

Sound recognition by Home AI assistant



30.10.2019
Katrina Bykova

Enhancing human interaction interface

- Alexa and Google-home recognize voice commands when called
 - Current home AI systems do not “see” the home environment
 - The interaction with a human can be improved by adding environmental sensing
-
- Can AI use sound to identify human activities?

Data source: Google AudioSet dataset

- Collection of human-labeled 10-second sound clips from YouTube videos
- **2,084,320** recordings
- **527** sound classes:
 - speech
 - music
 - footsteps
 - ...
- Features preprocessed by VGGish acoustic model are available

Home-related activities: 10 sound classes

Cooking/cleaning

- water
- microwave
- blender
- vacuum cleaner

Entertainment

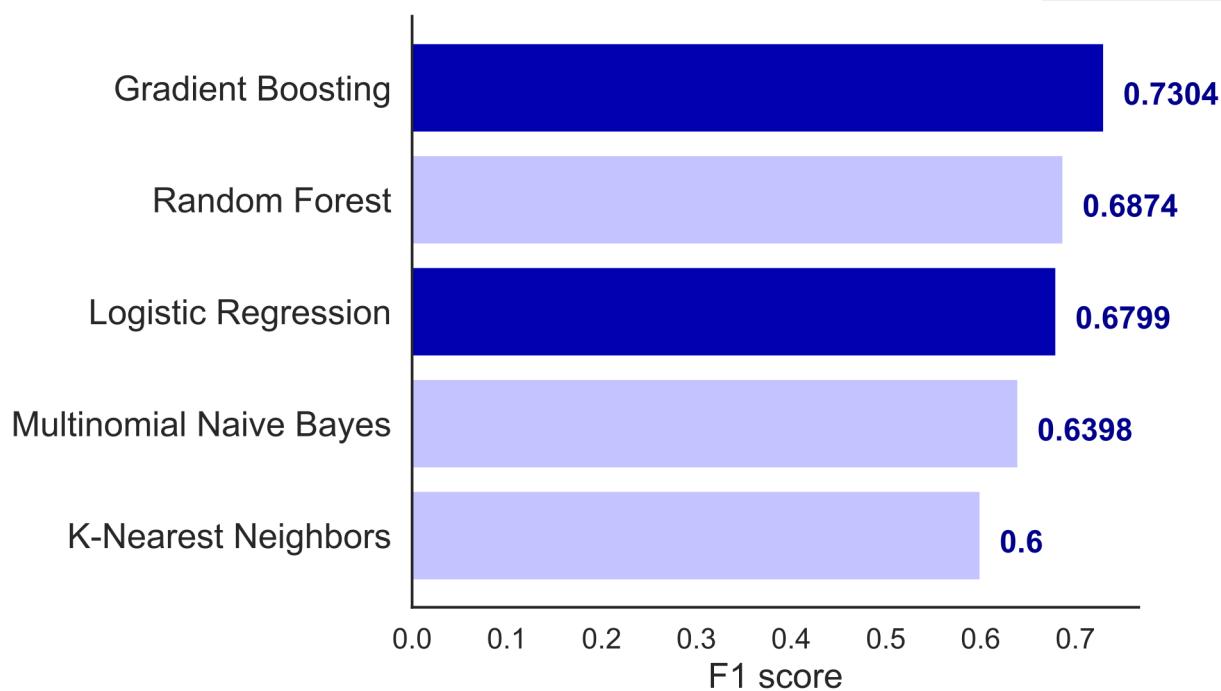
- music
- speech
- clarinet
- cat

Moving around

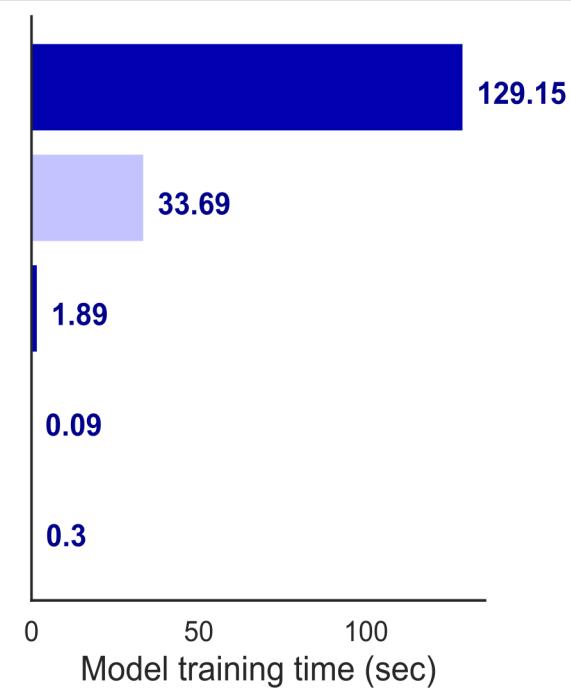
- footsteps
- opening doors

Sound recognition models

Model performance

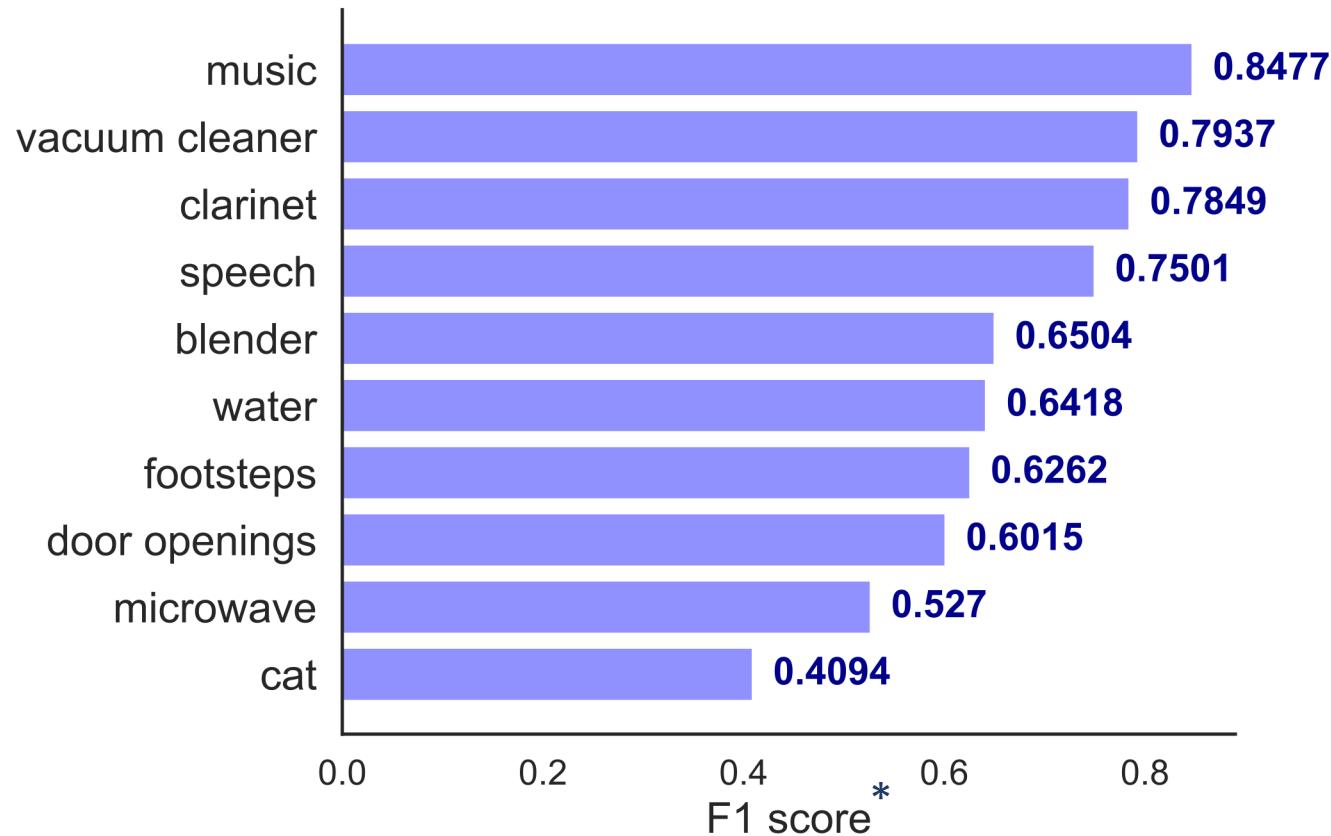


Model training time



- Gradient Boosting - best performance, longest model training time
