

TARTU ÜLIKOOL  
Arvutiteaduse instituut  
Informaatika õppekava

**Johanna Kaare**  
**Heliklipist signaali eraldamine ja minimaalse  
treeningandmestiku suuruse tuvastamine  
olmehelide klassifitseerimiseks**

**Bakalaureusetöö (9 EAP)**

Juhendajad:  
Tiit Sepp, MSc  
Tauno Palts, PhD

Tartu 2025

## **Heliklipist signaali eraldamine ja minimaalse treeningandmestiku suuruse tuvastamine olmelide klassifitseerimiseks**

### **Lühikokkuvõte:**

Kuulmislangusega inimesed ei pruugi kuulda olulisi helisid, mis võib raskendada nende igapäevast toimetulekut. Selleks on neile vajalikud helituvastustööriistad. Enne helide tuvastamist on vajalik signaali ja taustamüra sisaldavatest heliklippidest signaalid eraldada ning määrata helituvastuseks vajalik minimaalne treeningandmestiku suurus. Heliklipist signaali eraldamiseks töötati siinses töös välja uus algoritmi, mis jagunevad amplituudipõhisteks ja sageduspõhisteks. Kõik loodud algoritmid eraldasid heliklippidest signaali paremini kui Google'i loodud kõne aktiivsuse tuvastamise meetod WebRTC. Lisaks leiti Studenti t-testi põhjal, et minimaalne helifailide arv YAMNetil põhineva helituvastusmudeli ümbertreenimiseks on 10.

**Võtmesõnad:** Helide klassifitseerimine, heliklipist signaali eraldamine, minimaalne treeningandmestiku suurus

**CERCS:** P170 – Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine (automaatjuhtimisteooria); T121 – Signaalitöötlus

# **Extracting Signals from Audio Clips and Determining the Minimum Training Dataset Size for Household Sound Classification**

## **Abstract:**

People with hearing impairments may not hear important sounds, which can make daily life more challenging. For this reason, they require sound detection tools. Before detecting sounds, it is necessary to separate the signal from background noise in audio clips and determine the minimum size of the training dataset needed for sound detection. In this work, six algorithms were developed for extracting signals from audio clips, categorized into amplitude-based and frequency-based methods. All of the developed algorithms performed better at extracting signals from audio clips than Google's voice activity detection method WebRTC. Additionally, based on the Student's t-test, it was found that the minimum number of audio files required to retrain a YAMNet-based sound detection model is 10.

**Keywords:** Audio classification, signal extraction from an audio clip, minimum training set size

**CERCS:** P170 – Computer science, numerical analysis, systems, control; T121 – Signal processing

# **Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks**

Mina, Johanna Kaare,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose "Heliklipist signaali eraldamine ja minimaalse treeningandmestiku suuruse tuvastamine olmelide klassifitseerimiseks", mille juhendajad on Tiit Sepp ja Tauno Palts, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada Tartu Ülikooli digitaalarhiivi kuni autoriõiguse kehtivuse lõppemiseni;
2. annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi kaudu Creative Commons'i litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, alates 20.11.2030 kuni autoriõiguse kehtivuse lõppemiseni;
3. olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile;
4. kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Johanna Kaare

21.11.2025