MATLAB EXPO

Data-Centric AI for Signal Processing Applications

Gabriele Bunkheila, MathWorks

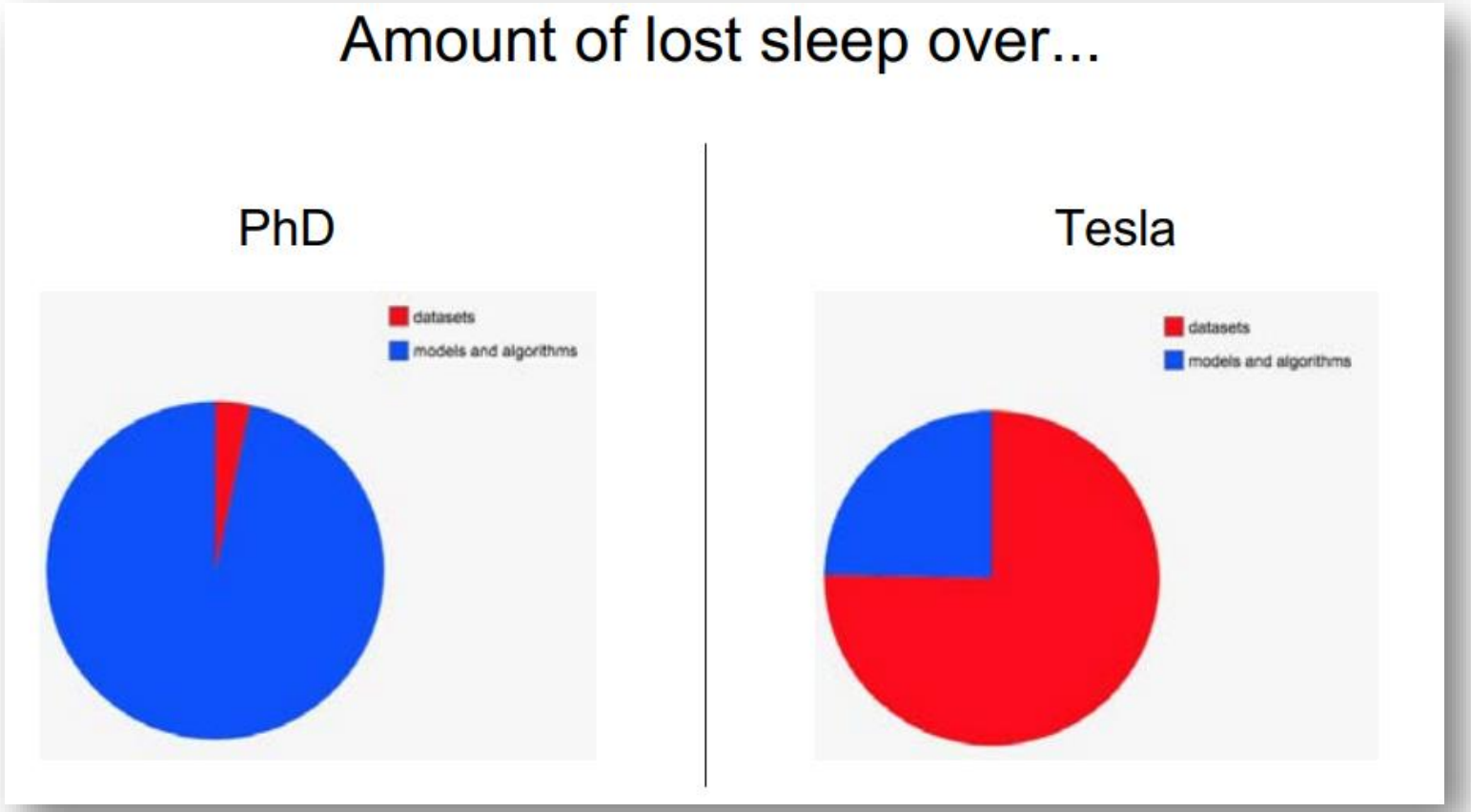
Frantz Bouchereau, MathWorks

MathWorks

linkedin.com/in/gabrielebunkheila/

linkedin.com/in/frantz-bouchereau-1673375/

Industry and Research Invest in AI in Different Ways
Better Models or Better Data?

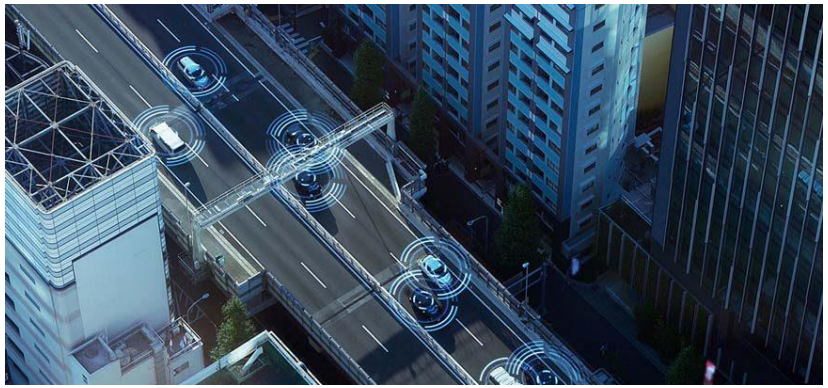


Andrej Karpathy – Building the Software 2.0 Stack (Spark+AI Summit 2018)

Data-Centric AI in 2022 – Trend Gaining Pace and Visibility

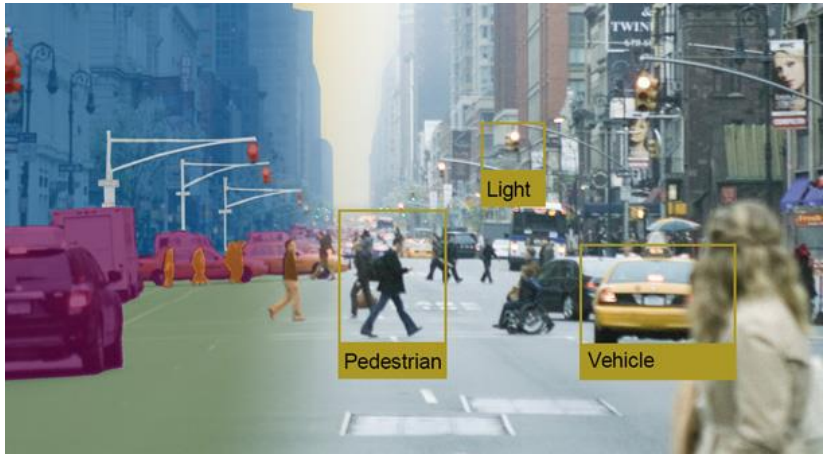
https://spectrum.ieee.org/andrew-ng-data-centric-ai

Most existing AI resources support few applications



Autonomous driving

Language modeling

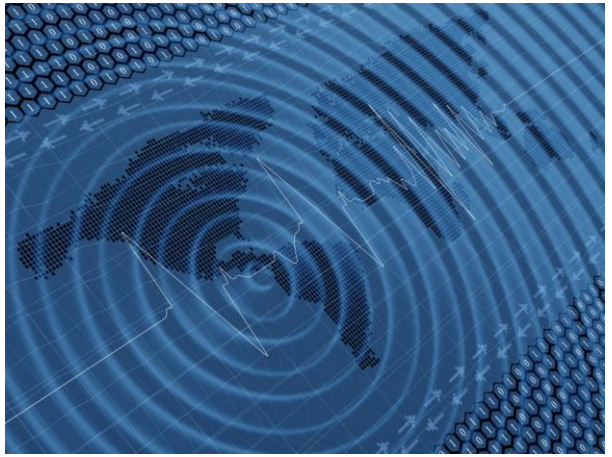


Computer vision

Speech recognition

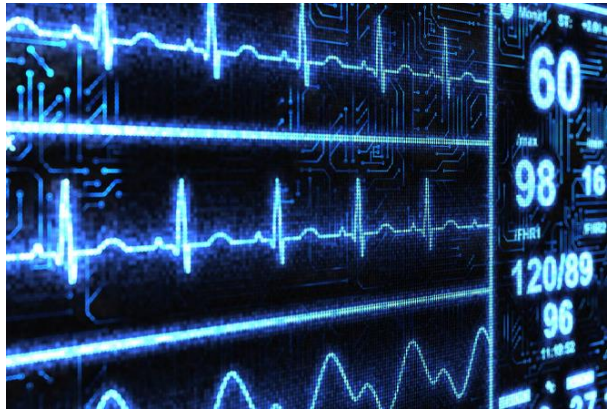
Most signal processing applications cannot count on many AI resources

