The new Neuromorphic Analog Signal Processor (NASP) concept and technology platform

Alexander Timofeev - Founder and Chief Executive Officer, Polyn.ai

June 21, 2022
Thank you, tinyML Strategic Partners*, for committing to take tinyML to the next Level, together.
Executive Strategic Partners
Powering tinyML Innovation

Arm AI Virtual Tech Talks

The latest in AI trends, technologies & best practices from Arm and our Ecosystem Partners.

Demos, code examples, workshops, panel sessions and much more!

Fortnightly Tuesday @ 4pm GMT/8am PT

Find out more: www.arm.com/techtalks
EDGE IMPULSE

The Leading Development Platform for Edge ML

dgeimpulse.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
Accelerate Your Edge Compute
SYNTIANT
Making Edge AI A Reality
www.syntiant.com
Platinum Strategic Partners
Fastest Video Analytics Solutions on Arm CPUs
High-Value or Safety-Critical Use Cases?
For your most important projects, use Reality AI Tools®

TinyML software that covers the full engineering lifecycle:
 Reality AI Tools®

- AutoML for non-visual sensing based on advanced signal processing math
- Hardware design analytics
- Explanation of TinyML models in terms of underlying physics
- Automated Data Readiness assessment

https://reality.ai/    @SensorAI    info@reality.ai    Reality AI
Renesas is enabling the next generation of AI-powered solutions that will revolutionize every industry sector.
Sony Semiconductor Solutions Corporation
Gold Strategic Partners
Where what if becomes what is.

Witness potential made possible at analog.com.
Making Over-the-Air Firmware and ML models Updates Simple and Accessible!

- Securely update your IoT devices regardless of their Hardware Platform (Silicon) Provider and physical location.
- Unlock TinyML business value through OTA Firmware and ML models update.
- Pay-as-you-go

www.fotahub.com
contact@fotahub.com
The Latent AI Efficient Inference Platform™ (LEIP) combined with application templates makes optimizing neural networks for production inference a cinch.

THE EASY BUTTON FOR DEVELOPERS resulting in faster time to market.

Reach out to info@latentai.com
TOGETHER, WE ACCELERATE THE BREAKTHROUGHS THAT ADVANCE OUR WORLD

www.nxp.com/ai
Deploy TinyML into the Real World - Plug and Play ML

Sense

- Sensors:
  - modulated and ready-to-use sensors to simplify the setup process
  - support 500+ grove modules

Train

- Codecraft:
  - no code Programming platform to Get Started With TinML
  - supports Arduino, Python, C or JavaScript etc.

- Edge Impulse:
  - to optimize data utilization and enable deploy a machine learning model faster than ever

Inference

- TensorFlow Lite:
  - to easily train low memory usage machine learning models

Applications

- Wio Terminal:
  - completed AI platform --- integrated with a 2.4” LCD Screen, onboard IMU (LIS3DHTR), microphone, buzzer, microSD card slot, light sensor, infrared emitter(IR 940nm)

- Tensorflows
  - Motion /Gesture/Speech /Smell/ Sports
  - Barcode/Face/image

- Artificial Nose
- AI Thermal Camera for Safe Camping
- Azure IoT Squirrel Feeder
The Right Edge AI Tools Can Make or Break Your Next Smart IoT Product

Analytics Toolkit Suite

AutoML

Data Collection

Test & Validation

Model Training

Code Generation

Team Collaboration

Version Control and Model Management

sensiml.com/tinyML
STMicroelectronics provides extensive solutions to make tiny Machine Learning easy
ENGINEERING EXCEPTIONAL EXPERIENCES

We engineer exceptional experiences for consumers in the home, at work, in the car, or on the go.

www.synaptics.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
Silver Strategic Partners
tinyML EMEA Innovation Forum 2022

Connect, Unify, and Grow the tinyML EMEA Community
October 10-12, 2022

https://www.tinyml.org/event/emea-2022

Event will be held in person in Cyprus.

EMEA 2022 Call for Presentations is open till August 1st, 2022.

