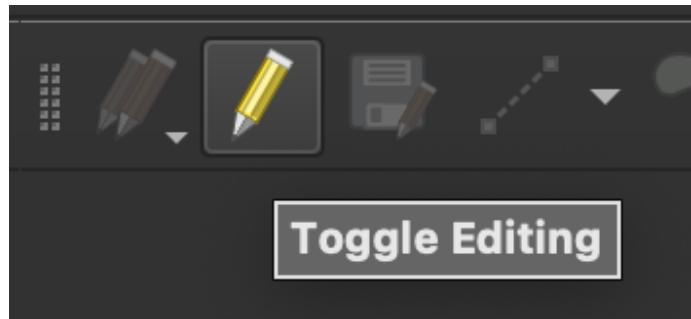
satImage = lingerie ID millele see poligon kuulub (sama mis number satelliidipildi nimes)

New Shapefile Layer

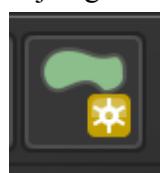
File name	/Users/kristjanluik/Desktop/krissu_ano.shp		
File encoding	UTF-8		
Geometry type	Polygon		
Additional dimensions	<input checked="" type="radio"/> None <input type="radio"/> Z (+ M values) <input type="radio"/> M values		
Project CRS: EPSG:3301 – Estonian Coordinate System of 1997			
New Field			
Name	[]		
Type	123 Integer (32 bit)		
Length	10		
Precision			
<input type="button" value="Add to Fields List"/>			
Fields List			
Name	Type	Length	Precision
id	Integer	10	
type	String	80	
satImage	Integer	10	

c)

d) Lülitage sisse töötlemine (editing):



e) Vajutage uue polügooni tegemise ikoonile:

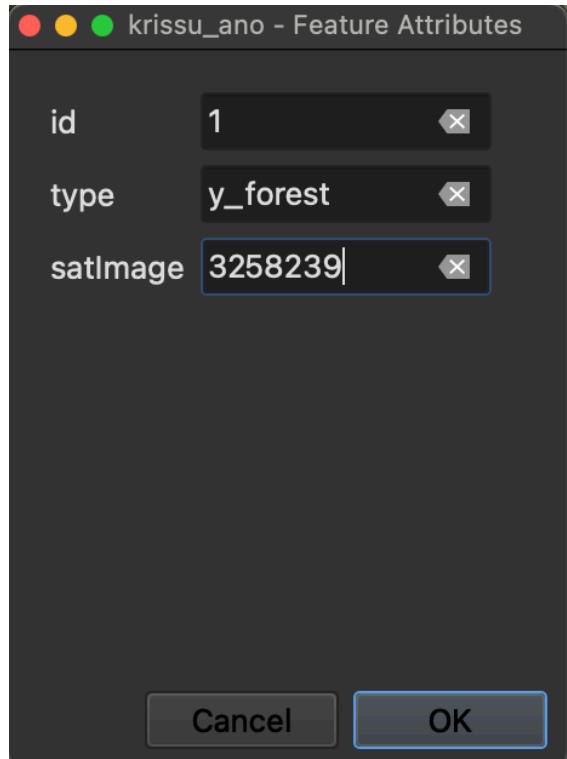


f) Looge haripunkte (vertex) millest polügoon koosneb:



(parem click et polügooni loomine lõpetada)

g) Sisestage polügooni väljad:



h) Polügoon on loodud.

Vektormaskide klasside tuvastamine

Olulised põhimõtted:

Klasside loomisel lähtuda õige ajavahemiku järgi valitud satelliidipildist.

Määramisel ja piiride seadmisel on abiks **ortofotod**.

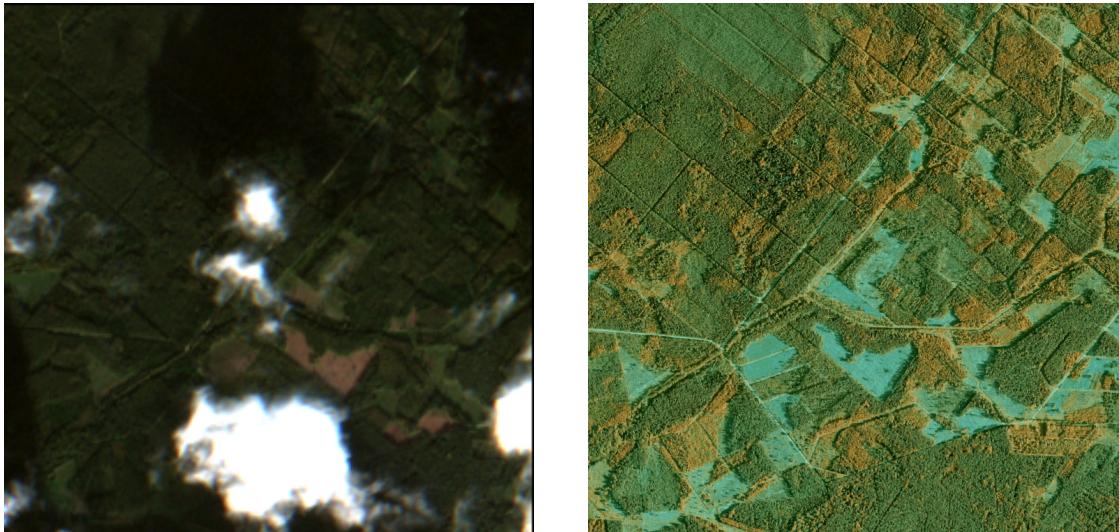
Kui ortofotol on näha raiet, kuid satelliidipildil juba noort metsa, siis lähtuda **satelliidipildist**.

