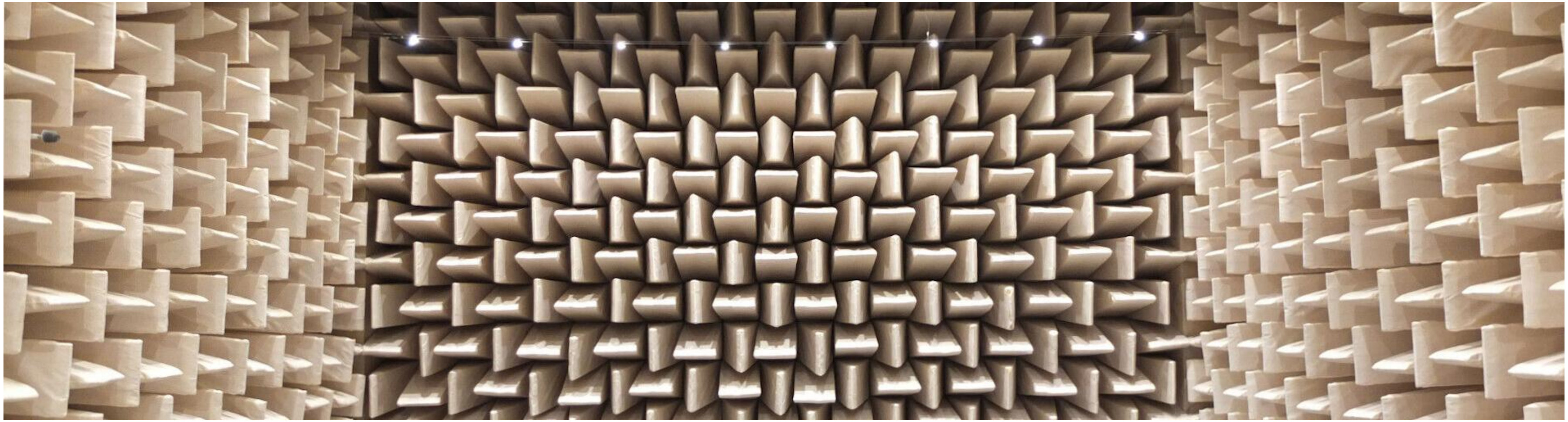
DAS | DAGA 2025
COPENHAGEN

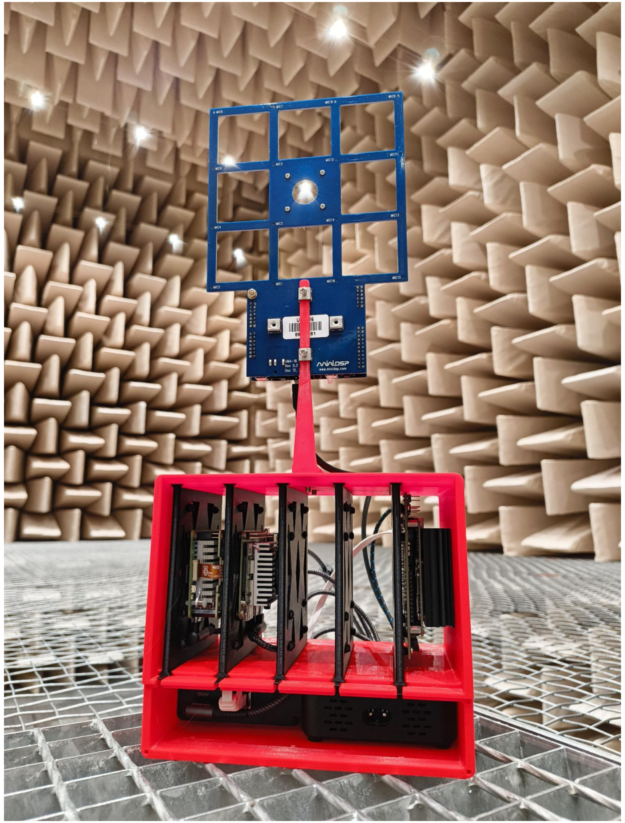
TECHNISCHE UNIVERSITÄT BERLIN



Comparison of Embedded Hardware Platforms for Optimized Machine Learning-Based Acoustic Imaging