Vibration analysis



Seismic analysis

Predictive maintenance



Digital health

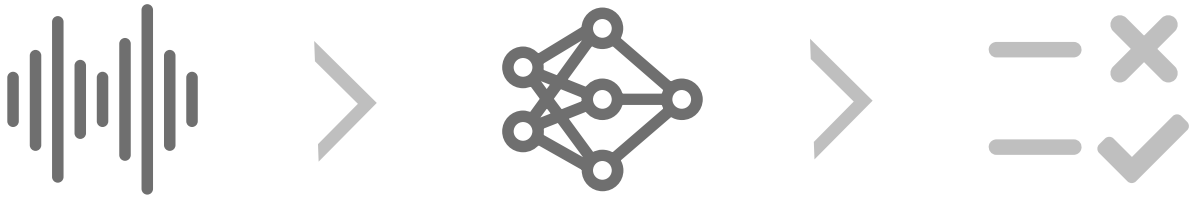
Machine health

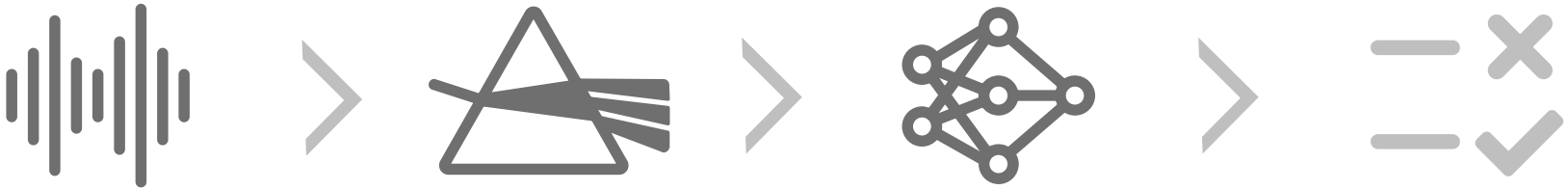
...

Time for a survey…

Which of these best describes your AI-related challenges?

- ○ Model Complexity
- ○ Data Complexity
- ○ AI Expertise

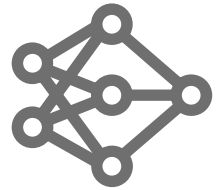




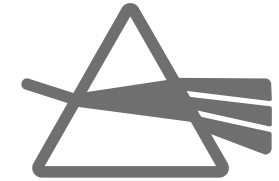
Data-Centric AI in Signal Processing Applications

Agenda – Three Practical Engineering Approaches

1. Transfer learning with pre-trained AI models



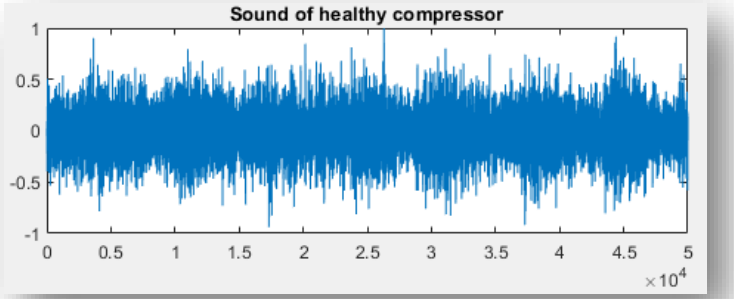
2. Feature extraction with simpler and smaller AI models

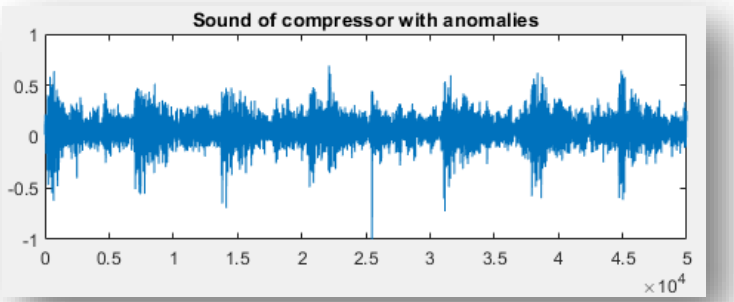


3. Better signal datasets, real or simulated

How can I apply transfer learning to detecting faults in an air compressors based on their noise

- Have dataset with labeled sound recordings
- One “healthy” class
- 7 different classes of faults
- 1800.wav files, 225 per class

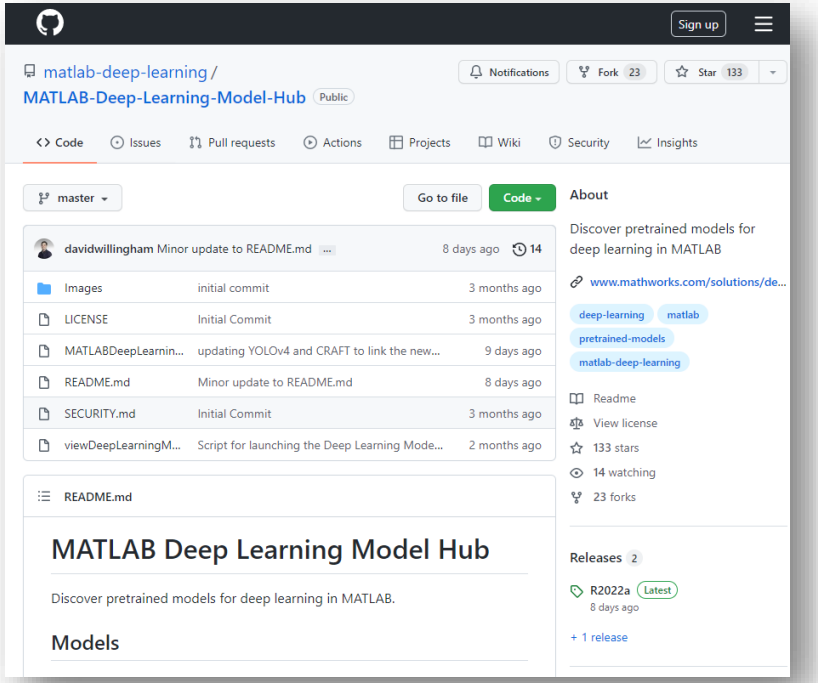




[Example: Transfer Learning with Pretrained Audio Networks in Deep Network Designer]

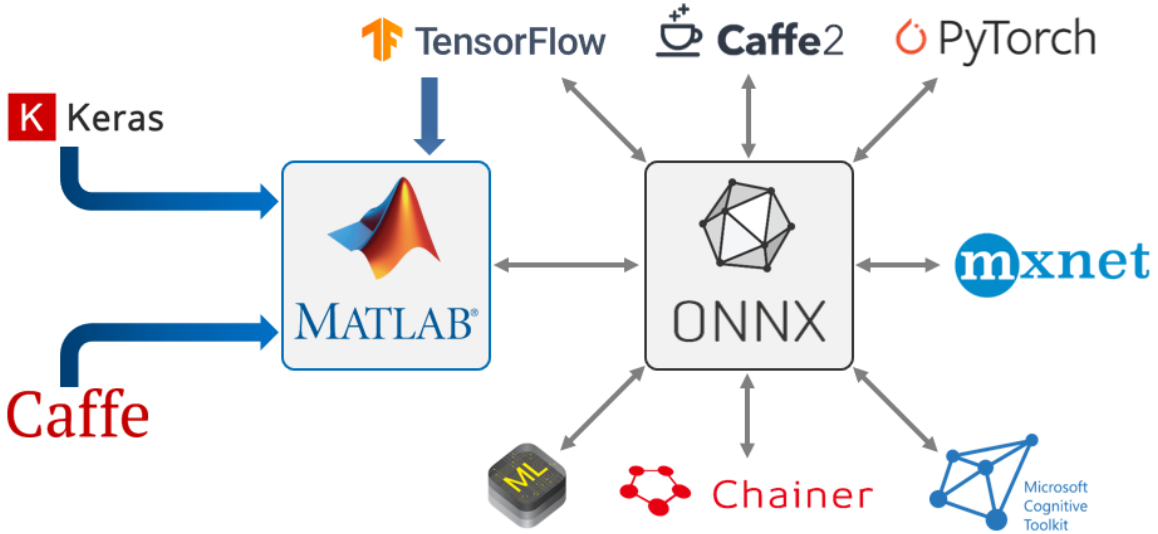
Finding a pre-trained deep learning network for Transfer Learning