Vaja on märkida vaid metsateatiste registris olevad alad.

Pilvedega kaetud alad jäätta märkimata.

Kui te pole milleski kindel, siis jäätta ala märkimata.

Järgnevalt on välja toodud näidispildid ja igast klassist üks näide.

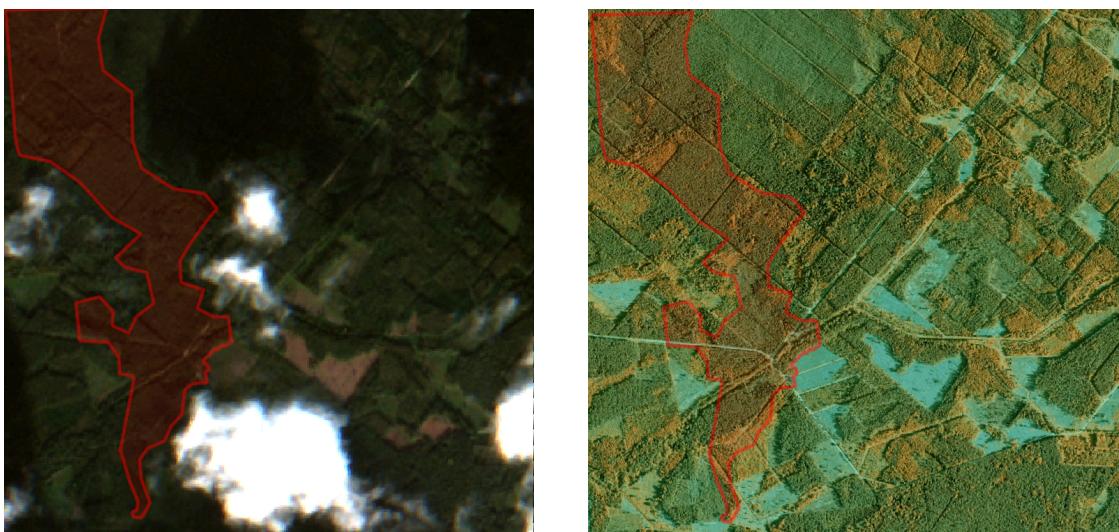


Klassid ja nende tunnused:

1) Mets (“forest”)

Satelliidipiltidel on mets ühtlast **tumedamat rohelist** värv, mis tuleneb taimestiku tihedusest. Metsa alad võivad piltidel olla **ebaühtlase** tekstuuriga, kuna metsa moodustavad puud on eri kõrguste ja võra kujudega. Metsaservad on tihti **ebakorrapärased** ja järgivad maastikku, seega ei paista metsad piltidel korrapäraste kujunditena.

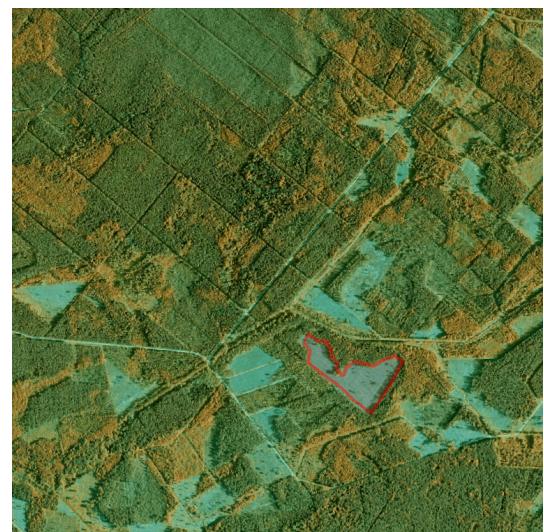
CIR-NGR ortofotodel on mets rohelist ja rohekas-kollast värv. Ortofotodel on suumides **näha puid**.



2) Lageraie (“clearcut”)

Satelliidipiltidel on lageraie alad **heledamates** toonides kui mets. Lageraie alad on kas **pruunid, helepruunid, beežid** või **hallikad**. Raiealade tekstuur on metsade omast ühtlasem, eriti kohe pärast raiet. Võib näha inimtegevusele viitavaid mustreid, raieteid või alles jäänud prahti. Lageraie on metsaga vörreledes kujult **korrapärasem** ja piiritletum, võib näha **sirgeid jooni** ja teravaid nurki.

CIR-NGR ortofotodel on lageraie ala sinakas-rohelist **tsüaan** tooni. Ortofotodel on suumides näha **lagedat** ala, kus on alles üksikud puud nö seemneks.



3) Noor mets (“y_forest”)

Satelliidipiltidel on noor mets **helerohelist** värv, erksamat tooni kui muidu mets ja rohelisem kui lageraie. Pärast raiet hakkab taimestik taastuma ja muutub vaikselt heledamaks laiguliseks roheliseks, sellisel juhul märkida ala juba noore metsa klassi. CIR-NGR ortofotodel on noor mets **kollakas-rohelist** värv. Ortofotodel on lähemalt näha vaid madalamaid puid.