Jakob Tschavoll | Engineering Acoustics

Contents



Motivation

Acoustic cameras for **industrial** and **urban** noise monitoring

https://www.youtube.com/@SoramaSoundImaging

Problem:

> 10.000 €

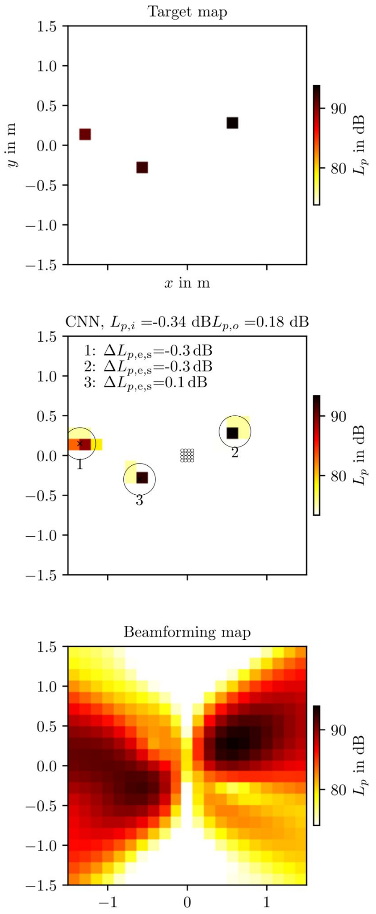
Motivation
Questions

1. Can the price be reduced to 10% or less?
2. Can such a device be made more available?
3. Can low-end devices perform this task?

Motivation

Advances in ML-based Acoustic Imaging

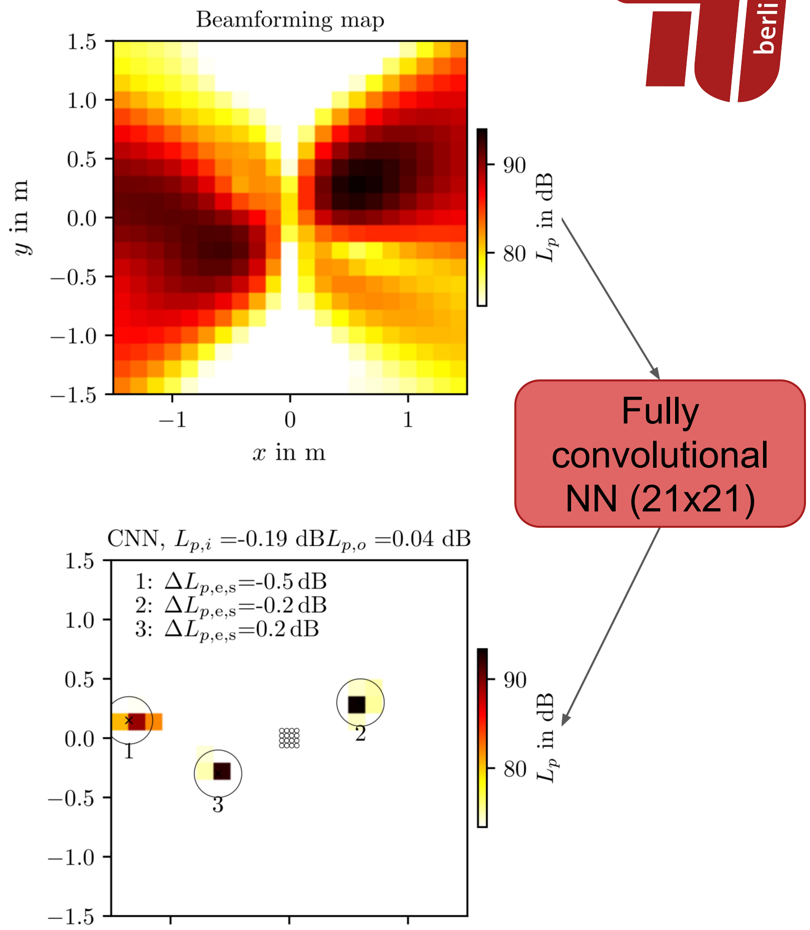
- DNN: Castellini et al. (2021)
- CNN: Ma and Liu (2019),
  Pinto et al. (2021),
  Pasha et al. (2021)
- SVM: Salvati et al. (2016)
- Other: Lee et al. (2022),
  Rashida et al. (2023),
  Kujawski and Sarradj (2022)



