Invited talk at Imperial College London

8 November 2024

Image source: www.economist.com

Elevating Health Awareness through Personalized Multimodal Sensing

[building a future]

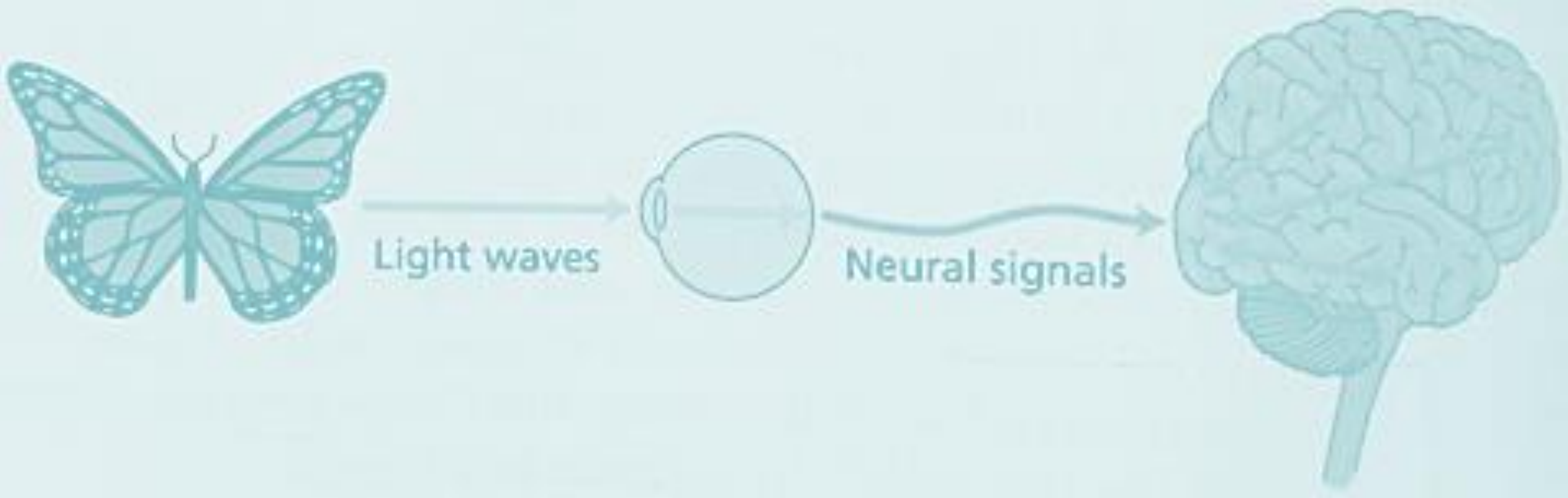
Mohammad Malekzadeh

Senior Research Scientist & Tech Lead

mmalekzadeh.github.io

NOKIA BELL LABS

Device Software Research



From **Sensation** **to** **Perception**

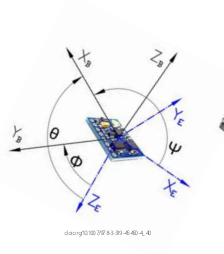
Sight, hearing, smell,
taste, touch, proprioception,
…