- Find one directly in MATLAB

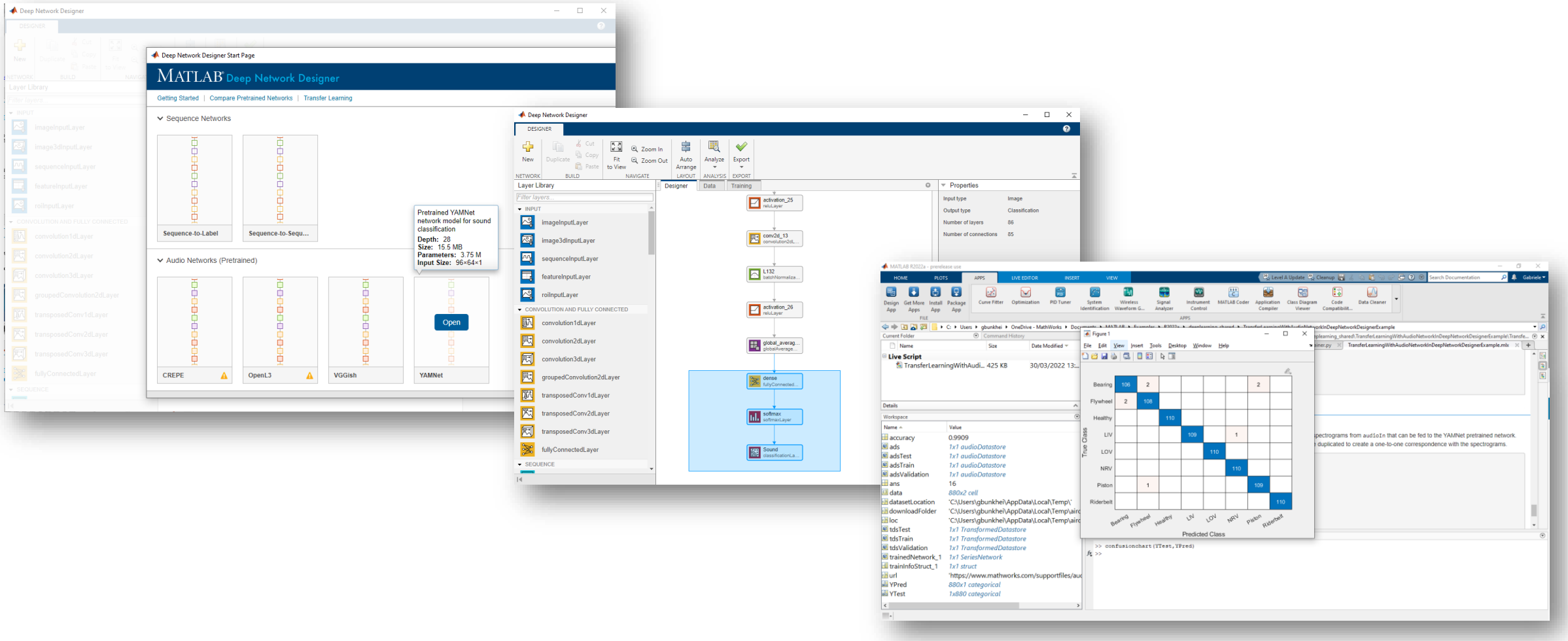


https://github.com/matlab-deep-learning/MATLAB-Deep-Learning-Model-Hub

- Import it from a known non-MATLAB repository



Demo – Transfer Learning with YAMNet for Fault Detection



[Example: Transfer Learning with Pretrained Audio Networks in Deep Network Designer]

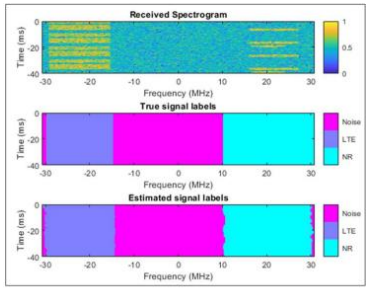
Transfer Learning – Handouts

Choosing the right model for transfer learning

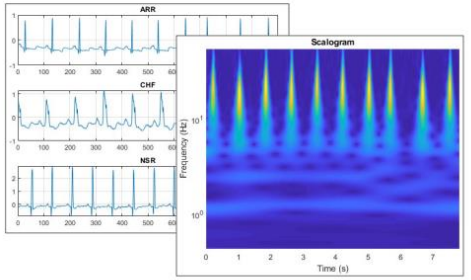
https://www.mdpi.com/2224-2708/10/4/72

Transfer Learning with models pre-trained on different types of data

- Spectrum Sensing with Deep Learning to Identify 5G and LTE Signals
- Network: ResNet-50 (Image segmentation)
- Input: 256-by-256-by-3 images
- Features: `spectrogram` of baseband waveforms



- Classify Time Series Using Wavelet Analysis and Deep Learning
- Network: GoogLeNet (Image object classification)
- Input: 224-by-224-by-3 images
- Features: `cwt` (scalogram) of ECG signals

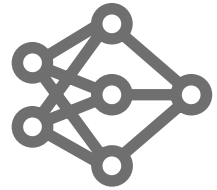


Download @ Journal of Sensors and Actuator Networks

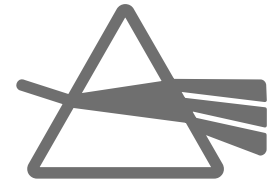
Data-Centric AI in Signal Processing Applications

Agenda – Three Practical Engineering Approaches

1. Transfer learning with pre-trained AI models

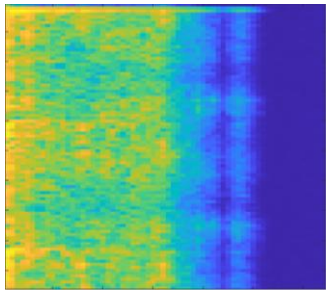


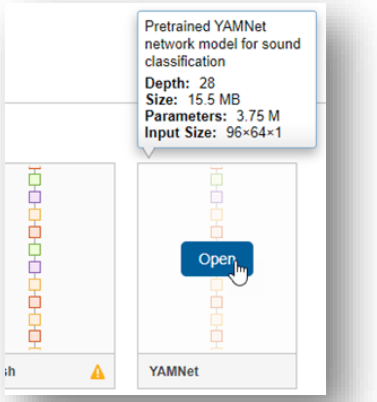
2. Feature extraction with simpler and smaller AI models

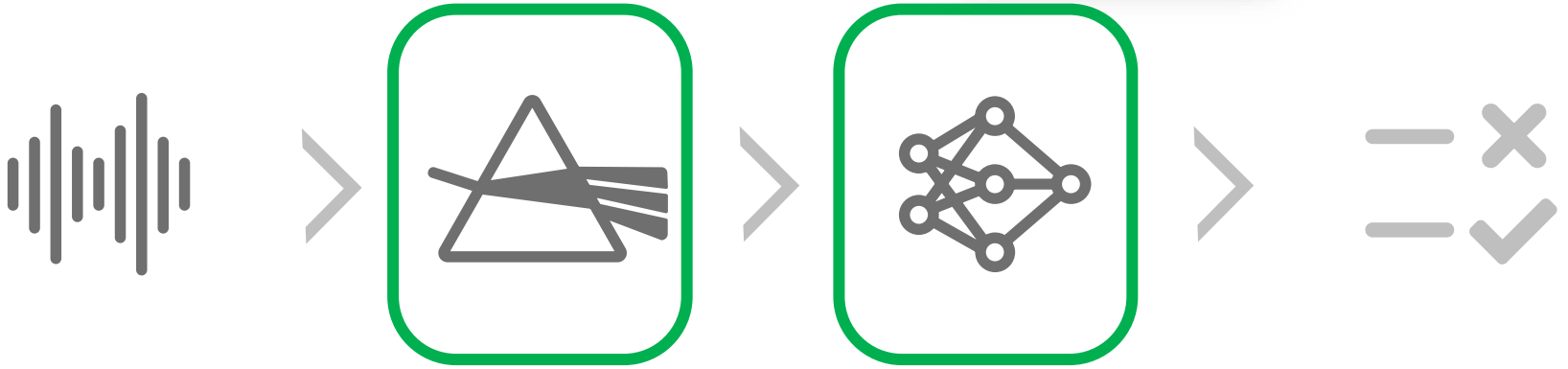


3. Better signal datasets, real or simulated

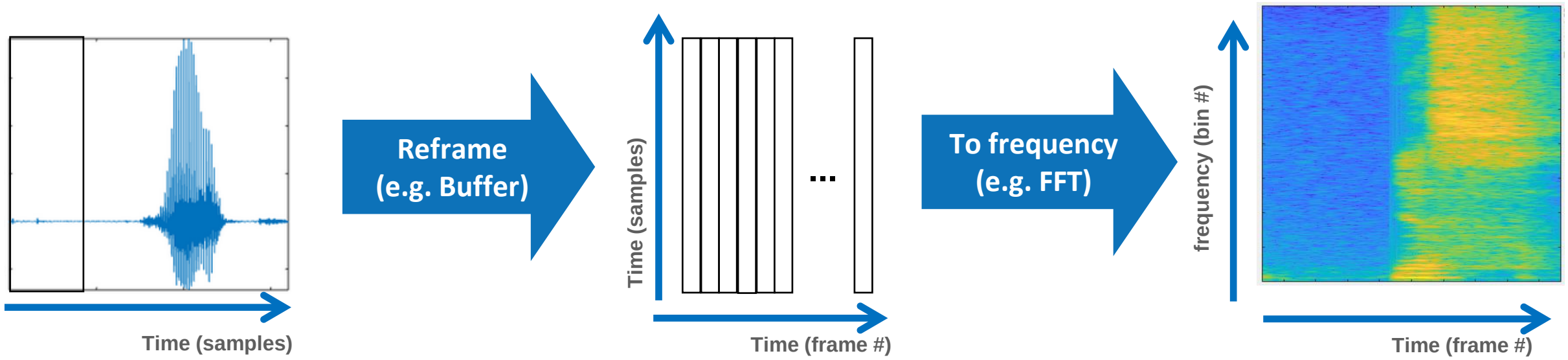
Deep networks most often don’t learn directly from raw signals