Fundamentals
Model Input

Approach by Pinto et al.:

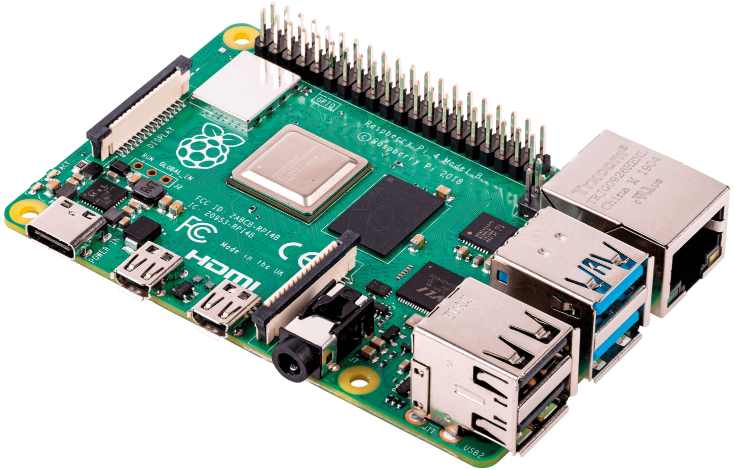
- Calculate low resolution beamformer from CSM
- Deconvolute map (image processing)
- Output quasi-sparse locations and source strengths



Fundamentals
Constraints

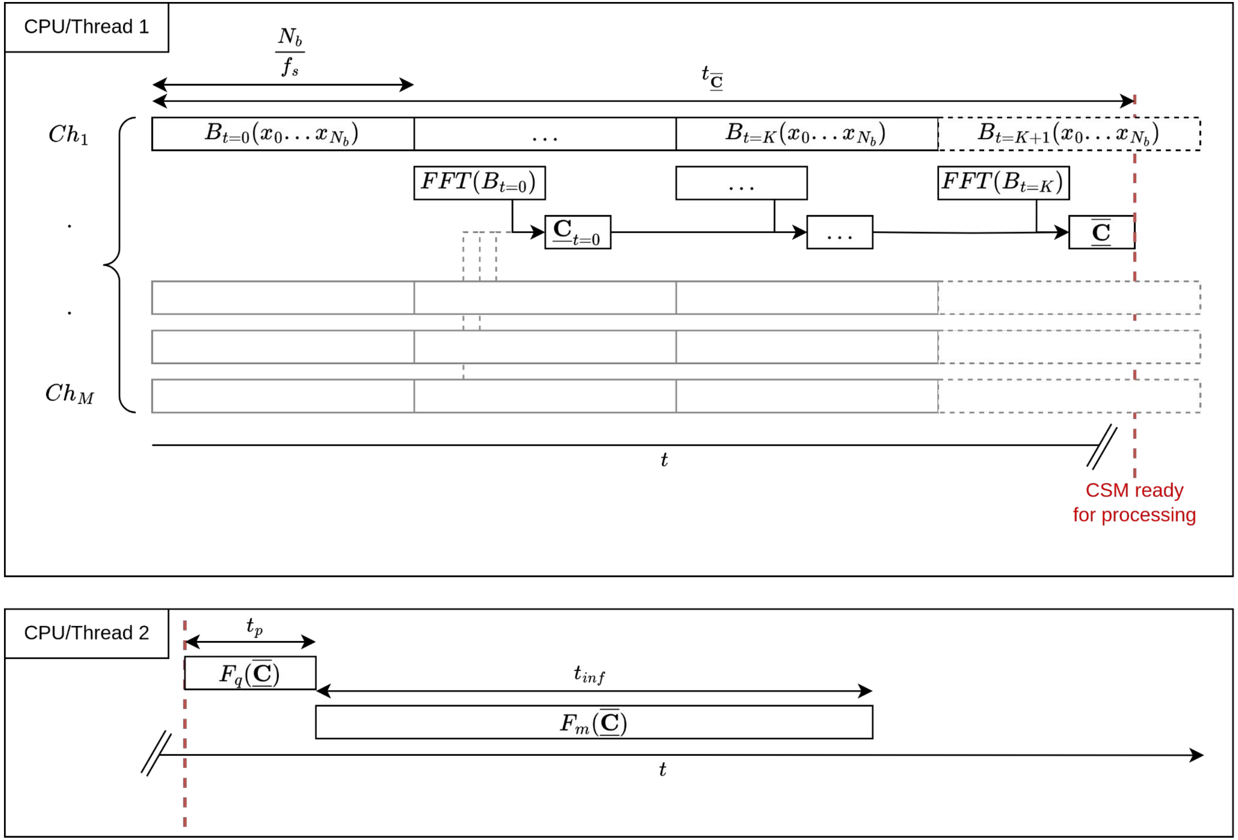
Possible embedded system constraints:

- Supported software
- Limited resources
- Greater processing time



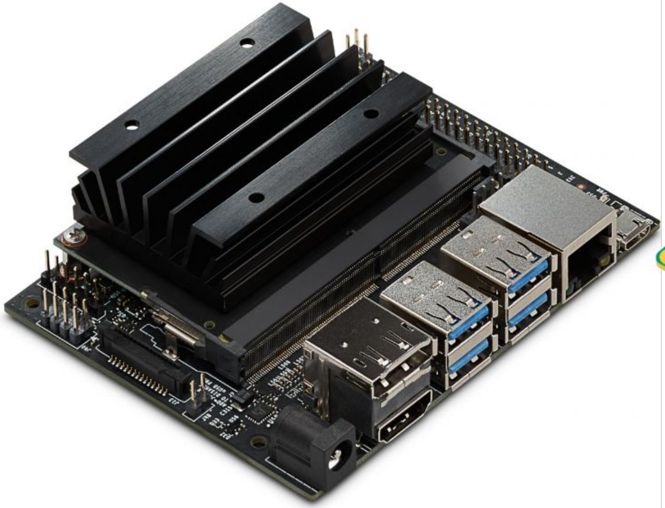
Fundamentals

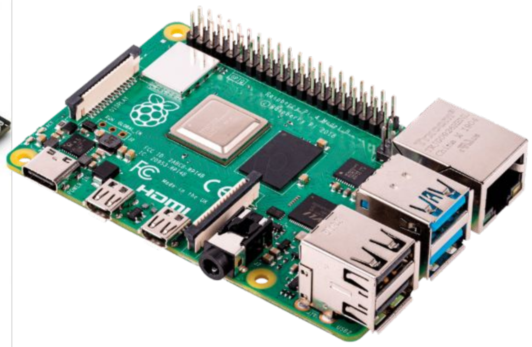
Constraints



Methods

Hardware selection

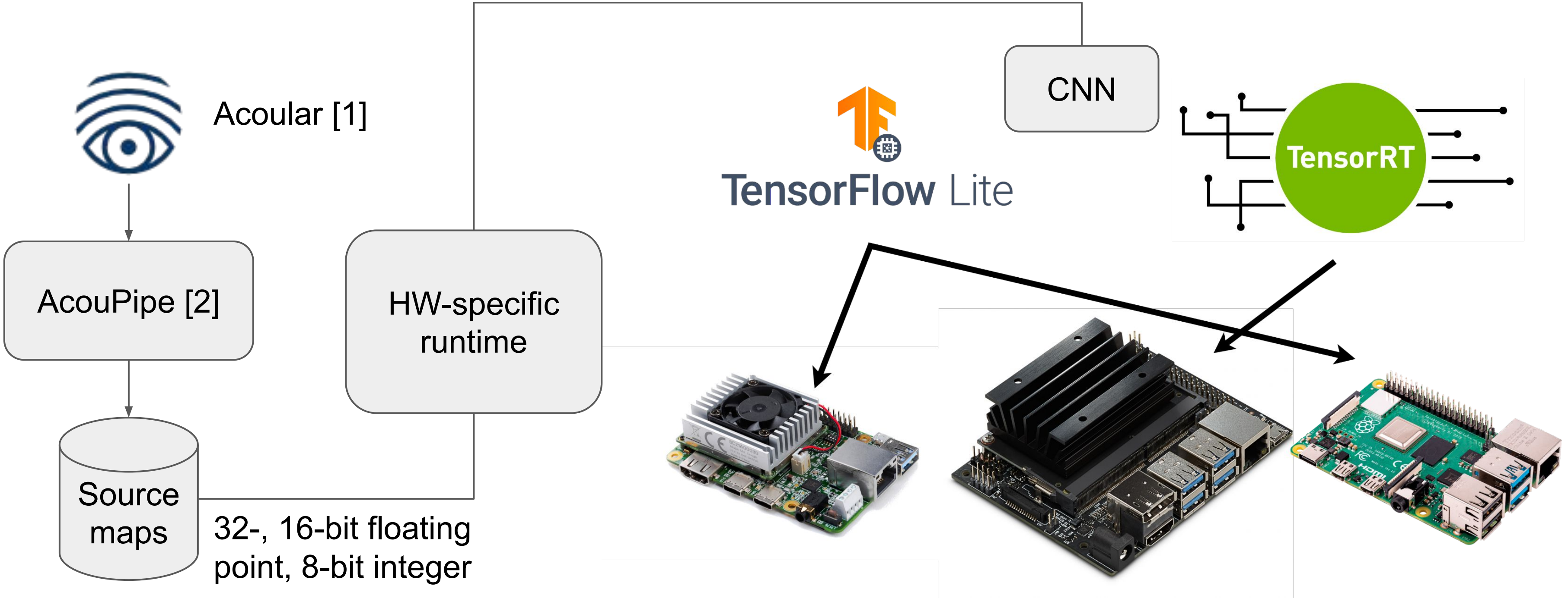




Google Coral-TPU Dev Board (150-180€), NVIDIA Jetson Nano (230-300€), Raspberry Pi 4 (70-130€)

Methods

Software selection



[1] Sarradj, E., & Herold, G. (2017). "A Python framework for microphone array data processing."
[2] Kujawski, A. and Pelling, A. J. R. and Jekosch, S. and Sarradj, E. (2023): "A framework for generating large-scale microphone array data for machine learning."