Memory, learning,
expectation, attention,
reflection, …

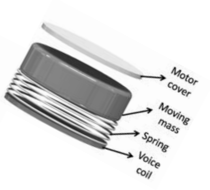
Machine Sensation

Acoustic Sensor

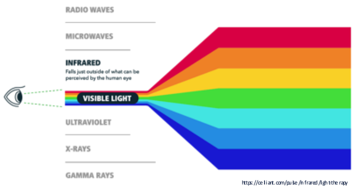
Optical Sensor

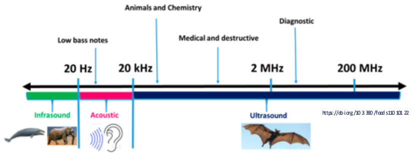


Motion Sensor



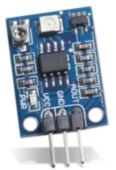
Haptic Sensor



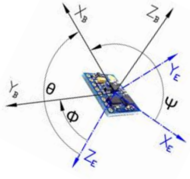


Machine Sensation

Acoustic Sensor

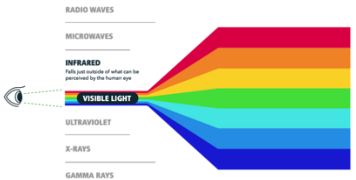


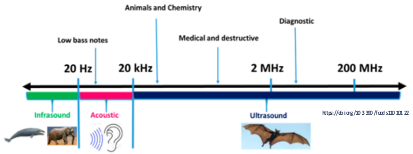
Optical Sensor



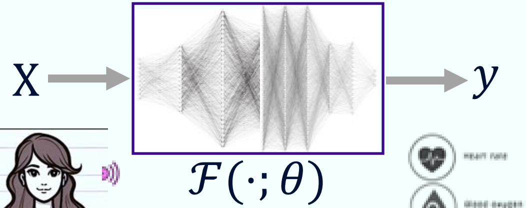
Motion Sensor

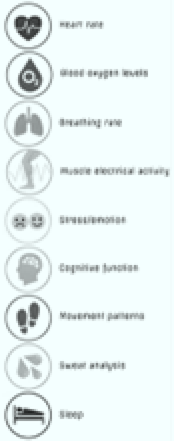
Haptic Sensor



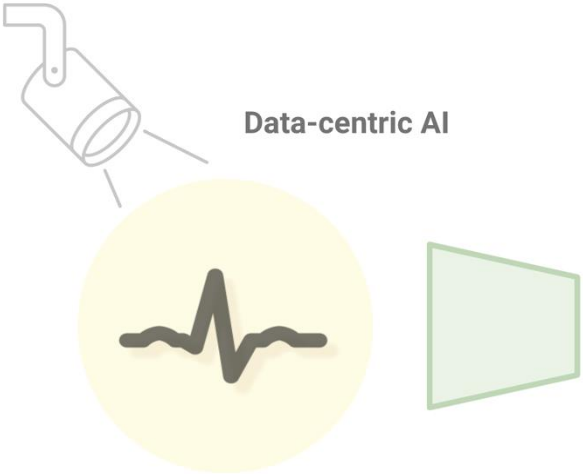


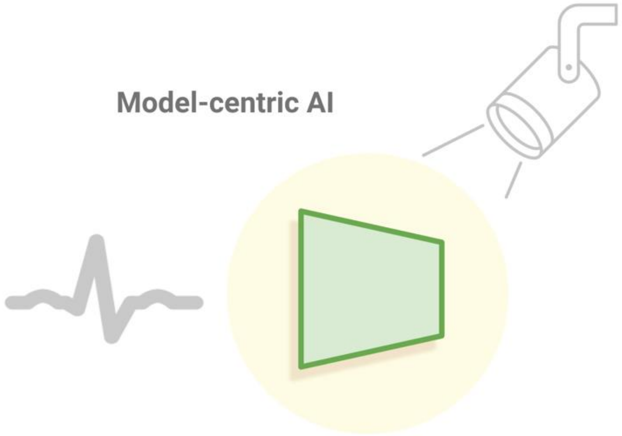
Machine Perception



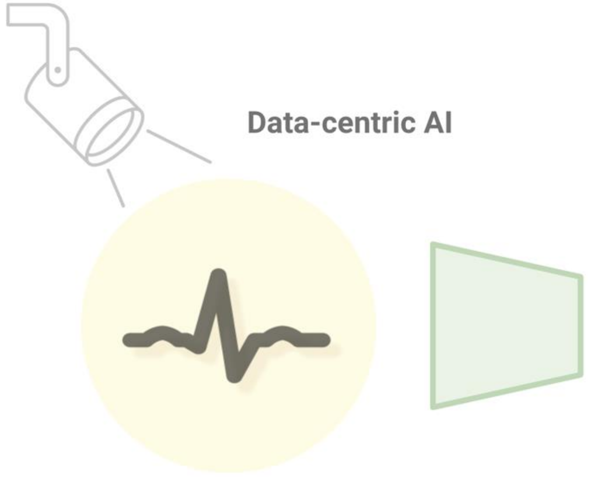


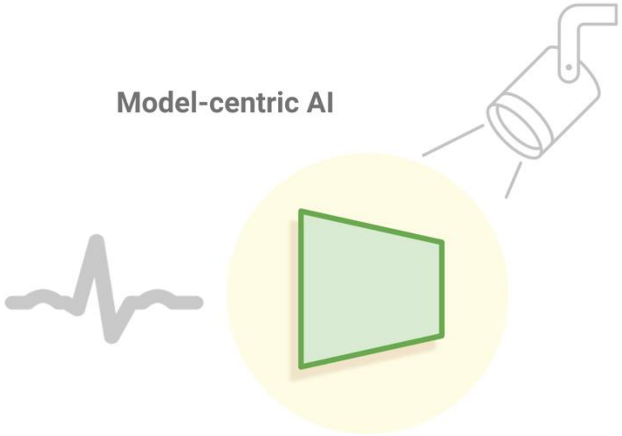
Machine Learning: Data-centric & Model-centric