- Logistic Regression - less optimal performance, short model training time

Recognition of individual sound classes



* Gradient Boosting model

Recognition of home activities: examples



Recognition of home activities: examples



vacuum cleaner ✓



clarinet ✓

← model classification

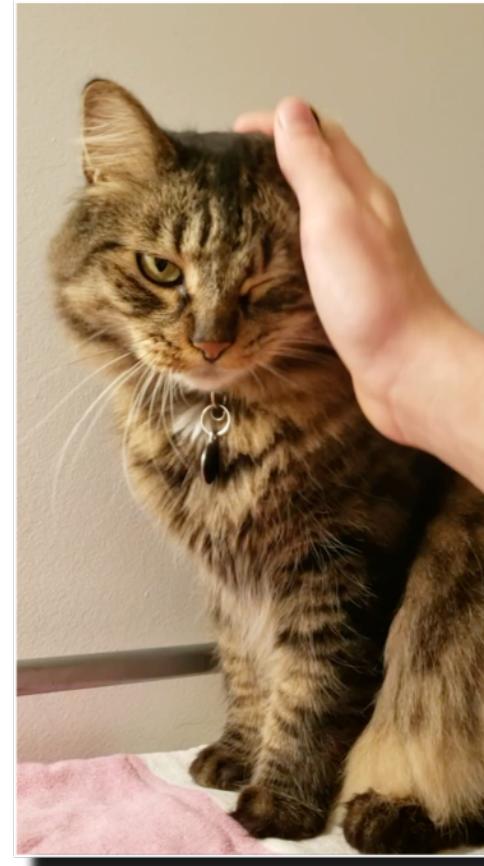
Recognition of home activities: examples



Recognition of home activities: examples



cat ✓



cat ✗ → human voice

- Training on "human making cat sound" and "cat meow" sounds would help with the classification