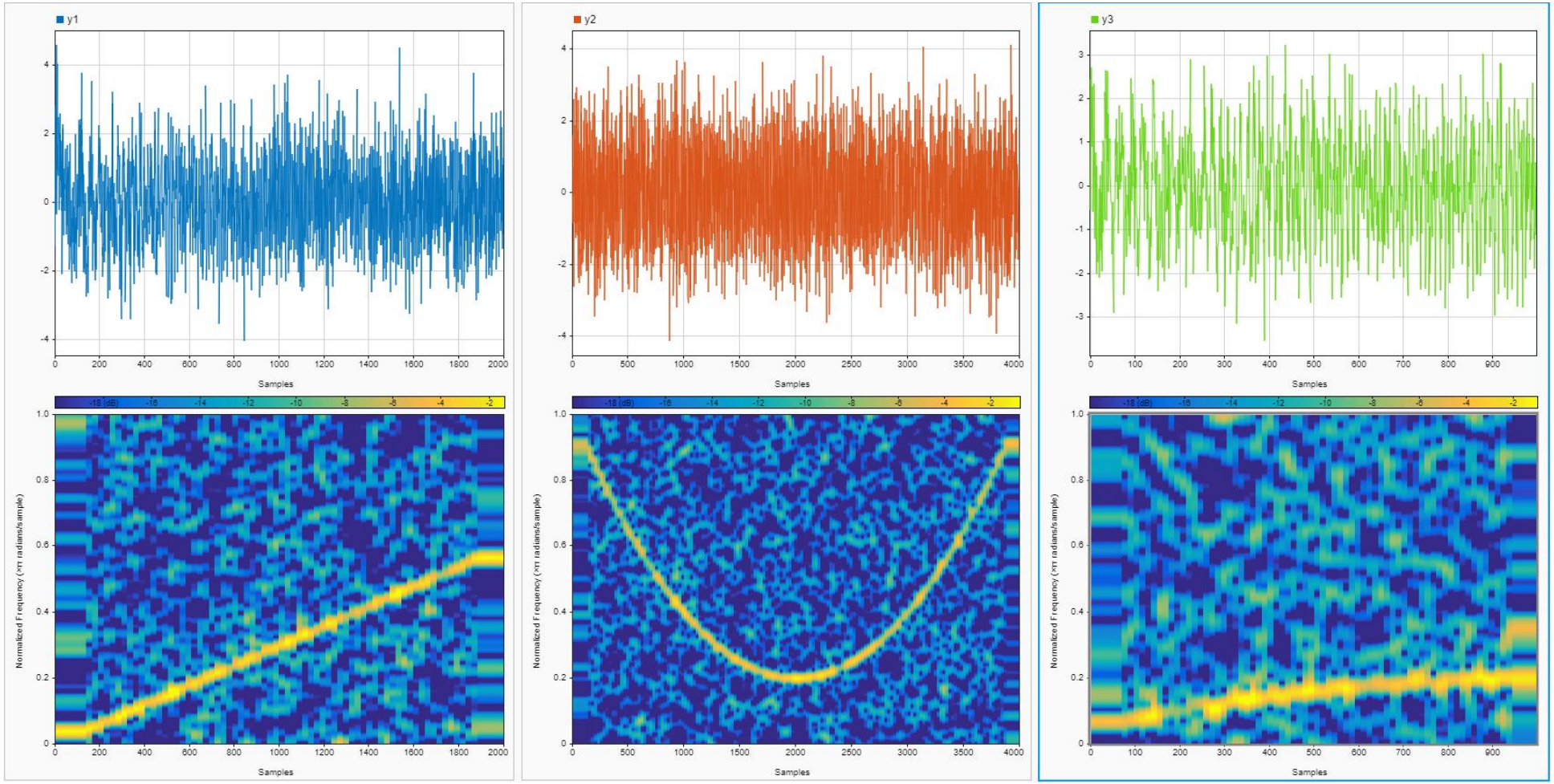


Time-frequency transformations are popular feature extraction methods



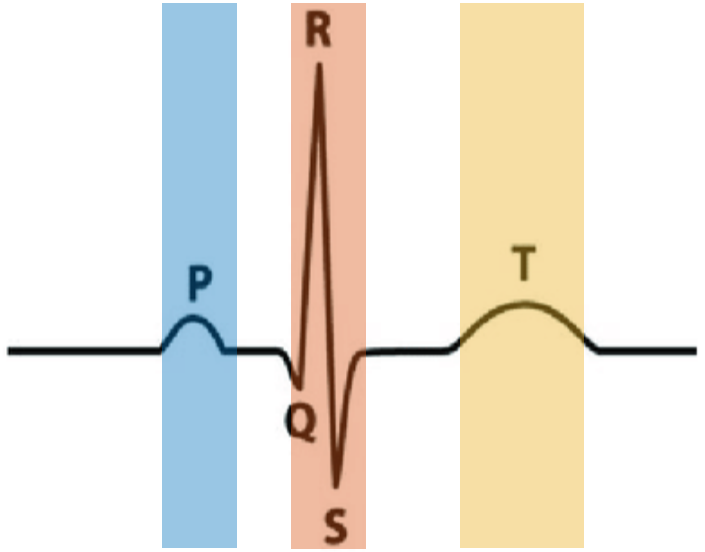
Time-frequency transforms make signal characteristics more evident

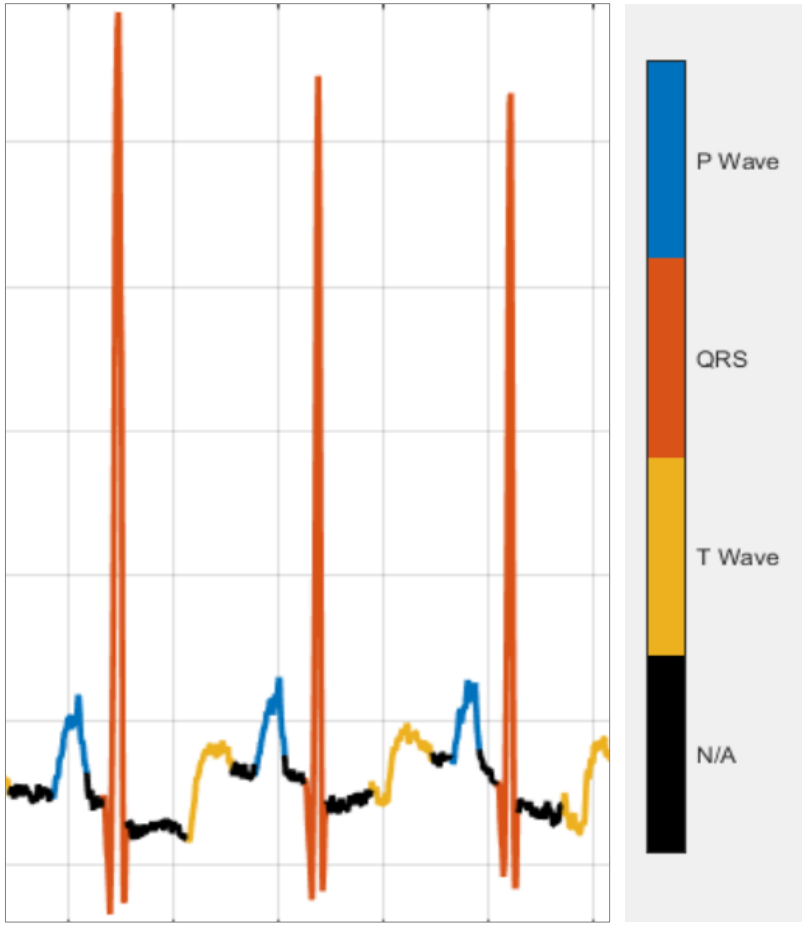
Time-frequency transform



How to use feature extraction to segment ECG signals?