More sponsorships are available: sponsorships@tinyml.org
Our next tinyML Trailblazers Series
Success Stories with Yoram Zylberberg
(CEO, Emza Visual Sense)

LIVE ONLINE July 6th, 2022 at 8 am PDT

Register now!
Join Growing tinyML Communities:

tinyML - Enabling ultra-low Power ML at the Edge

The tinyML Community
https://www.linkedin.com/groups/13694488/

10k members in 45 Groups in 36 Countries

2.9k members & 7.6k followers
Subscribe to tinyML YouTube Channel for updates and notifications (including this video)
www.youtube.com/tinyML
## Next tinyML Talks

<table>
<thead>
<tr>
<th>Date</th>
<th>Presenter</th>
<th>Topic / Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wednesday, June 29</td>
<td>Thomas Basikolo / Marco Zennaro /</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alessandro Grande</td>
<td>TinyML Challenge 2022: Smart weather station</td>
</tr>
</tbody>
</table>

Webcast start time is 8:00 am Pacific time

Please contact [talks@tinyml.org](mailto:talks@tinyml.org) if you are interested in presenting
Reminders

Slides & Videos will be posted tomorrow

Please use the Q&A window for your questions

tinyml.org/forums  youtube.com/tinyml
Alexander Timofeev

Alexander is a serial entrepreneur in high-tech startups and a venture builder, with more than 20 years in the industry and vast technical and business experience in industries such as nanotechnology, autonomous systems, and super-hard materials. Prior to founding PolyN, Alexander founded iGlass Technology Inc., a company that developed novel electrochromic smart glass technology. He built the core team and product concept and successfully sold the company at the end of 2020 to a strategic investor. Alexander is also the founder (in 2010) and managing partner at FPI Innovation Fund VC, an early-stage venture investment management company. The fund focuses on earlier-stage innovative companies, developing clear product concepts and strategies and working with venture firms and partners for subsequent funding rounds. As a managing partner, Alexander is responsible for greenfield development from early stage to growing business.
A revolutionary solution for an Analog Neuromorphic Product era

Bridging software and hardware worlds

Infinite opportunities
Did you know how amazingly efficient our brain is in sensing the outside world?

Your brain’s fantastic ability to perceive and reason is based on its **ULTRA-COMPRESSED SENSING CAPABILITIES:**

(A) with 100,000x data reduction and

(B) a low operation power

We are focused not for brain activity simulation, like big ML system, our focus is how to achieve highly efficient sensor raw data pre-processing.
TECHNOLOGY CHANGING THE PARADIGM

NASP technology will impact AI/ML markets in a way the DSP influenced signal processing markets 40 years ago.

DSP invented 1980 was revolutionized widely used analog signal pre-processing in audio, telecommunications, digital image processing, radar, sonar, and speech recognition systems, and in common consumer electronic devices such as mobile phones, disk drives, and high-definition television (HDTV) products.

NASP invented 2022 add AI pre-processing on the sensor edge.

Signal processing was revolutionized in the 1980s by the wide adoption of the MOS transistor.

PolyN

NASP: practical analog neuromorphic AI processing
We developed new, much better, way to extract sense of the outside world — for mobile devices. We learned how to embed AI into chips with 100% equivalence of the initial NN* — as Tiny AI**. The ever-reinforcing trends drive demand for always-on raw data processing on sensor level.

THE CONTEXT OF OUR BREAKTHROUGH TECHNOLOGY

Raw data tsunami: e.g. proliferation of IoT devices, sports, digitalization of health, etc.

*NN = Neural Network.
**The Tiny AI™ is always-on smart analytic level.
Today’s Neural Networks capabilities to solve problems in mobile devices – by far exceed the hardware capabilities practically available in mobile devices

Excellent solutions available

Processing of Digital Big Data in CLOUD
CPU, GPU, TPU and digital neuromorphic accelerator

Good solutions available

EDGE
Digital low-power MCU

No adequate solution. But the market requires:

- Ultra low-power
- Low-latency
- Always-on

Tiny AI

RAW-DATA STREAM

Lidar/Radar
Iot Camera
Microphone
IoT Sensor
Bio Sensor
NAEP: THE FIRST EVER PRACTICAL ANALOG NEUROMORPHIC SOLUTION

We successfully developed an affordable, mass-market technology which fully meets the customers’ needs.

