Seeed Studio

XIAO ESP32S3 Sense

Prof. Marcelo José Rovai
UNIFEI - Federal University of Itajubá, Brazil
TinyML4D Academic Network Co-Chair
Powerful MCU Board: ESP32S3 32-bit, dual-core, Xtensa processor chip operating up to 240 MHz.

Elaborate Power Design: Lithium battery charge management capability (deep sleep mode with power consumption as low as 14μA)

Great Memory for more Possibilities: Offer 8MB PSRAM and 8MB FLASH

Outstanding RF performance: Support 2.4GHz Wi-Fi and BLE dual wireless communication, support 100m+ remote communication when connected with U.FL antenna

Thumb-sized Compact Design: 21 x 17.5mm, adopting the classic form factor of XIAO, suitable for space-limited projects like wearable devices

Advanced Functionality: Detachable OV2640 camera sensor for 1600*1200 resolution, compatible with OV5640 camera sensor, integrating an additional digital microphone and an SD card slot for external 32GB FAT memory.
# Hardware (Dev. Boards)

<table>
<thead>
<tr>
<th></th>
<th>Raspberry Pico (W)</th>
<th>Arduino Nano Sense</th>
<th>Espressif ESP 32</th>
<th>Seeed XIAO ESP32S3 Sense</th>
<th>Arduino Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>32Bits CPU</td>
<td>Dual-core Arm Cortex-M0+</td>
<td>Arm Cortex-M4F</td>
<td>Xtensa LX6 Dual Core</td>
<td>Xtensa LX7 Dual Core</td>
<td>Dual Core Arm Cortex M7/M4</td>
</tr>
<tr>
<td>CLOCK</td>
<td>133MHz</td>
<td>64MHz</td>
<td>240MHz</td>
<td>240MHz</td>
<td>480/240MHz</td>
</tr>
<tr>
<td>RAM</td>
<td>264KB</td>
<td>256KB</td>
<td>520KB (part available)</td>
<td>8MB (PSRAM)</td>
<td>1MB</td>
</tr>
<tr>
<td>ROM</td>
<td>2MB</td>
<td>1MB</td>
<td>2MB</td>
<td>8MB</td>
<td>2MB</td>
</tr>
<tr>
<td>Radio</td>
<td>(Yes for W)</td>
<td>BLE</td>
<td>BLE / WiFi</td>
<td>BLE / WiFi</td>
<td>BLE / WiFi</td>
</tr>
<tr>
<td>Sensors</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes (Nicla)</td>
</tr>
<tr>
<td>Bat. Power Manag.</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Price</td>
<td>$</td>
<td>$$$</td>
<td>$</td>
<td>$</td>
<td>$$$$</td>
</tr>
</tbody>
</table>

TinyML Application
Examples
Machine Learning

- Supervised
  - Task driven
  - Regression
  - Classification
  - Object Detection

- Unsupervised
  - Data driven
  - Anomaly Detection

- Reinforcement
  - Learns to react to an environment
  - Autonomous Navigation
Create a new project

Enter the name for your new project:
XIAO-ESP32S3-CAM-Fruits-vs-Veggies-v1-ESP-NN

Choose your project type:
- Developer
  20 min job limit, 4GB or 4 hours of data, limited collaboration.
  
- Enterprise
  No job or data size limits, higher performance, custom blocks.
  Create under organization: Edge Impulse Experts

Create new project
Fruits and Vegetables Image Recognition Dataset (Kaggle)
An impulse takes raw data, uses signal processing to extract features, and then uses a learning block to classify new data.

Image data

- Input axes: image
- Image width: 96
- Image height: 96
- Resize mode: Fit shortest

For optimal accuracy with transfer learning blocks, use a 96x96 or 160x160 image size.

Image

- Name: image
- Input axes (1): image

Transfer Learning (Images)

- Name: Transfer learning
- Input features:
  - Image
- Output features:
  - 3 (apple, banana, potato)

Output features

- Name: 3 (apple, banana, potato)
- Save impulse

Add a processing block

Add a learning block
Initialization

Main Loop

Image provider

Data:
- ESP-NN (Input: 96 x 96 x 3) – RGB

Detection responder

Model

Read Image

Crop & Convert

(96 x 96 x 3) – RGB

Serial Monitor
<table>
<thead>
<tr>
<th>Device</th>
<th>Processor</th>
<th>Frequency</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>XIAO ESP32S3</td>
<td>Xtensa LX7</td>
<td>240 MHz</td>
<td>135 ms</td>
</tr>
<tr>
<td>ESP - CAM</td>
<td>Xtensa LX6</td>
<td>240 MHz</td>
<td>171 ms</td>
</tr>
<tr>
<td>ARDUINO Pro</td>
<td>ARM H7</td>
<td>480 MHz</td>
<td>45 ms</td>
</tr>
</tbody>
</table>
Sound

