TensorFlow Lite for microcontrollers

Deploy machine learning models on tiny devices

Daniel Situnayake
@dansitu
TensorFlow Lite is our production ready, cross-platform framework for deploying ML on mobile devices and embedded systems
TensorFlow Lite is our production ready, cross-platform framework for deploying ML on mobile devices and embedded systems.

Open Source
TensorFlow Lite for microcontrollers

TensorFlow provides you with a single framework to deploy on Microcontrollers as well as phones.
Training workflow

- TensorFlow (keras or estimator)
- SavedModel
- TF Lite converter
- TF Lite model
Inference workflow

1. Load your model
2. Preprocess data
3. TFLite interpreter
4. Use the resulting output
Where does it run?
Supported architectures

- 32-bit
- Tested extensively with Arm Cortex-M series
- Designed to be easy to port
- Device-specific hardware optimizations
Memory requirements

- 16 KB runtime
- 22 KB including enough ops for speech
- Speech recognition yes/no model: 20 KB
- Vision-based person detection: 250 KB
Development environment

- C++ 11, no standard libraries
- TensorFlow Lite flatbuffer
- Generate projects for Make, Mbed, Keil
- Arduino library
Op support

- 43 ops supported (up from 3 in April)
- No RNNs yet
- Rapidly adding more
- **Contribute**: Help us prioritize!
New things since last time:
Hello World

- Get up and running
- Train a model to replicate a sine function
- Integrate it into an application
- Can use as a project template
Micro speech

- Speech hotword detection
- 20 KB model recognizing yes/no
- Application captures audio
- Reusable preprocessing code for audio classification
- Training scripts and dataset
Person detection

- Image classification (96x96 grayscale input)
- 250 KB model recognizing person/not person
- Application captures images from image sensor
- Training scripts and dataset (COCO)
Magic wand

- Time series data classification
  (3-axis accelerometer)
- 20 KB model recognizing 3 gestures
- Application captures accelerometer data
- Training scripts coming soon!
Arduino library

- **Launch of official Arduino library** - run example code directly from desktop and web IDEs onto Arduino hardware

- **Speech detection in 5 minutes** - open source models available to get started quickly on Arduino
Get started with microcontrollers

This document will help you get started using TensorFlow Lite for Microcontrollers. It explains how to run the framework’s example applications, then walks through the code for a simple application that runs inference on a microcontroller.

Get a supported device

To follow this guide, you’ll need a supported hardware device. The example application we’ll be using has been tested on the following devices:

tensorflow.org/lite/microcontrollers
TinyML book

- First reference book on ML for MCUs - coming soon from O’Reilly, a comprehensive resource with examples and code, all using TensorFlow

tinyurl.com/tinyml-book
How to get started

Documentation
tensorflow.org/lite/microcontrollers
How to get involved

Micro special interest group
tensorflow.org/community/forums

Pull requests welcome
github.com/tensorflow/tensorflow
One more thing...