Processing of Digital Big Data in CLOUD
CPU, GPU, TPU and digital neuromorphic accelerator

EDGE
Digital low-power MCU

Tiny AI
moves raw data AI processing from CPU to sensor level

NASP is an excellent solution:
• Ultra low-power
• Low-latency
• Always-ON

and even more:
• 4-7x faster time to market
• Cost efficient

Neuromorphic Analog Signal Processing

NAEP
embedded on sensor

RAW-DATA STREAM

Lidar/Radar
IoT Camera
Microphone
IoT Sensor
Bio Sensor
We automatically synthesize a true neuromorphic AI chip layout from a trained neural network

1. Select trained neural net or train it for our customer
2. Convert neural net into NASP Math Model
3. Convert Math Model into chip layout
4. Arrange chip production at a semiconductor foundry

Customer gets a true AI chip for the end-product

BUSINESS MODEL: WE ARE A FABLESS SEMICONDUCTOR COMPANY THAT SUPPLIES NEURAL NETWORKS THAT ARE TRAINED TO SOLVE OUR CUSTOMERS’ PROBLEMS – IN THE FORM OF READY-TO-USE AI CHIPS

We help

Trained Neural Net in a form of a mathematical model in any format, like Keras, TensorFlow or other

Math Model generated with NASP Compiler - Neural Network-to-Chip Layout
D-MVP – on CPU Simulation model

We execute with our fully automated process

We use prepared Layout for the Chip or IP Block (GDSII) for taping out at foundries (TSMC, Global Foundries, Samsung, etc.)

We execute with our fully automated process

NASP chip is built with standard analog neurons of our proprietary design
Foundries use their standard process for NASP chip production
3rd parties make chips to our order

while the production is done by 3rd parties

1-min concept overview
The fundamental NN property: after few thousands of training epochs, the first ~90% of deep neural net layers stop changing regardless of future training cycles. So, we can freeze this part and only update the last ~10% layers for up-train classification if needed.
We can convert the first ~90% of NN layers into fix Analog Neuromorphic Core with high power efficiency. The last ~10% are still in a digital format with full flexibility. The NASP core reduces data more than 1000 times and allows to use ultra-low-power small digital core for post-processing.
We leverage the ecosystem of developers of neural networks.

DIGITAL BLOCK is replaced –

By our proprietary ANALOG NEURON

We convert trained Neural Networks into chip design documentation ready for fabs to produce the chips.

NASP speeds up the Chip Design stage by 4-7x

Results: faster time to market at a lower cost

3rd party semiconductor fabs produce NASP chips with existing equipment and processes.

WHAT IS DIFFERENT FROM THE INDUSTRY ROUTINE?

We use the existing industry technology framework, including equipment.
We generate custom neuromorphic products from an idea to a chip. Not vice versa.

**Semiconductor companies:**
- Do not tailor the product to customer’s true needs
- "Squeeze" customers needs into existing general chips
- Work in a traditional approach from chip to its programming

**Customers:** do not know or understand well the neuromorphic technologies

---

**Traditional way**
- Customers needs for desired performance, functionality, and resources get sacrificed to fit into a pre-build chip

---

**Polyn**
- Start from the problem and neural net training to a Digital MVP
- Synthesized chip for true needs: faster process at mass-market cost efficiency
- NN size and structure always 100% fit to the customers’ task — the fundamental advantage

**Nasp**
- Custom process enables much faster emergence of new analog neuromorphic-based products
NASP target area

Applications

Performance

HPC clouds
Server
Desktops

Engineering
Manufacturing
Robotics

Edge, IOT
Customer electronics

Automotive

Low power

Latency

Source: IMEC
**NASP platform will drive the analog neuromorphic revolution of Tiny AI for numerous applications**

**The Vibration Sensors Market is expected to reach $3.0 billion by 2025 | CAGR of 8.2%, with almost Billion of machines and IoT devices connected**

**Market Pain:**
- Most of the IIoT communications are implemented base on narrow band technologies (LoRa, NB-IoT..) which cannot support full sensor data transmission toward the cloud.
- The industry looking for energy harvesting solution to optimize CAPEX and OPEX, which is impossible within the current approach.

**Smartwatches and Fitness Trackers Shipments Worldwide to reach 259M units in 2025, with 21% CAGR in the smartwatch segment**

**Market Pain:**
- PPG sensing and heart rate calculation is resource hungry.
- Heart rate calculation has low accuracy during person’s activities.
- Human activity learning can be done in cloud-only settings with huge raw data transfer.
- AI is a trend, but it is not used in products, as it needs excessive resources and cloud connection.