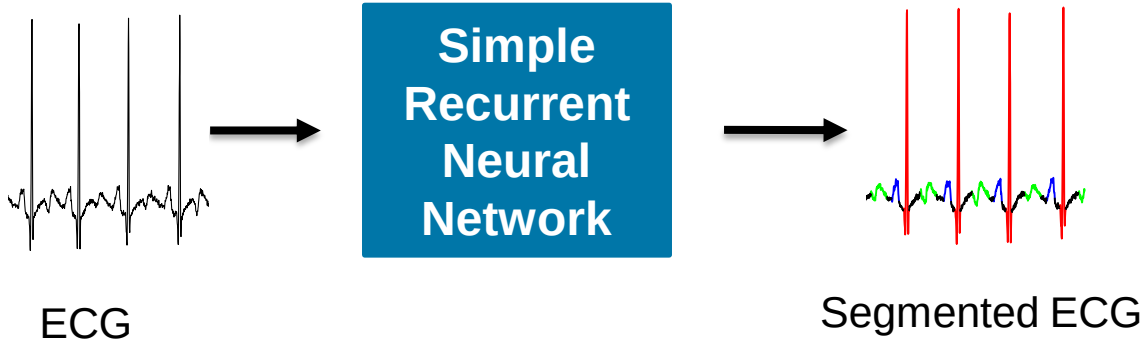
- Have dataset with signals labeled by cardiologists
- 3 types of wave events
- 210 ECG recordings (total ~15 minutes)

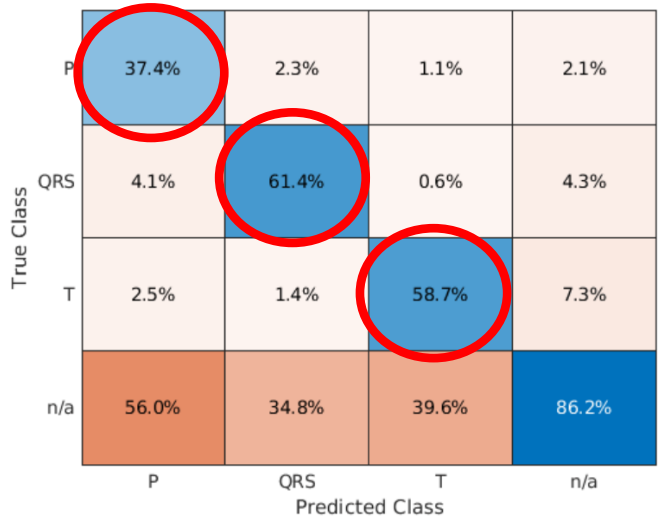


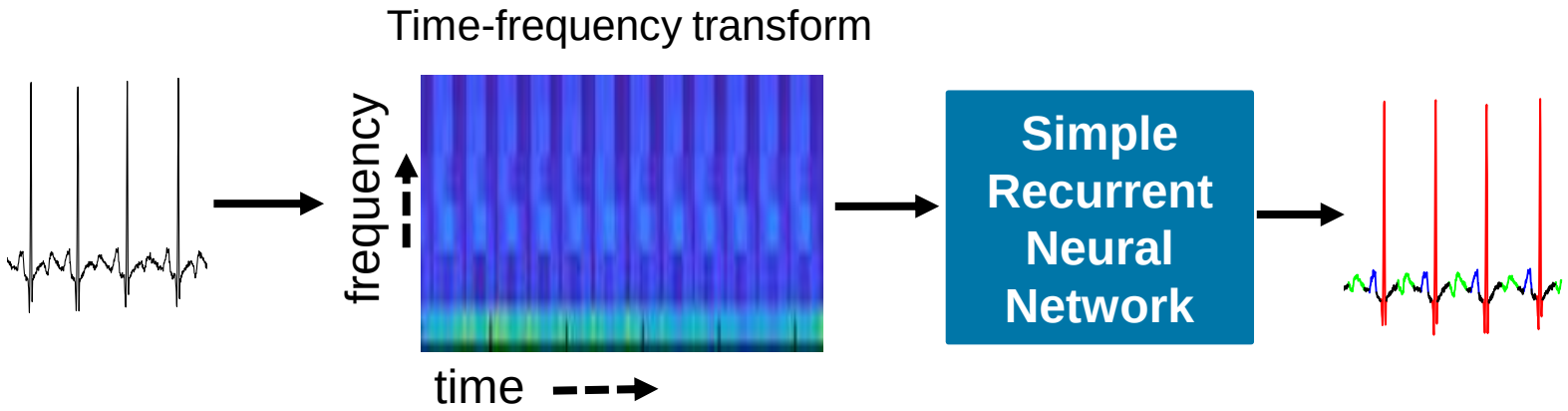


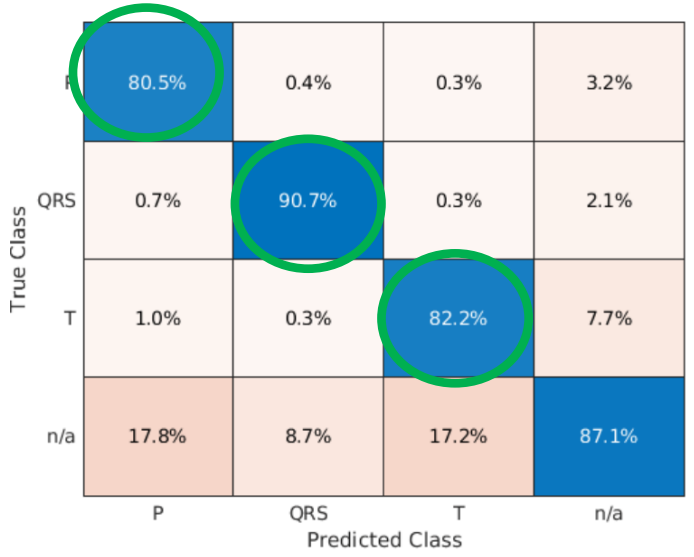
[Example: Waveform Segmentation Using Deep Learning]

