“Building TinyML Applications Using Rune”

Kartik Thakore - HOTG

May 25, 2021
tinyML Talks Sponsors

Additional Sponsorships available – contact Olga@tinyML.org for info
Arm: The Software and Hardware Foundation for tinyML

1. Connect to high-level frameworks
2. Supported by end-to-end tooling
3. Connect to Runtime

Profiling and debugging tooling such as Arm Keil MDK

Application

Optimized models for embedded

Runtime (e.g. TensorFlow Lite Micro)

Optimized low-level NN libraries (i.e. CMSIS-NN)

RTOS such as Mbed OS

Arm Cortex-M CPUs and microNPUs

Stay Connected

@ArmSoftwareDevelopers
@ArmSoftwareDev

Resources: developer.arm.com/solutions/machine-learning-on-arm
WE USE AI TO MAKE OTHER AI FASTER, SMALLER AND MORE POWER EFFICIENT

Automatically compress SOTA models like MobileNet to <200KB with little to no drop in accuracy for inference on resource-limited MCUs

Reduce model optimization trial & error from weeks to days using Deeplite's design space exploration

Deploy more models to your device without sacrificing performance or battery life with our easy-to-use software

BECOME BETA USER bit.ly/testdeeplite
TinyML for all developers

Dataset

Acquire valuable training data securely

Enrich data and train ML algorithms

Edge Device

Real sensors in real time
Open source SDK

Embedded and edge compute deployment options

Test

Test impulse with real-time device data flows

Impulse

Copyright © EdgeImpulse Inc.

www.edgeimpulse.com
Maxim Integrated: Enabling Edge Intelligence

**Advanced AI Acceleration IC**

The new MAX78000 implements AI inferences at low energy levels, enabling complex audio and video inferencing to run on small batteries. Now the edge can see and hear like never before.

[www.maximintegrated.com/MAX78000](http://www.maximintegrated.com/MAX78000)

**Low Power Cortex M4 Micros**

Large (3MB flash + 1MB SRAM) and small (256KB flash + 96KB SRAM, 1.6mm x 1.6mm) Cortex M4 microcontrollers enable algorithms and neural networks to run at wearable power levels.

[www.maximintegrated.com/microcontrollers](http://www.maximintegrated.com/microcontrollers)

**Sensors and Signal Conditioning**

Health sensors measure PPG and ECG signals critical to understanding vital signs. Signal chain products enable measuring even the most sensitive signals.

[www.maximintegrated.com/sensors](http://www.maximintegrated.com/sensors)
Qeexo AutoML

Automated Machine Learning Platform that builds tinyML solutions for the Edge using sensor data

Key Features

- Supports 17 ML methods:
  - Multi-class algorithms: GBM, XGBoost, Random Forest, Logistic Regression, Gaussian Naive Bayes, Decision Tree, Polynomial SVM, RBF SVM, SVM, CNN, RNN, CRNN, ANN
  - Single-class algorithms: Local Outlier Factor, One Class SVM, One Class Random Forest, Isolation Forest
- Labels, records, validates, and visualizes time-series sensor data
- On-device inference optimized for low latency, low power consumption, and small memory footprint applications
- Supports Arm® Cortex™- M0 to M4 class MCUs

End-to-End Machine Learning Platform

Target Markets/Applications

- Industrial Predictive Maintenance
- Smart Home
- Wearables
- Automotive
- Mobile
- IoT

For more information, visit: www.qeexo.com
Advancing AI research to make efficient AI ubiquitous

Power efficiency
- Model design, compression, quantization, algorithms, efficient hardware, software tool

Personalization
- Continuous learning, contextual, always-on, privacy-preserved, distributed learning

Efficient learning
- Robust learning through minimal data, unsupervised learning, on-device learning

Perception
- Object detection, speech recognition, contextual fusion

Reasoning
- Scene understanding, language understanding, behavior prediction

Action
- Reinforcement learning for decision making

A platform to scale AI across the industry
Add Advanced Sensing to your Product with Edge AI / TinyML

Pre-built Edge AI sensing modules, plus tools to build your own

**Reality AI solutions**
- Prebuilt sound recognition models for indoor and outdoor use cases
- Solution for industrial anomaly detection
- Pre-built automotive solution that lets cars “see with sound”

**Reality AI Tools® software**
- Build prototypes, then turn them into real products
- Explain ML models and relate the function to the physics
- Optimize the hardware, including sensor selection and placement

https://reality.ai    info@reality.ai    @SensorAI    Reality AI
Build Smart IoT Sensor Devices From Data

SensiML pioneered TinyML software tools that auto generate AI code for the intelligent edge.

