

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond

Kristjan Luik 211809IAPM

Eesti metsaraie tuvastamine masinõppe meetoditega

Magistritöö

Juhendaja: Juhan-Peep Ernits

PhD

Tallinn 2025

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: Kristjan Luik

05.03.2025

Annotatsioon

[ANNOTATSIOONI TEKST LÄHEB SIIA]

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 7 leheküljel, 5 peatükki.

Abstract
Thesis Title

[YOUR TEXT GOES HERE]

The thesis is in Estonian and contains 7 pages of text, 5 chapters.

Lühendite ja mõistete sõnastik

API	Rakendusliides (<i>Application Programming Interface</i>)
CPU	Keskseade (<i>Central Processing Unit</i>)
IDE	Integreeritud programmeerimiskeskkond (<i>Integrated Development Environment</i>)
IOT	Asjade Internet (<i>Internet Of Things</i>)
VM	Virtuaalmasin (<i>Virtual Machine</i>)

Sisukord

1	Sissejuhatus.....	9
2	Valdkonna ülevaade	11
2.1	Metsandus.....	11
2.2	Copernicus ja EstHub	11
2.2.1	Sentinel.....	11
2.2.2	Lainepikkuste spekter	12
2.3	Masinõppe meetodite kasutus kaugseires	12
3	Lahendus	13
3.1	Töövahendid.....	13
3.2	Andmestiku loomine	13
3.3	Mudelite võrdlus.....	13
4	Tulemuste analüüs	14
5	Kokkuvõte.....	15
	Kasutatud kirjandus	16
	Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks.....	18
	Lisa 2 – Something	19
	Lisa 3 – Something Else	20

Jooniste loetelu

Tabelite loetelu

1 Sissejuhatus

Antud magistritöö põhieesmärgiks on võrrelda masinõppe meetodeid ja tuua välja täpseim mudel, mis suudaks tuvastada metsaraieid satelliidipiltidelt. Mida aeg edasi seda rohkem on riigid hakanud mõistma kui tähtis on metsamajandus, metsade säilitamine ja hoidmine. Tehnoloogia pideva arenguga on hakatud ka 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, vabatahtlike kaasavaid rakendusi (Crowdsourcing Applications) [1].

Praegusel hetkel kasutatakse Eestis mõni aasta tagasi Keskkonnaagentuuri ja Tartu Ülikooli koostöös väljatöötatud statistika mudelit, mis raie tuvastamiseks kasutab suvasalu (Random forest) algoritmi [2] satelliidi piltidelt. Selle mudeli esmased tulemused olid paljulubavad, aga peale mõndaaegset kasutamist pole see ikkagi rahuldavaid tulemusi andnud ja mudeli kasutajad on sunnitud siiski manuaalseid viise kasutama.

Euroopa Liidu kaugseireprogrammt Copernicus võimaldab Eesti riigil koguda satelliidi pilte andmekeskusesse Esthub [3]. Lisaks muule infole, mida hallatakse Copernicus-es ja seeläbi ka Esthub-is, on kasutusel informatsioon mis tuleb erinevatelt Sentineli nime kandvatelt satelliitidelt [4]. Kuna Sentinel-2 on juba 2015. aastast töös olnud, sisaldab laia valikut valgusribasid ning on tiheda korduskülastus sagedusega [5], siis keskendub käesolev magistritöö peamiselt sellele satelliidi tüübile.

Sellest tulenevalt ona üheks alam eesmärgiks luua Python programm, mis hõlbustaks satelliidi piltide allalaadimist ja töötlemist. Peale andmete kogumist on plaan läbi viia tänapäevaste masinõppe mudelite võrdlus raiete tuvastamiseks. Raiet hinnatakse piksli põhise täpsusega üle pildi. Hiljuti on tehtud mitmeid uuringuid selles valdkonnas, kus kasutatakse ka suvasalu, XGBoost ja U-Net'il põhinevaid mudeli arhitektuure [6], [7]. Mõlemas uurimistöös on ka mudelite võrdlus välja toodud, aga need keskenduvad erinevatel suundadel. Esimese puhul ehitatakse mudelid kasutades rohkem pilte läbi aja, et mudel

saaks paremini tuvastada muutust. Teise puhul keskendutakse erinevate lainepikkuste kombineerimisele, et tabada muutusi.

Peale mudelite treenimist samadelt lähteandmetelt on antud magistritöös välja toodud tulemuste mõõtmine. 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.

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 saamiseks ning umbes 70 miljonit, sealhulgas paljud põlisrahvad, peavad metsi oma koduks [10]. Metsad varustavad meid hapnikuga, pakuvad varjualust, töökohti, puhast 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

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 kosmosesest. See info on kõigile kättesaadav tasuta. Selle programmiga seotud satelliit kutsutakse **Sentineliks**. [12]

2.2.1 Sentinel

Sentinel-1 on radaripõhine satelliit, mis võimaldab jälgida maapinna vajumist, struktuuride kahjustusi ning geohazarde nagu maavärinad ja maalihked. Samuti on see ideaalne mere- ja Arktika seireks, sealhulgas laevade jälgimiseks ning naftareostuse tuvastamiseks. [13]

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. [14]

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. [15]

Sentinel-5P on esimene Copernicuse missioon, mis on pühendatud atmosfääri seirele. See kannab tipptasemel **Tropomi** instrumenti, mis kaardistab mitmeid gaase nagu lämmastikdioksiid, osoon, formaldehüüd, vääveldioksiid, metaan, vingugaas ja aerosoolid - kõik need mõjutavad meie hingatavat õhku, tervist ja kliimat. [16]

2.2.2 Lainepikkuste spekter

2.3 Masinõppe meetodite kasutus kaugseires

3 Lahendus

3.1 Töövahendid

3.2 Andmestiku loomine

3.3 Mudelite võrdlus

4 Tulemuste analüüs

third

5 Kokkuvõte

summ

Kasutatud kirjandus

- [1] 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).
- [2] 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-statistilised-meetodid-eesti-metsaressursi-hindamiseks> (vaadatud 11.01.2025).
- [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. Podoprigozova *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] *S1 Applications*. URL: <https://sentiwiki.copernicus.eu/web/s1-applications> (vaadatud 04.03.2025).
- [14] *S2 Applications*. URL: <https://sentiwiki.copernicus.eu/web/s2-applications> (vaadatud 04.03.2025).
- [15] *S3 Mission*. URL: <https://sentiwiki.copernicus.eu/web/s3-mission> (vaadatud 04.03.2025).
- [16] *S5P Applications*. URL: <https://sentiwiki.copernicus.eu/web/s5p-applications> (vaadatud 04.03.2025).

Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹

Mina, Kristjan Luik

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose “Eesti metsaraie tuvastamine masinõppe meetoditega”, mille juhendaja on Juhan-Peep Ernits
 - 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 veebikeskonna 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.

05.03.2025

¹Lihtlitsents 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 tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.

Lisa 2 – Something

```
<!DOCTYPE html>  
<html>  
<body>  
  
<h1>Example Title </h1>  
  
<p>Some text here </p>  
  
</body>  
</html>
```

Lisa 3 – Something Else

Pythagorean theorem

$$x^n + y^n = z^n \tag{1}$$

Normal distribution

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{2}$$