Vibration

Vision

Exploring Machine Learning with the new XIAO ESP32S3
MJRoBot (Marcelo Rovai)
MPU6050 SCL → XIAO D5
MPU6050 SDA → XIAO D4
MPU6050 VCC → XIAO 3.3V
MPU6050 GND → XIAO GND
Terrestrial Lift
Maritime
Idle
Terrestrial
2 second window

300 Raw Features

Raw Data from sensor

Manual Feature Extraction

Spectral Analysis

Features
- RMS
- SKEW
- KURT
- FFT
- PSD

NN Classifier

Classes
- Lift
- Terrestrial
- Maritime
- Idle

TinyML under the hood: Spectral Analysis
An impulse takes raw data, uses signal processing to extract features, and then uses a learning block to classify new data.
### Feature importance

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>accZ RMS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accZ Spectral Power 0.78 - 2.34 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accX RMS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accY Spectral Power 10.16 - 11.72 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accY Spectral Power 17.97 - 19.53 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accX Spectral Power 7.03 - 8.59 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accY RMS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accX Spectral Power 0.78 - 2.34 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accZ Spectral Power 2.34 - 3.91 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accY Spectral Power 3.91 - 5.47 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accY Spectral Power 11.72 - 13.28 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accY Spectral Power 8.59 - 10.16 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accY Spectral Power 0.78 - 2.34 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accY Spectral Power 14.84 - 16.41 Hz</td>
<td></td>
<td></td>
</tr>
<tr>
<td>accY Spectral Power 16.41 - 17.97 Hz</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Feature explorer

- **idle**
- **lift**
- **maritime**
- **terrestrial**
Preprocess

Data

units = 63

63 Features

• RMS
• SKEW
• KURT
• FFT
• PSD

Classes

• Lift
• Terrestrial
• Maritime
• Idle

Input Layer

Hidden Layer 1

Hidden Layer 2

Output Layer

• Dense
  kernel (63x20)
  bias (20)
  activation = relu
  units = 20
  ReLU

• Dense
  kernel (20x10)
  bias (10)
  activation = relu
  units = 10
  ReLU

• Dense
  kernel (10x4)
  bias (4)
  activation = softmax
  units = 4
  Softmax

input → y_pred
Configure your deployment

You can deploy your impulse to any device. This makes the model run without an internet connection, minimizes latency, and runs with minimal power consumption. Read more.

**SELECTED DEPLOYMENT**

Arduino library

An Arduino library with examples that runs on most Arm-based Arduino development boards.

**MODEL OPTIMIZATIONS**

Model optimizations can increase on-device performance but may reduce accuracy.

<table>
<thead>
<tr>
<th>Option</th>
<th>Quantized (int8)</th>
<th>Unoptimized (float32)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable EON Compiler</td>
<td>Selected</td>
<td>Select</td>
</tr>
<tr>
<td>Latency</td>
<td>2 ms</td>
<td>7 ms</td>
</tr>
<tr>
<td>RAM</td>
<td>1.7K</td>
<td>1.7K</td>
</tr>
<tr>
<td>Flash</td>
<td>10.8K</td>
<td>11.8K</td>
</tr>
<tr>
<td>Accuracy</td>
<td>98.86%</td>
<td>97.54%</td>
</tr>
</tbody>
</table>

Build results for Cortex-M4 80MHz: Change target

![Building Arduino Library](image)

Built Arduino Library

Full build successful in the Arduino IDE.

[More information](link)
TinyML Made Easy: KeyWord Spotting (KWS)
MJRoBot (Marcelo Rovai)
Personal Assistant

KWS
KeyWord Spotting (KWS) - Inference

Obtains an input

“Yes”

Digital Mic

16KHz / 16 bits Sample: [1s]

Pre-Process

MFCC Feature Converter

Output: Image [49, 40, 1]

Runs model

Prob ‘Noise’
Prob ‘Unknown’
Prob ‘Yes’
Prob ‘No’

Output Dim [1, 4]

Post-Processes

If Probability of YES is greater than 80%
Take actions

Make things happen

Inference
MJRoBot (Marcelo Rovai) / XIAO-ESP32S3-KWS

This is your Edge Impulse project. From here you acquire new training data, design impulses and train models.

About this project

This public Edge Impulse project does not have a README yet. Clone this project to add new data or retrain this project, or to deploy this project to a device.