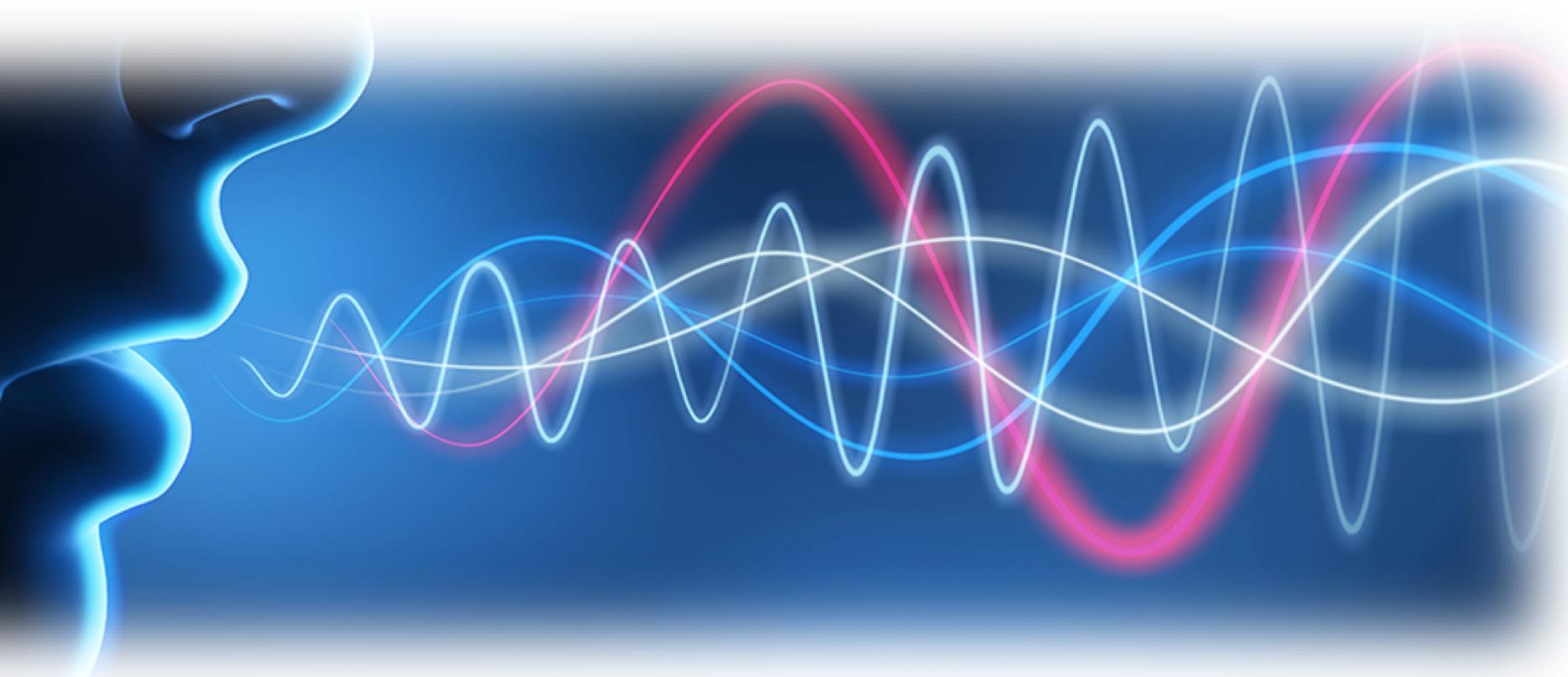
Conclusions

- Home-related activities can be identified by AI through sound recognition
- Recommendations:
 - Gradient Boosting model for the best performance
 - Logistic Regression for faster modeling time