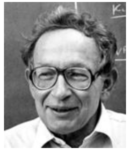




Machine Learning: Data-centric & Model-centric





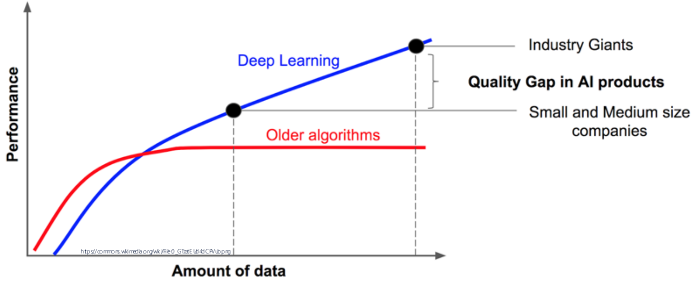


Anderson

“More is Different”

The **larger** is the **quantity**,
the more **unexpected** is the **quality**

Anderson, Philip W. "More is different: broken symmetry and the nature of the hierarchical structure of science." *Science* (1972)



Labelling!

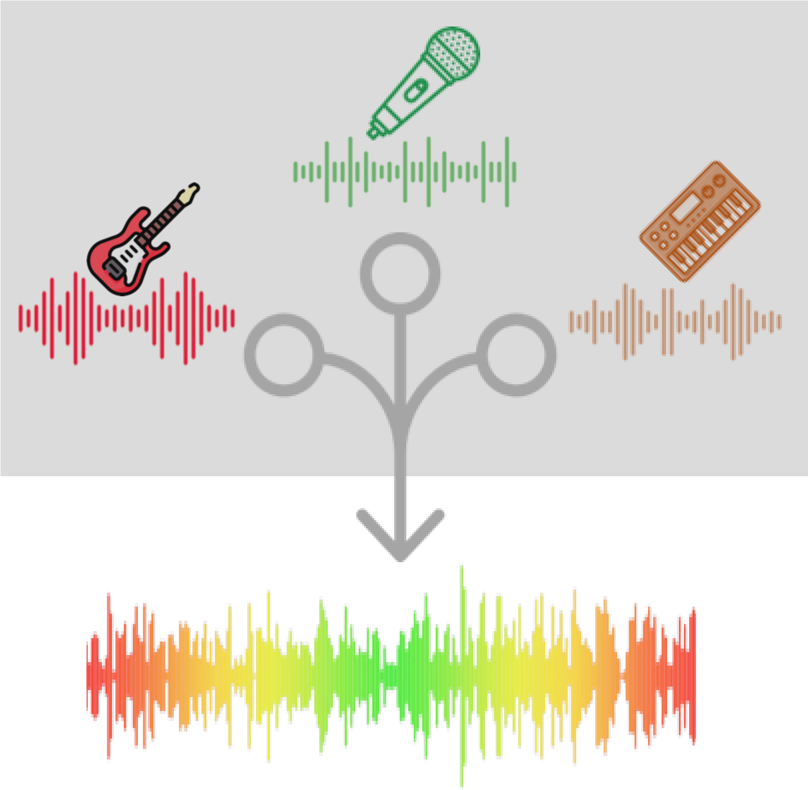
https://youtu.be/pbmTDxprHaM?feature=shared