- End-to-end AI workflow
- Multi-user auto-labeling of time-series data
- Code transparency and customization at each step in the pipeline

We enable the creation of production-grade smart sensor devices.

sensiml.com
SynSense builds sensing and inference hardware for ultra-low-power (sub-mW) embedded, mobile and edge devices. We design systems for real-time always-on smart sensing, for audio, vision, IMUs, bio-signals and more.

https://SynSense.ai
Focus on:
(i) developing new use cases/apps for tinyML vision; and (ii) promoting tinyML tech & companies in the developer community

Submissions accepted until August 15th, 2021
Winners announced on September 1, 2021 ($6k value)
Sponsorships available: sponsorships@tinyML.org
https://www.hackster.io/contests/tinyml-vision
Successful tinyML Summit 2021:

- **5 days of tutorials, talks, panels, breakouts, symposium**
  - 4 tutorials
  - 6 keynotes & 6 plenary tinyTalks (more in breakouts)
  - 2 panel discussions
  - 5 disruptive news presentations
  - 17 breakout/partner sessions
  - 6 Best Product and Innovation Award Finalists & Presentations
  - 89 Speakers

- **5006** registered attendees representing:
  - 104 countries, 1000+ companies and 400+ academic institutions

- **26** Sponsoring companies

www.youtube.com/tinyML with 150+ videos
June 7-10, 2021 (virtual, but LIVE)
Deadline for abstracts: May 1

Sponsorships are being accepted: sponsorships@tinyML.org
Next tinyML Talks

<table>
<thead>
<tr>
<th>Date</th>
<th>Presenter</th>
<th>Topic / Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday, June 6</td>
<td>Vikram Shrivastava, Knowles Corporate</td>
<td>Dedicated Audio Processors at the Edge are the Future of AI</td>
</tr>
</tbody>
</table>

Webcast start time is 8 am Pacific time

Please contact talks@tinyml.org if you are interested in presenting
Reminders

Slides & Videos will be posted tomorrow

tinyml.org/forums  youtube.com/tinyml

Please use the Q&A window for your questions
Kartik Thakore is an entrepreneur, founder and an experienced AI & Engineering leader in technology with a focus on healthcare. Since 2019, he has served as a core contributor in producing new patentable technologies leveraging medical data using data engineering, data science & clinical modelling. Currently, Kartik is the co-founder of HOTG-ai and Director of AI Engineering at Doc.ai. Prior to his role, he served as Sr. Data Engineer at BetterDoctor (now Quest Analytics), HumanAPI and Head of Data Science at Newtopia. Kartik previously founded AiMED, a startup focused on accelerating medical statistics (data science) in Clinical Research in Canada. Kartik is also involved in initiatives around decentralized clinical trials and serves as the Vice Chair of the IEEE DCT Committee. Kartik received his Masters in Biomedical Engineering (2014) and Bachelor of Engineering (2011) from Western University in Canada.
Building TinyML & EdgeML Applications using Rune

Kartik Thakore
Howdy!

- Director of AI at Doc.ai (recently acquired)
- Open source advocate and founder of HOTG (Tinyverse)
- Enthusiastic about making application development more accessible
Rune Contributors and Maintainers Team
Containerization is the key for TinyML Applications

Leveraging lessons from CloudML
Round 1: TinyML Applications & Deployments

- COVID pandemic and TinyML
- Built a Cough Detector for public spaces
- However, building a robust application required a lot of leg work
- Only knew cloudML approaches
- Needed a consistent code driven approach to building and testing TinyML components for applications
- So I built one!
TinyML: Integrating ML into applications
Let’s make a Rune!

- Declarative definition of TinyML pipelines
- Deploying onto Phones
- Testing live on Phone using Runic Mobile app
- Integrating RuneVM Library into mobile applications
Integration into Realtime Applications (updates)
Use case 1: Full spectrum ML

- How can edge and cloud devices collaborate?
- Large models can recognize thousands of types of objects but the edge device will only see a subset regularly.
- We train a very small model to quickly do a coarse classification on the edge and if needed, we ask the cloud for a more fine grained result.
How can you get started today?

Start building Runes for mobile apps today!

Visit our developer website: https://hotg.dev

Connect with us (resources in last slide)
Roadmap

More devices
Signal, Audio and NLP models
Improved developer experience
Ongoing performance updates
Community support
Connect with us!

Email: community@hotg.ai

Developer website: https://hotg.dev

Twitter: @hotg_ai

Github: https://github.com/hotg-ai

Community: https://tinyverse.discourse.group/
Copyright Notice

This presentation in this publication was presented as a tinyML® Talks webcast. The content reflects the opinion of the author(s) and their respective companies. The inclusion of presentations in this publication does not constitute an endorsement by tinyML Foundation or the sponsors.

There is no copyright protection claimed by this publication. However, each presentation is the work of the authors and their respective companies and may contain copyrighted material. As such, it is strongly encouraged that any use reflect proper acknowledgement to the appropriate source. Any questions regarding the use of any materials presented should be directed to the author(s) or their companies.

tinyML is a registered trademark of the tinyML Foundation.

www.tinyML.org