Results

Static values

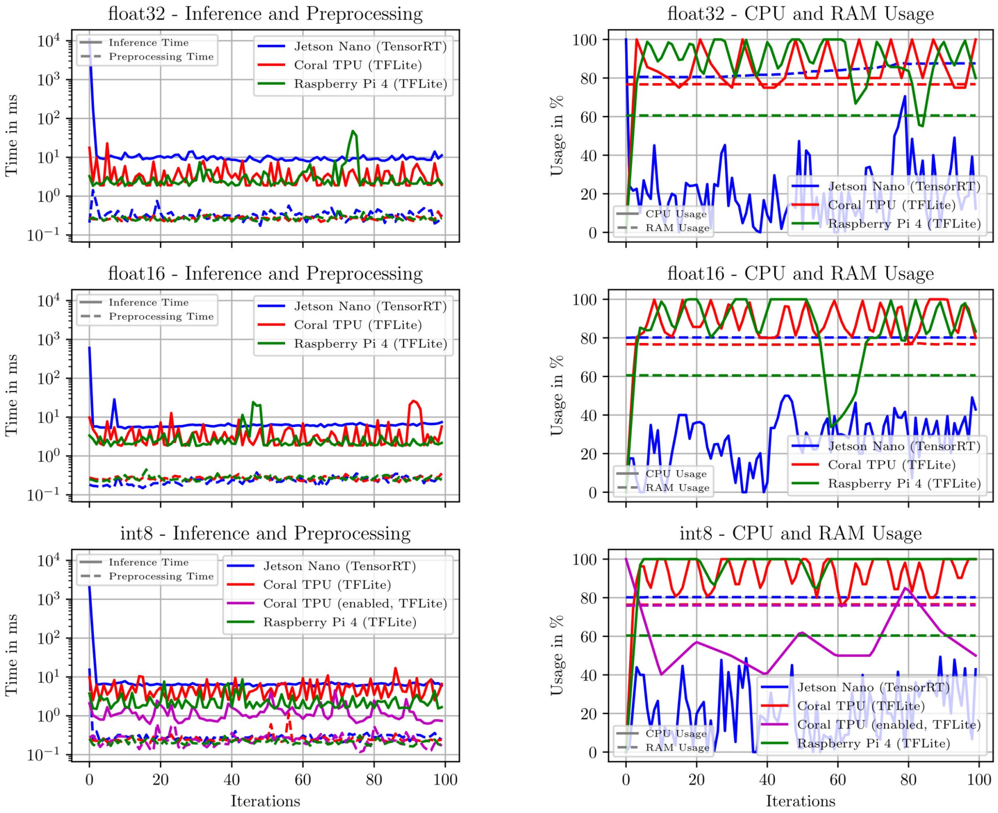
Model Sizes in MB	**32-bit f.p.**	**16-bit f.p.**	**8-bit int.**
TFLite	0.2	0.1	0.05/0.1**
TRT*	1.5	1.3	1.3

Load times in ms	**32-bit f.p.**	**16-bit f.p.**	**8-bit int.**
Coral TPU	5	5.5	4.5/7**
Jetson Nano	43 **in s**	40 **in s**	40 **in s**
Raspberry Pi	1.5	1.9	2.7

*For TRT, the entire folder is measured.
**TPU compilation.