Feature extraction allows getting high accuracy from simple AI models

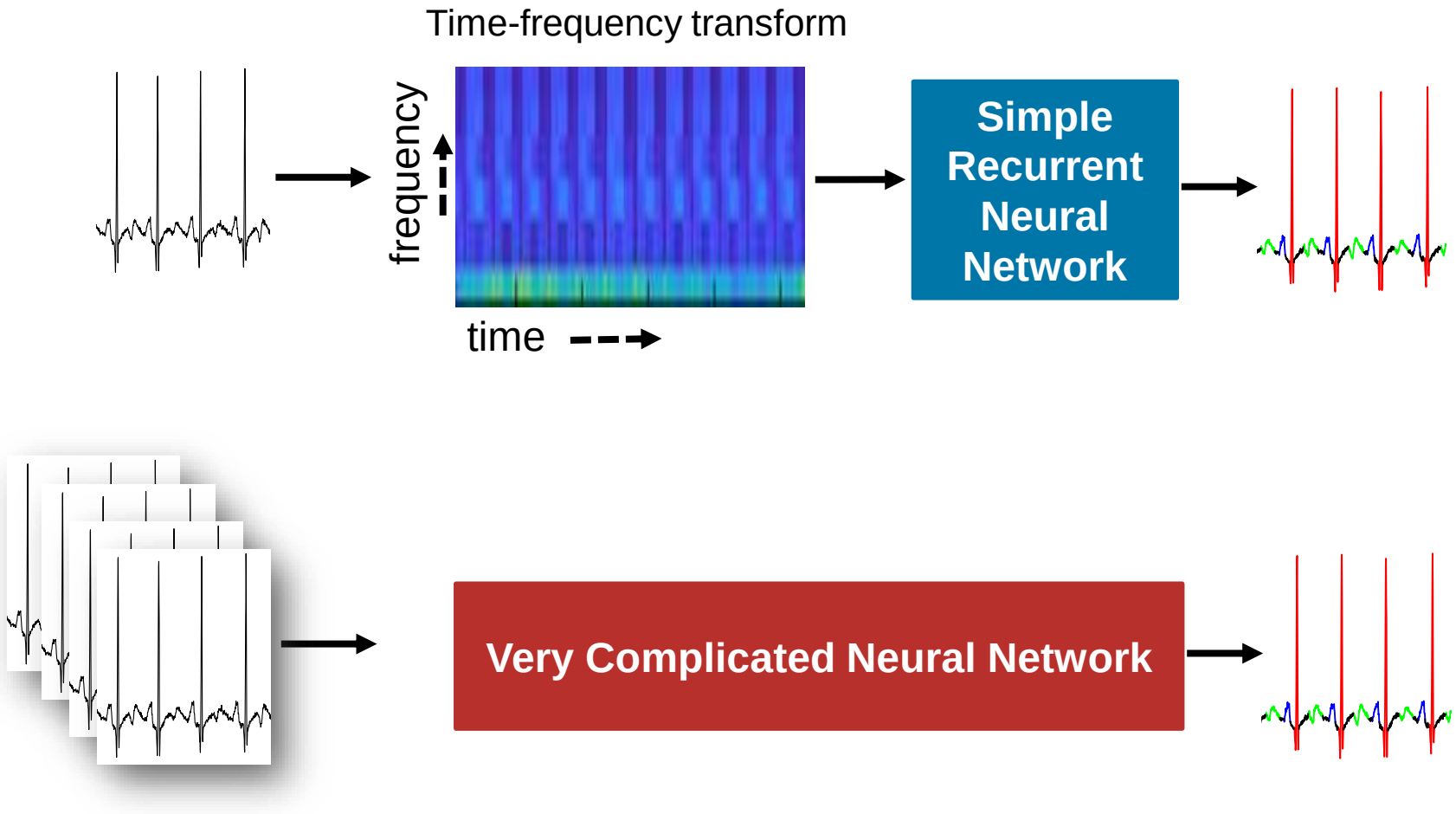


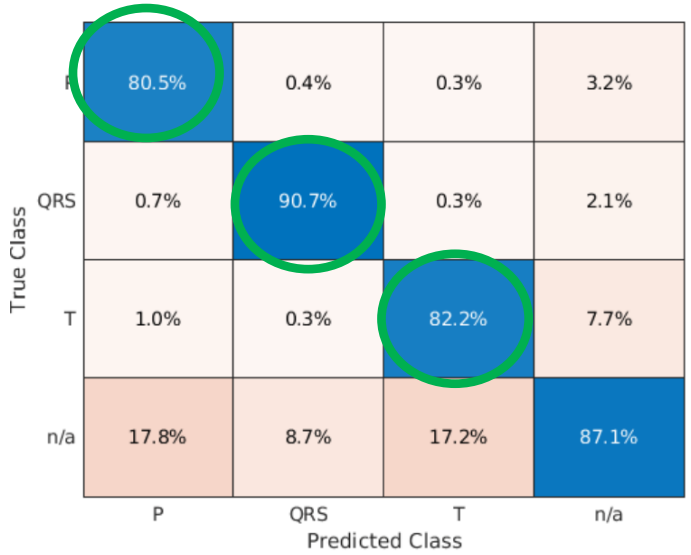






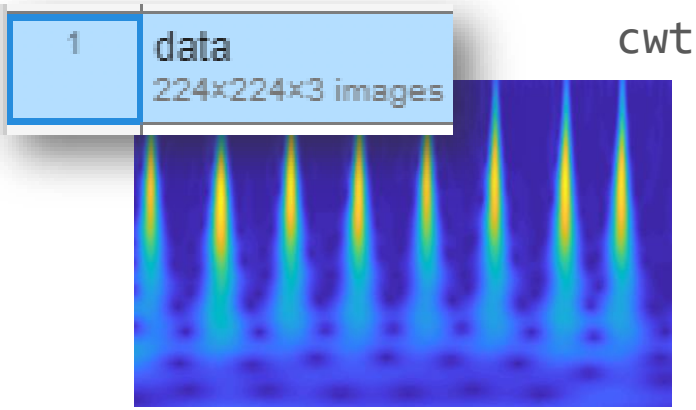
Feature extraction reduces model and data complexity





Domain experts are best placed to select feature extraction algorithm

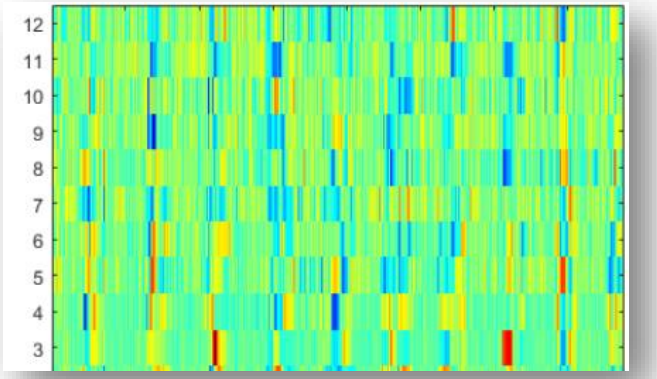
Model size, signal patterns



`cwt`

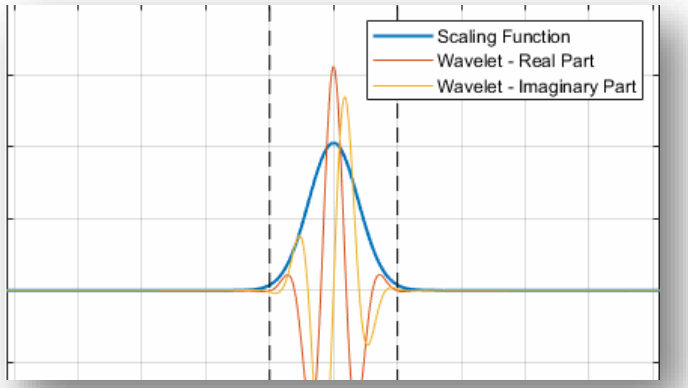
Application and signal type

`mfcc`



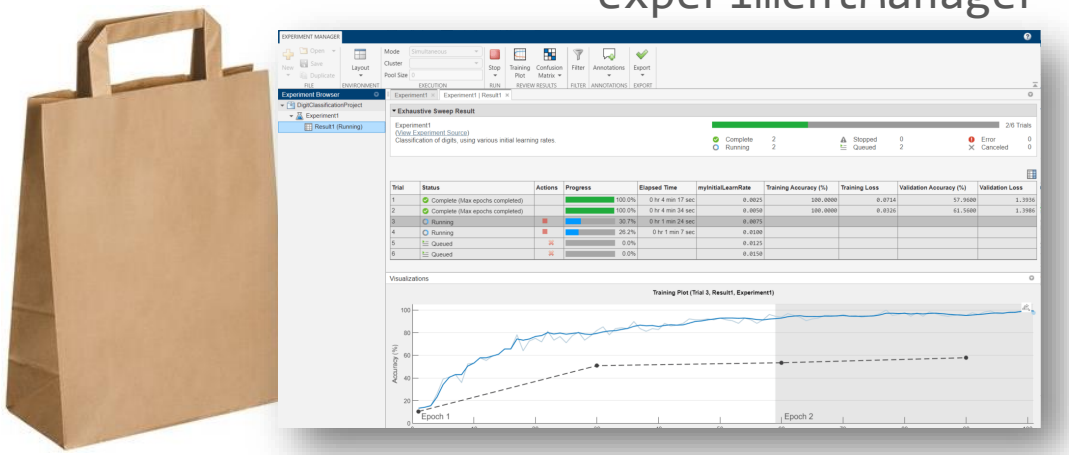
Automated methodology

`waveletScattering`



Test-based experiments

`experimentManager`



`signalFrequencyFeatureExtractor`

Feature Extraction – Handouts

MathWorks Wins Geoscience AI GPU Hackathon

The following post is from Akhilesh Mishra, Mil Shastri and Samvith V. Rao from MathWorks here to talk about their participation and in a Geoscience hackathon. Akhilesh and Mil are Applications Engineers and Samvith is the Industry Marketing Manager supporting the Oil and Gas industry.

Background

SEAM (SEG Advanced Modeling Corp.) is a petroleum geoscience industry body that fosters collaborations among industry, government, and academia to address major Geological challenges. Their latest event was a hackathon (SEAM AI Applied Geoscience GPU Hackathon) that sought to explore the use of AI to improve both qualitative and quantitative interpretation of geophysical images of Earth's interior, and speed up the applications using NVIDIA GPUs.

A total of 7 teams participated from all over the world, including commercial companies (Chevron, Total, Petrobras) and a mix of industry and university students. Each team was assigned a mentor who is an expert geoscientist working for a top oil and gas company.

