# Embedded Deep Learning

MER

#### Fabio Frazao

MERIDIAN, Institute for Big Data Analytics, Dalhousie University, Halifax, Canada



- Deep learning on the edge
- Platforms
- Constraints

## Deep learning on the edge



• Conventional pipeline



#### Deep learning on the edge



• Edge computing



Data processing happens near the source, with the processing algorithms embedded into the hardware

#### Deep learning on the edge

• What is deep learning?

## Deep learning is an approach to **machine learning** that uses **deep neural networks**



Everyday applications:

Speech recognition & synthesis



Translation



#### Face Recognition



Object detection





# 

#### Luo et al.: JASA Express Letters Published Online 4 January 2019 https://doi.org/10.1121/1.5085647 Convolutional neural network for detecting odontocete echolocation clicks Wenyu Luo, Wuyi Yang,<sup>a)</sup> and Yu Zhang Key Laboratory of Underwater Acoustic Communication and Marine Information Technology of the Ministry of Education. College of Ocean and Earth Science. niversity, Xiamen, China wyyang@xmu.edu.cn, yuzhang@xmu.edu.cn Marine Mammal Species Classification using Convolutional Neural Networks and a Novel Acoustic Representation **SCIENTIFIC** Mark Thomas<sup>1</sup>, Bruce Martin<sup>2</sup>, Katie Kowarski<sup>2</sup>, Briand Gaudet<sup>2</sup>, and Stan REPORTS Matwin<sup>1,3\*</sup>

Deep learning on the edge

#### ORCA-SPOT: An Automatic Killer Whale Sound Detection Toolkit Using Deep Learning

Christian Bergler<sup>1</sup>, Hendrik Schröter<sup>1</sup>, Rachael Xi Cheng<sup>2</sup>, Volker Barth<sup>3</sup>, Michael Weber<sup>3</sup>, Elmar Nöth<sup>3</sup>, Heribert Hofer<sup>0,2,4,5</sup> & Andreas Maier<sup>0</sup>

#### Performance of a Deep Neural Network at Detecting North Atlantic Right Whale Upcalls

Oliver S. Kirsebom,<sup>1,-10</sup> Fabio Frazao,<sup>1</sup> Van Simard,<sup>2,3</sup> Nathalie Rog,<sup>3</sup> Stan Matoin,<sup>1,4</sup> and Samuel Giard,<sup>3</sup> <sup>1</sup> Anistatic for Big Data Analytics, Dahonise Uliverstrij, Rajdiar, Nova Sotio BH JR. Condo <sup>2</sup> Fibbrics and Ocoans Canada Chair in undervater acoustics applied to cosystem and marine mammads, Marine Sciences, Folistite, Fibbrieria Jd Queber data Minosaki, Rimosaki, Qedee, Canada <sup>3</sup> Marries Lamontagne Institute, Fibbrieria Jd Qoeans Canada, Mori-Joli, Qedee, Canada <sup>4</sup> Marries Lamontagne Institute, Fibbrieria and Ocoans Canada, Mori-Joli, Qedee, Canada <sup>4</sup> Marries Analysis Anadami, Jd Soinerse, Warnaw, Poland



of Sperm Whale Bioacoustics Peter C. Bermant<sup>1</sup>, Michael M. Bronstein<sup>1,2,7</sup>, Robert J. Wood<sup>1,4</sup>, Shane Gero<sup>6</sup> &

David F. Gruber 01,6

Staple deep neural networks are useful for underwater acoustics.

- + Better performance in many tasks
- + High adaptability and reusability
  - Requires more resources (data, computing power, etc)
  - Often hard to interpret

Can they be embedded into PAM systems?



Deep learning requires more resources during the training phase than during the deployment phase



Training





#### Running



#### Small computers







Raspberry Pi

Banana Pi

Rock Pi



Deep learning-specific processors

Google Coral edge TPU

			System on Module	olderable module
Carrier board	PCIe modules	USB accelerator		





Google Coral edge TPU



Carrier board



Carrier boards (development, evaluation boards)

Device



Carrier board



More general

User's application



More specific (Power budget, environmental conditions, additional hardware)



Nvidia Jetson

Deep learning-specific processors







Jetson Nano

Jetson Tx2

Jetson Xavier



Jetson Nano Dev board



Microcontrollers

- A number of basic components you would find in a computer (processors, memory, oscillators, etc) in one chip
- Lower power consumption and computing power



#### Microcontrollers

DFRobot Firebeetle



With Espressif's esp32



With Nordic's nRF52840



With ST's STM32F746

Arduino Portenta H7



With ST's STM32H747



#### Software



#### Software





#### Software







#### Software





(https://www.st.com/en/embedded-software/x-cube-ai.html)





#### Conclusions



- The kind of model you can run in the field will depend on what devices your system can afford
- Small computers and DL SoMs are capable of running relatively complex models
- Microcontrollers will run simpler models, which might be enough for some applications



Thank You!