Results

Benchmarks

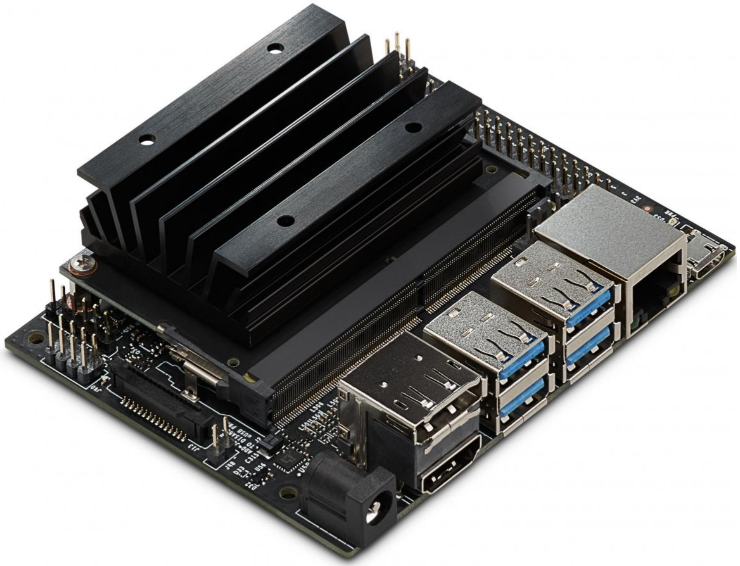


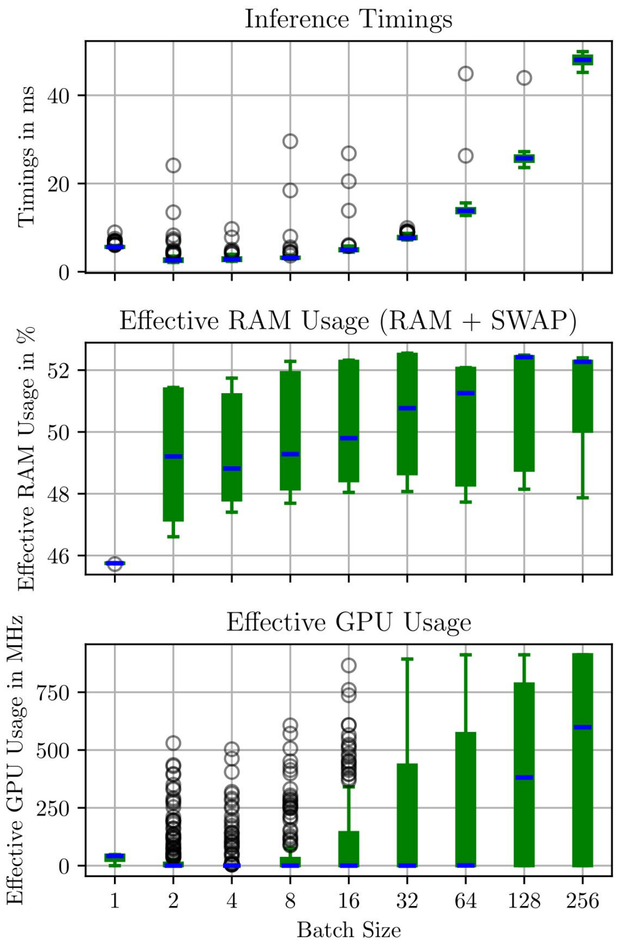
Timing

Overall	**Best per device**	**Results**
Best	Coral-TPU (enabled) 8-bit	μ = 1.23, σ = 0.64 ms
Good	Raspberry Pi 8-bit	μ = 2.42, σ = 1.25 ms
Worst	Jetson Nano 16-bit	μ = 12.3, σ = 59.57ms