Download block output

<table>
<thead>
<tr>
<th>TITLE</th>
<th>TYPE</th>
<th>SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFCC training data</td>
<td>NPY file</td>
<td>4830 windows</td>
</tr>
<tr>
<td>MFCC training labels</td>
<td>NPY file</td>
<td>4830 windows</td>
</tr>
<tr>
<td>MFCC testing data</td>
<td>NPY file</td>
<td>1435 windows</td>
</tr>
</tbody>
</table>

Run this model

Scan QR code or launch in browser

Launch in browser

Summary

DATA COLLECTED
1h 42m 36s

XIAO-ESP32S3-KWS (Edge Impulse)
Speech Commands Dataset (reduced set)

Dataset

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>Label</th>
<th>Added</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>unknown.1f5l59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 13:30</td>
<td>1 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
<tr>
<td>unknown.1f6e59_...</td>
<td>unknown</td>
<td>May 22, 2023, 14:35</td>
<td>12 sec</td>
</tr>
</tbody>
</table>

Metadata

No metadata.
Pre-Processing (MFCC)

1 second sample@16KHz raw data -> 16,000 features

Processed features -> 637 features (13 x 49)
An impulse takes raw data, uses signal processing to extract features, and then uses a learning block to classify new data.
Neural network architecture

Architecture presets
- 1D Convolutional (Default)
- 2D Convolutional

Input layer (637 features)

Reshape layer (13 columns)

1D conv / pool layer (8 neurons, 3 kernel size, 1 layer)

Dropout (rate 0.25)

1D conv / pool layer (16 neurons, 3 kernel size, 1 layer)

Dropout (rate 0.25)

Flatten layer

Add an extra layer

Output layer (4 classes)
unknown: 0.625000
yes: 0.039063
Predictions (DSP: 514 ms., Classifier)
no: 0.000000
noise: 0.000000
unknown: 0.019518
yes: 0.980469
To learn more…

- IESTI01 TinyML - Machine Learning for Embedding Devices (Videos: Pt)
- WALC 22 – Applied AI - TinyML (Videos in Spanish)
- Professional Certificate in Tiny Machine Learning (TinyML) – edX/Harvard
- Introduction to Embedded Machine Learning - Coursera/Edge Impulse
- Computer Vision with Embedded Machine Learning - Coursera/Edge Impulse
- "Deep Learning with Python" book by François Chollet
- "TinyML" book by Pete Warden, Daniel Situnayake
- "TinyML Cookbook" by Gian Marco Iodice
- "AI at the Edge" book by Daniel Situnayake, Jenny Plunkett

On the TinyML4D website, you can find lots of educational materials on TinyML. They are all free and open-source for educational uses – we ask that if you use the material, please cite them! TinyML4D is an initiative to make TinyML education available to everyone globally.
**TinyML4D Show&Tell Presentations**

<table>
<thead>
<tr>
<th>Date</th>
<th>Thread</th>
<th>Video</th>
</tr>
</thead>
<tbody>
<tr>
<td>August 31st, 2023</td>
<td>TBD</td>
<td>Video here when ready</td>
</tr>
<tr>
<td>May 25th, 2023</td>
<td>Thread here</td>
<td>Video here when ready</td>
</tr>
<tr>
<td>April 20th, 2023</td>
<td>Thread here</td>
<td><a href="https://youtu.be/uoM_ljXjDFY">https://youtu.be/uoM_ljXjDFY</a></td>
</tr>
<tr>
<td>March 30th, 2023</td>
<td>thread here</td>
<td><a href="https://youtu.be/UQO1-SwBwUY">https://youtu.be/UQO1-SwBwUY</a></td>
</tr>
<tr>
<td>February 23rd, 2023</td>
<td>thread here</td>
<td><a href="https://youtu.be/BAEdil7X68Y">https://youtu.be/BAEdil7X68Y</a></td>
</tr>
<tr>
<td>January 26th, 2023</td>
<td>thread here 17</td>
<td><a href="https://youtu.be/-0xRZ-5UYUc_9">https://youtu.be/-0xRZ-5UYUc_9</a></td>
</tr>
<tr>
<td>December 1st, 2022</td>
<td>thread here 2</td>
<td><a href="https://youtu.be/e49pkjnIMIQ_8">https://youtu.be/e49pkjnIMIQ_8</a></td>
</tr>
<tr>
<td>October 27th, 2022</td>
<td>thread here 2</td>
<td><a href="https://youtu.be/s8_hKpOWUwY_1">https://youtu.be/s8_hKpOWUwY_1</a></td>
</tr>
</tbody>
</table>

**TinymML4D Academic Network Show and Tell Main Index.**

The TinyML4D Academic Network Students should use this form to sign up for the latest presentations.
https://forms.gle/ic52HZMqVv4pBrkP7_2

The Show and Tell are typically held at 2 pm UTC on the last Thursday of each month and will take place in this Zoom room.
https://zoom.us/j/95229860797_1
Meeting ID: 952 2986 0797
Passcode: 141278
Thanks