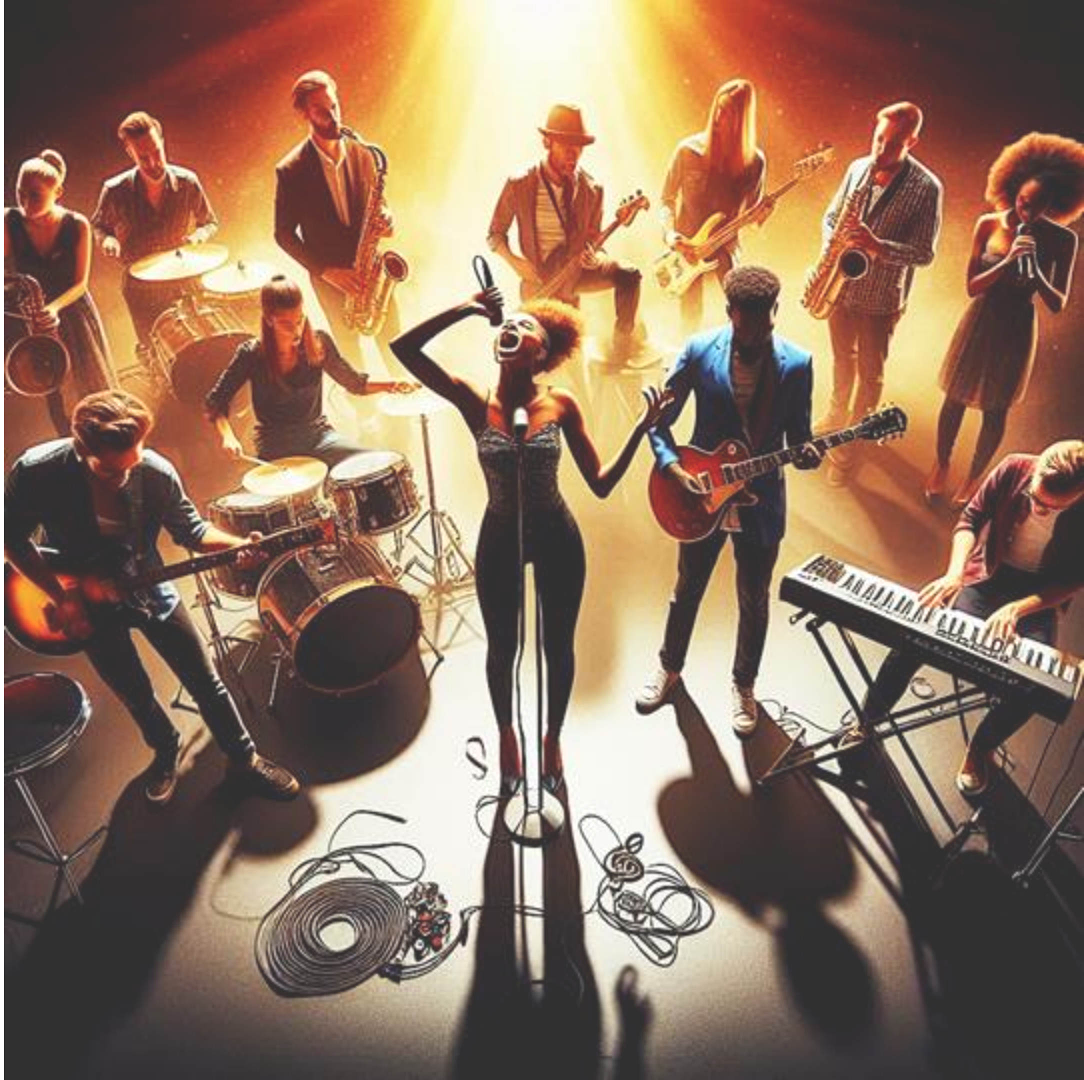
Annotation!

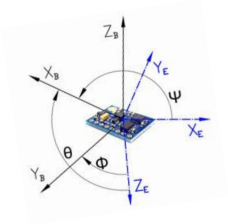
- ✓ Tree
- ✓ Bicycle
- ✓ Kid
- ✓ Female
- ✓ Male
- ✓ Family
- ✓ Park
- ✓ Green
- ✓ Outdoor
- ✓ Cycling
- ✓ Summer
- ✓ …

A happy family of three—parents and their young son—cycling together through a green park enjoying the outdoors in a sunny day and creating lasting memories with every ride. 🚴❤️🚴