Results

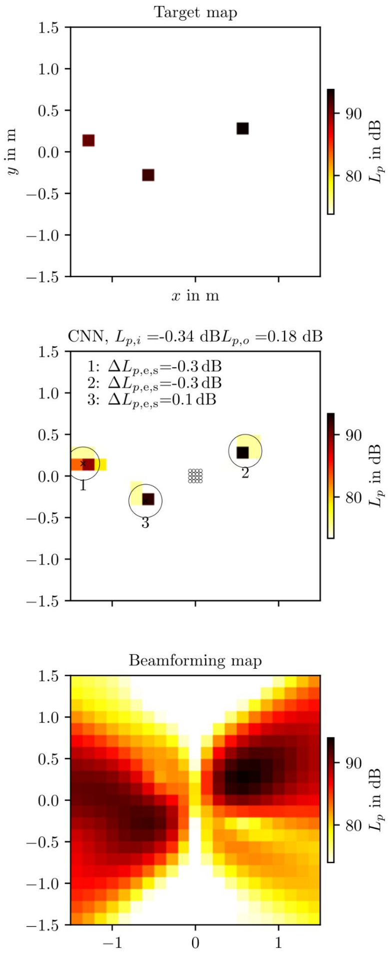
Parallelization experiments

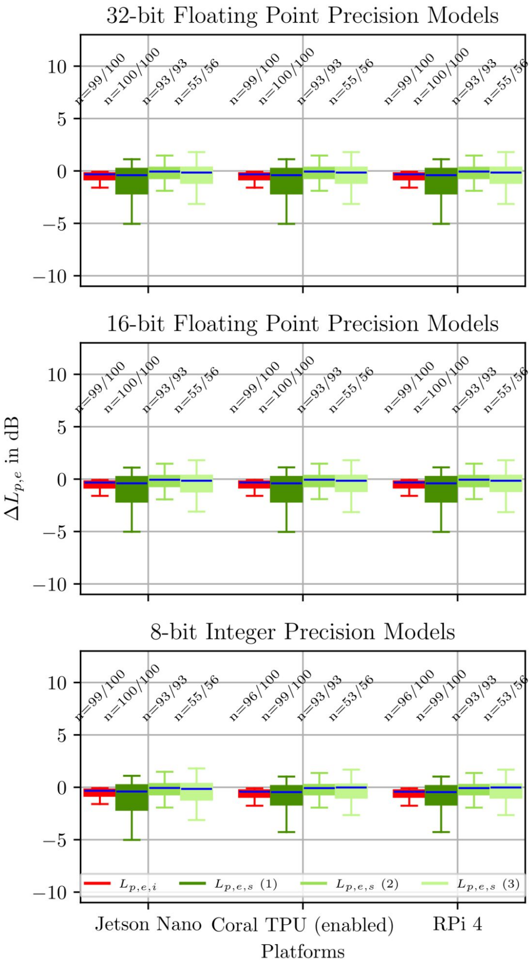




Results

Output quality





No significant difference in output quality.

Discussion

Platforms

	Coral-TPU	Jetson Nano	Raspberry Pi 4
✅	Best performance Speed increase Instruction offloading	16-bit f.p. native precision High batch sizes	Good results without ML acceleration unit
❌	Requires model recompilation EOL (no next gen)	Not suited for this task EOL	High CPU load
❔	Performance with other models	Possibly fastest when pre-processing on GPU	Possible increase with Pi 5

Discussion

Validity & Usability

✅	❌
Established error metric	Comparison with original model difficult because of hardware
Logging performed on separate thread and with established tools (htop, psutil, tegrastats)	Self-referential issues with utilization and timing calls
Conversion logs state successful conversion	Devices have a reduced runtime (without Acoular)
Program can be deployed on most arm64 devices	Used devices are EOL

Discussion

Key takeaways

- ML-based acoustic imaging works on embedded devices
- Embedded GPU/TPU increases performance
- Coral-TPU performs best
- Jetson Nano is not suited for this task
- Embedded systems still need specialized runtimes
- Real-time capabilities are to be explored

References

- P. Castellini et al. “A neural network based microphone array approach to grid-less noise source localization”. In: Applied Acoustics 177 (June 1, 2021), p. 107947.
- W. Ma and X. Liu. “Phased microphone array for sound source localization with deep learning”. In: Aerospace Systems 2.2 (Dec. 1, 2019), pp. 71–81.
- W. Gonçalves Pinto et al. “Deconvoluting acoustic beamforming maps with a deep neural network”. In: INTER-NOISE and NOISE-CON Congress and Conference Proceedings 263.1 (Aug. 1, 2021), pp. 5397–5408.