The Challenge

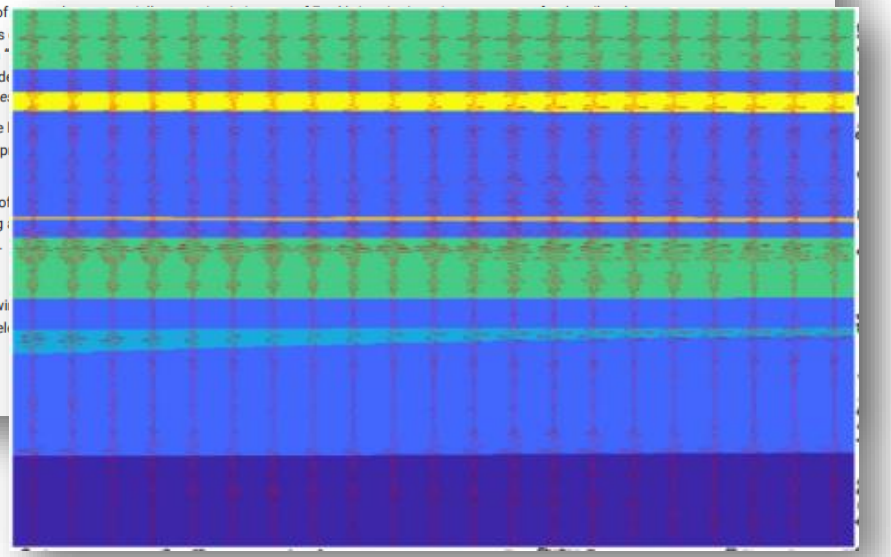
Geologic interpretation of
industry. Seismic images
summarized by the term "
and abandonment of unde
often called *seismic facie*

This process is still done
display. Successful interp
features.

The problem statement o
automatically, producing
up human interpretation.

The Data

We were given the followi
public and has been label



[MathWorks Deep Learning Blog Post]

Daihatsu Uses AI to Classify Engine Sounds

Challenge

Develop an AI solution that can judge the level of engine knocking sound, which only skilled workers could judge

Solution

Create classification models and easy-to-use interface with MATLAB, making it possible to examine features multiple times

Key Outcomes

- Performed knocking sound analysis with the same accuracy as skilled workers
- Increased AI expertise through MATLAB training
- Promoted visualization of AI and increased awareness of AI

Daihatsu used AI to identify knocking sounds from its engines.

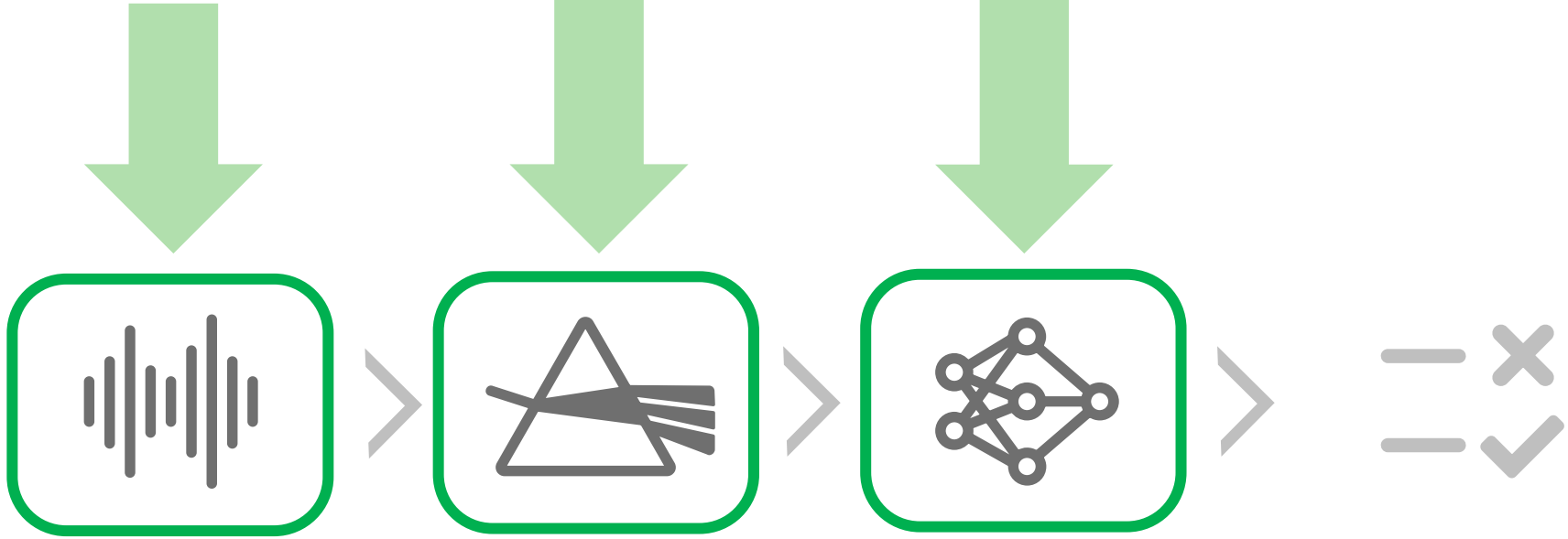
"Although we tried other programming languages, it was hard to implement. We decided to use MATLAB, which allows us to easily import the necessary data by dragging and dropping, and we could easily see the result by ourselves."
- Takuya Kumagae, Daihatsu Motor Co., Ltd.

Link to case study

[Daihatsu User Story]

Requiring smaller datasets multiplies the impact of data engineering

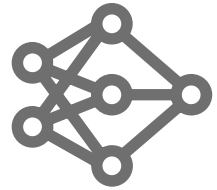
- Using **transfer learning**…
- …or **feature extraction** with simple models…
- …leads to requiring **much smaller labeled datasets** for model training



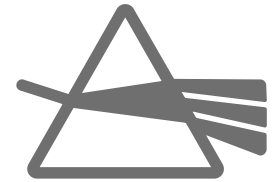
Data-Centric AI in Signal Processing Applications

Agenda – Three Practical Engineering Approaches

1. Transfer learning with pre-trained AI models



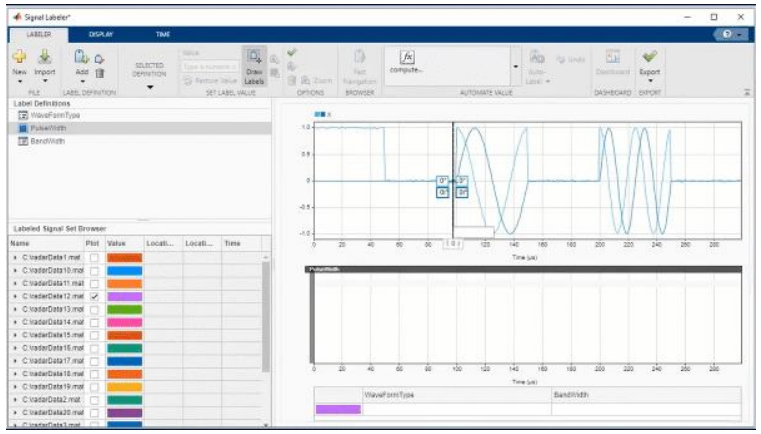
2. Feature extraction with simpler and smaller AI models



3. Better signal datasets, real or simulated

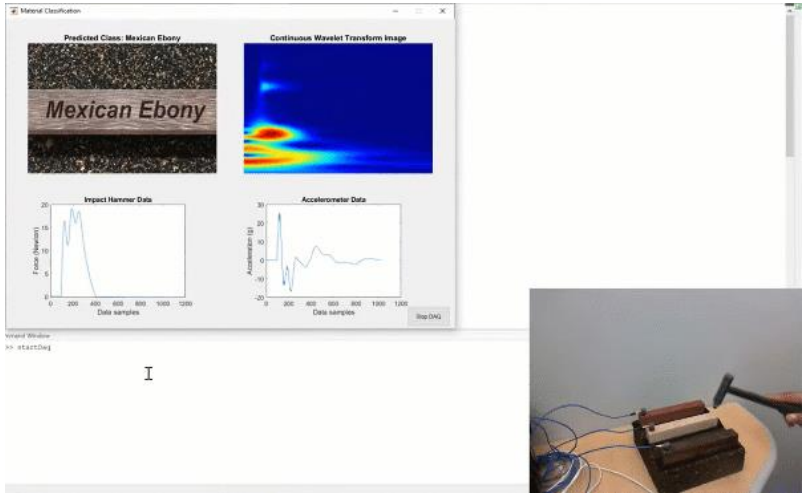
How can I enhance the quality of my training signal data?