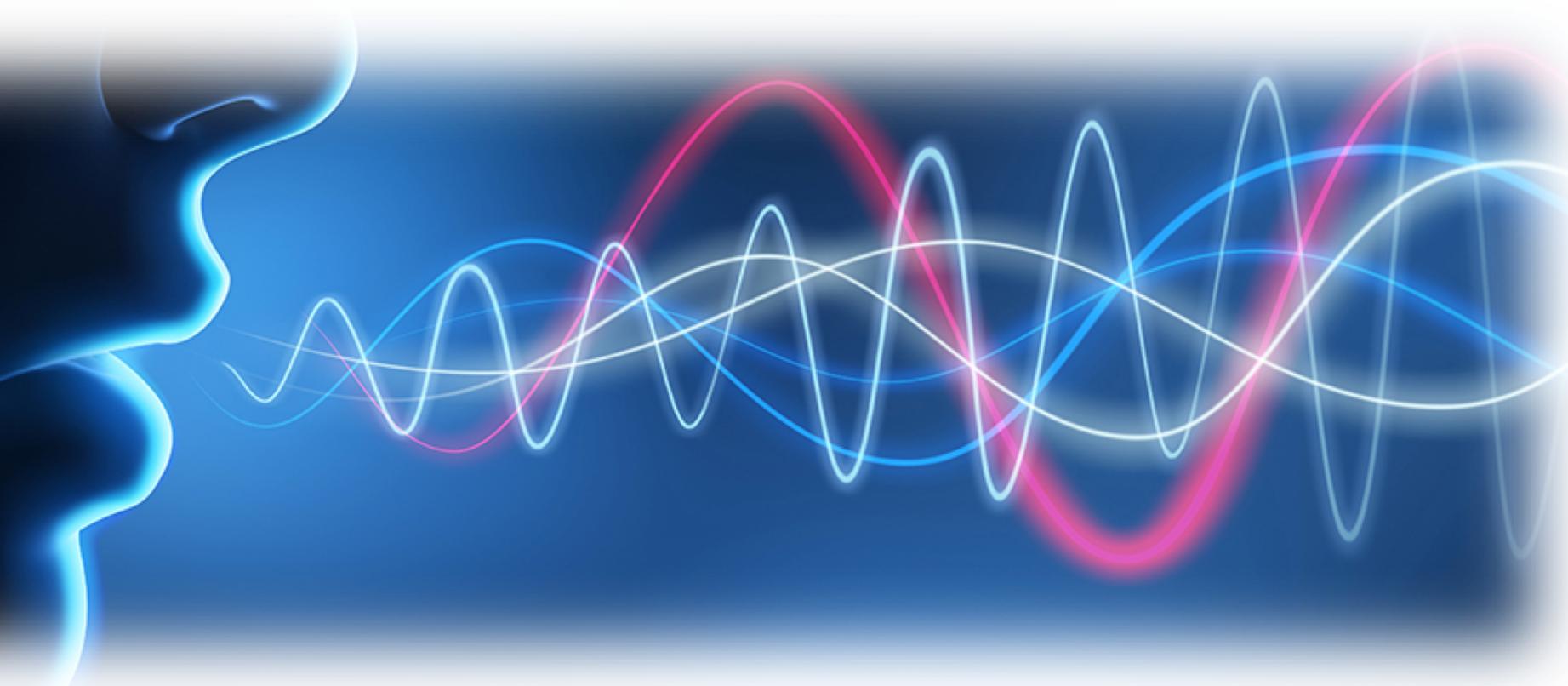
Future development

- Build Alexa skills
 - Children school morning AI assistant:
 - No “microwave sound” between 6:00 AM and 6:15 AM →
AI: “Have you packed your lunch?” Repeat every 5 minutes.
 - Repetitive “cat” sound in the morning →
AI: “Please feed the cat.” Repeat until the “cat” sound is no longer registered.
- Improve sounds recognition algorithm
 - Expend sound library
 - Develop alternative sound processing models

Thank you



Additional slides



Home-related activities: 10 sound classes

Cooking/cleaning

- water tap
- microwave
- blender
- vacuum cleaner

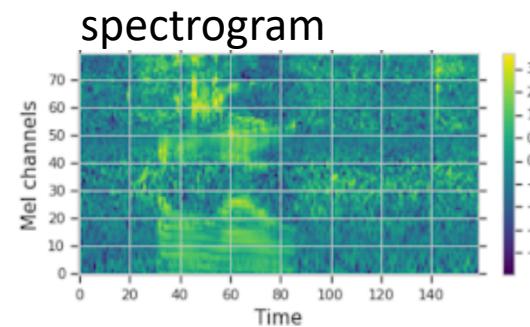
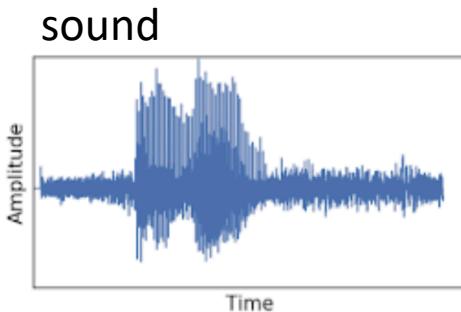
Entertainment

- music
- speech
- clarinet
- cat

Moving around

- footsteps
- opening doors

Features:



VGGish neural
networks model

1,280
features

10 sound classes: observations

speech	4042
music	3781
vacuum_cleaner	3054
blender	1884
door	1868
footsteps	1492
clarinet	960
microwave	894
water_tap	458
meow	388

Recognition of home activities: examples



vacuum
cleaner ✓



clarinet ✓

Recognition of home activities: examples



Recognition of home activities: examples



Recognition of home activities: examples



microwave



blender

- Adding additional observations to “microwave” and “blender” classes might help with the classification

← model classification