- S. Pasha et al. “Machine-learnt Beamforming for Large Aperture 3D Microphone Arrays, An Industrial Application”. In: 2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP). 2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP). ISSN: 2473-3628. Oct. 2021, pp. 1–6.
- D. Salvati et al. “On the use of machine learning in microphone array beamforming for far-field sound source localization”. In: 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP). 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP). Sept. 2016, pp. 1–6.
- S. Young Lee et al. “Deep Learning-Enabled High-Resolution and Fast Sound Source Localization in Spherical Microphone Array System”. In: IEEE Transactions on Instrumentation and Measurement 71 (2022). Conference Name: IEEE Transactions on Instrumentation and Measurement, pp. 1–12.
- K. Rashida et al. “Comparison of Machine Learning Classifier Models for Microphone Array based Acoustic Source Localisation”. In: 2023 International Symposium on Ocean Technology (SYMPOL). 2023 International Symposium on Ocean Technology (SYMPOL). ISSN: 2326-5566. Dec. 2023, pp. 1–5.
- E. Sarradj and G. Herold. A Python framework for microphone array data processing. Version 24.05”. Pages: 50–58 Publication Title: Applied Acoustics Volume: 116 original-date: 2015-01-23T11:02:57Z. 2017.
- A. Kujawski et al. “A framework for generating large-scale microphone array data for machine learning”. In: Multimedia Tools and Applications 83.11 (Sept. 25, 2023), pp. 31211–31231.
- A. Kujawski and E. Sarradj. “Fast grid-free strength mapping of multiple sound sources from microphone array data using a Transformer architecture”. In: The Journal of the Acoustical Society of America 152.5 (Nov. 1, 2022), pp. 2543–2556.