Define accurate data labels



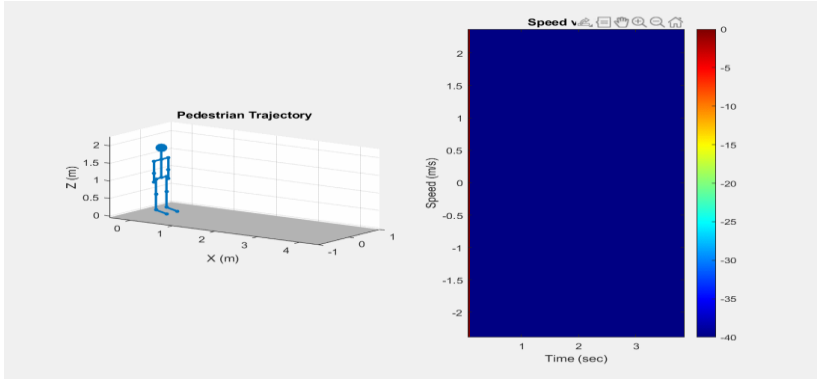
`signalLabeler`

Record and label new data via Apps and Hardware

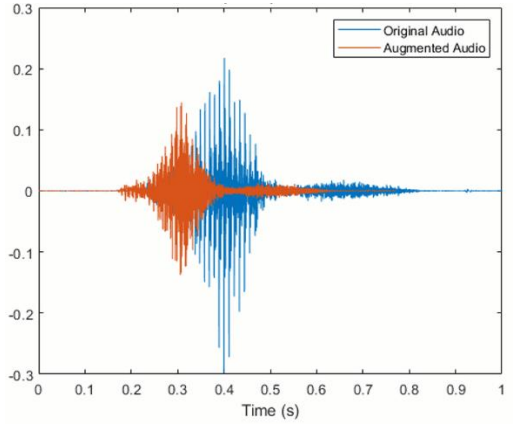


`App Designer`

Synthesize data via simulation



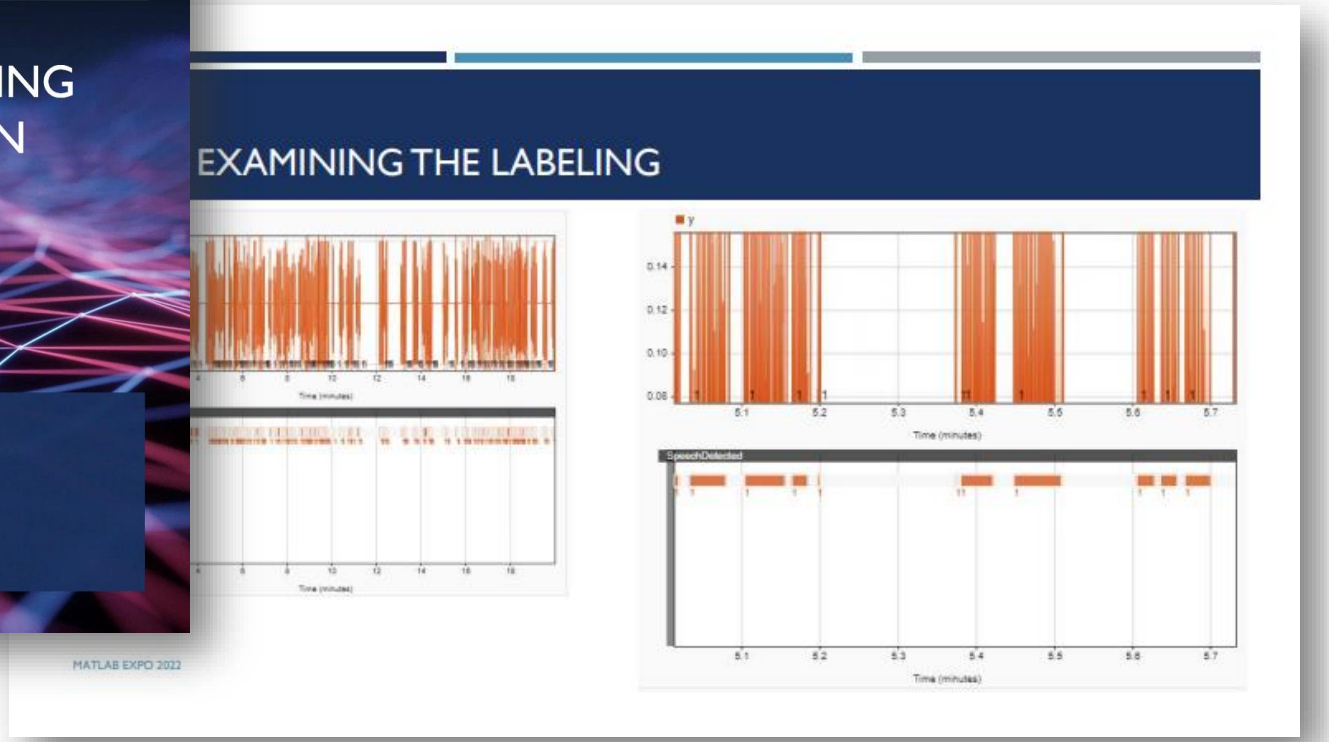
Augment data via signal processing



`audioDataAugmenter`

MATLAB EXPO 2022 Talk – Honeywell Technology Solutions

Automating Audio Labeling Workflow Using Pre-Trained Deep Learning Models for Voice Activity Detection

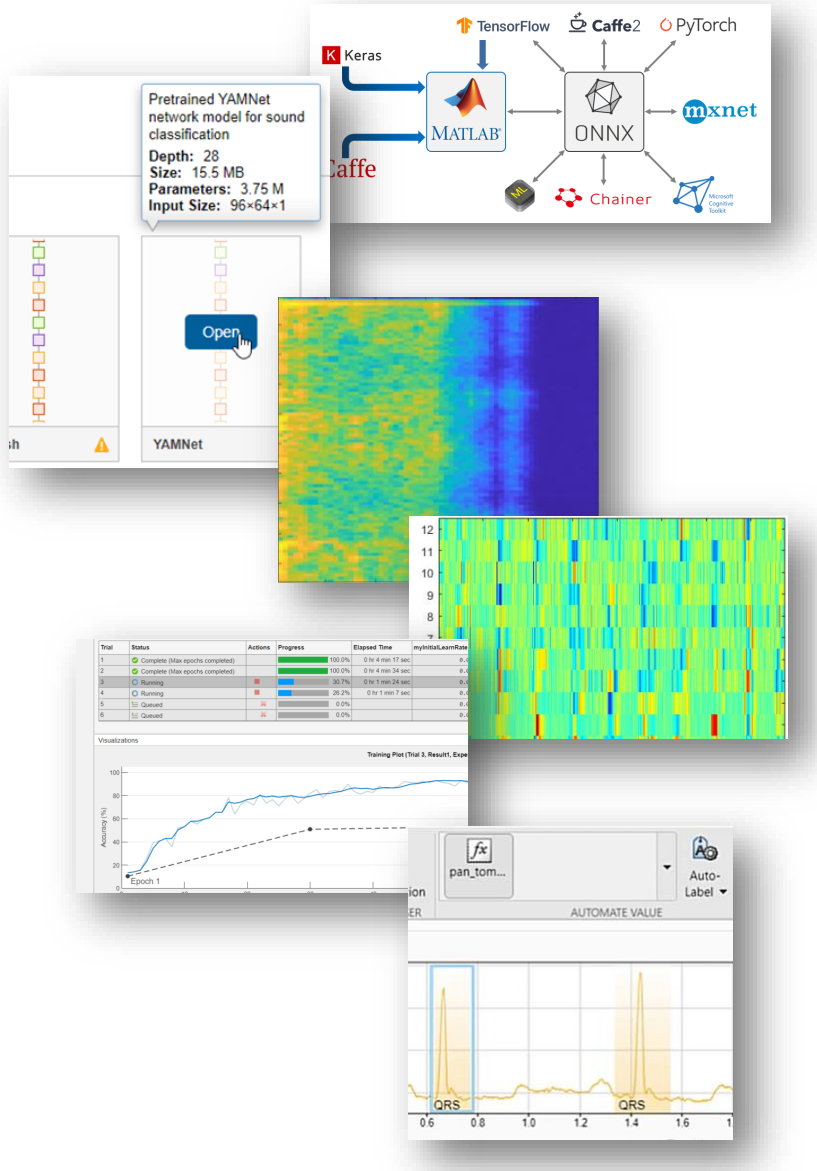


Track: **AI in Engineering**

Data-Centric AI accelerates AI adoption by domain experts

The “unbiggen AI” effect

- ~~Model Complexity~~ » MATLAB
- ~~Data Complexity~~ » Signal Processing
- ~~AI Expertise~~ » Domain Expertise



AI + Signal Processing

MATLAB EXPO

Thank you