**Voice Extraction, KWS, WUW is a huge market for special hearing aids and LMRS, and Laptops, earbuds, smartphones: multi-Billion units market**

**Market Pain:**
- The noise cancelation system does not work properly for voice extraction solutions with irregular noise.
- Most of the state-of-the-art implementations are based on digital processing has high latency and power requirement.
- Additional voice processing features like VAD,KWS, WUW have to work in always-on mode which is hard to implementing in digital way.
NEUROSENSE CHIP OVERVIEW

NASP NeuroSense — tiny AI chip for wearables:

- Replaces sensor MCU
- Drops power consumption
- Boosts accuracy heart rate extraction
- Any human activity recognition and full flexibility on application level

Next Steps

- PPG signal quality monitoring
- Dynamic AFE configuration adopting to usage environment
- Arrhythmia signs detection
- Oxygen saturation level

SEAMLESS INTEGRATION

Key parameters:

- Ultra-low power consumption: below 100µW.
- High accuracy
- Perfectly works at 25Hz
- Small packaged IC footprint: 7x7 to 8x8 mm.
- Cloudless operation

![Image of chip and diagram showing seamless integration and key parameters]
Digital imprint — unique digital matrix generated by NASP neuro core from IMU sensor data flow

Similar to unique QR code

Accuracy up to 94%
UNLIMITED FLEXIBILITY

Activities detected, stored and managed on the fly on application level

Detecting Activity Code

More Options...

Group/re-group activities in classes

Redefine classes

Find similar activities

Number of classes is unlimited

New Class

Rest

Moderate

Intensive

Run

Classes

Learned activities

Cycling

Standing

Walking

Swimming

Commutate

Routine

CrossFit

Water
A trained Deep Neural Network isolates human voice from various noise just like the human brain.

* Including background voices
The NASP Voice Isolation chip, powered with a Neural network, is integrated into the device and enables clear communication for both the speaker and the listener, no matter whether by app or by phone. A direct analog interface – no need of complicated integration, ultra-low power consumption.

**Today**

- Human voice extraction
- Dropped out

**Future**

- ANC/DSP
  - Recognizes background noise and eliminates it by producing anti-noise

**Transmitting side**

- Human (useful) voice
- Irregular noise
- Background noise*

---

* Including unwanted voice, sudden noise, animal noise, baby’s cry, etc.
The NASP VibroSense chip is integrated into the IIoT device for signal pre-processing. It enables ultra narrow-band communication to the Cloud with ultra-low power consumption.

- Support input bandwidth of up to 20KHz
- The data transfer volume drop more than 1000 times to support UNB (ultra narrowband) with ultra-low power consumption (~100uW) to support long battery life

**TODAY:** No compression and pre-processing on data – all traffic to the cloud

**With POLYN:** data compression and/or pre-processing on sensor

**ultra low bandwidth**

**long battery life**
Digital imprint — unique digital array generated by NASP neural core from IMU or piezo sensor data flow.

Digital imprint analysis at the application level makes it possible to identify unusual signal pattern from any source of vibration.

NASP VibroSense level
Vibration Pattern Digital Imprint

500-700 kB/sec

APP level
Alarm if digital imprint differs from normal vibration pattern
Predefined list of normal and abnormal vibrations patterns
Ability to implement anomaly detection on-sensor level without cloud

100 B/sec

Data compression factor more than 1000
NASP platform will drive the analog neuromorphic revolution of Tiny AI for the numerous applications

**VibroSense™**
- Industry 4.0
- Long battery life

**NeuroVoice™**
- Voice Processing
- Voice Authorization
- Key Word Sputtering
- Voice Activation

**NeuroDrive™**
- BLDC motor controllers

**3D Mapping**
- Indoor navigation for robots and drones

**NeuroSense™**
- HR monitoring
- HAR
- Glucose monitoring
- Blood pressure

**Safe and secure face recognition:**
- NO need to send your biometrics over internet

**UNLIMITED OPPORTUNITIES**

NASP inside creates the future with AI all around – sensing continuously and making our lives easier, safer and more pleasant
Copyright Notice

This multimedia file is copyright © 2022 by tinyML Foundation. All rights reserved. It may not be duplicated or distributed in any form without prior written approval.

tinyML® is a registered trademark of the tinyML Foundation.

www.tinylml.org
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