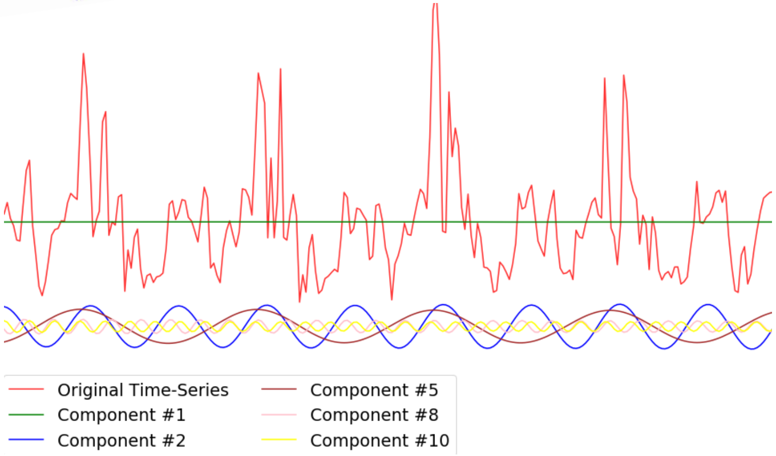
Acoustic Sensor

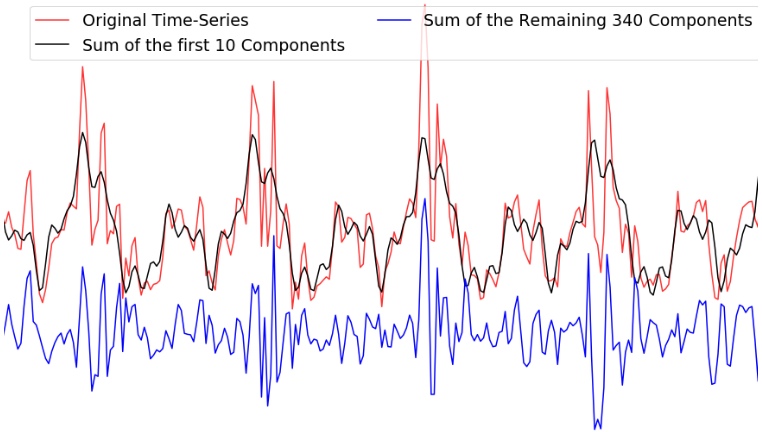






Motion Sensor: Accelerometer





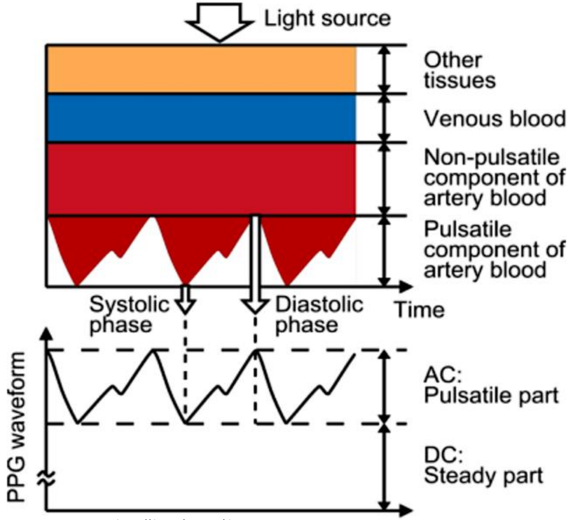
labelling

(A) *Walking*

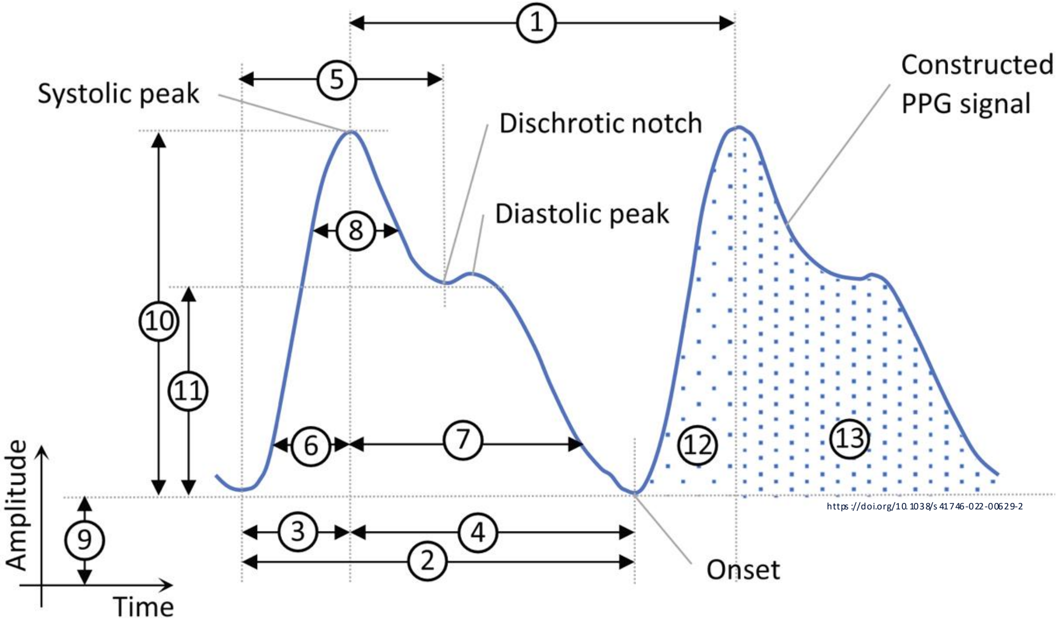
(B) An *<u>early thirties,</u>* *<u>tall</u>* & *<u>skinny</u>* *<u>man</u>* is *<u>walking</u>*

age height weight gender activity

Optical Sensor: PhotoPlethysmoGram (PPG)



https://doi.org/10.3390/electronics3020282

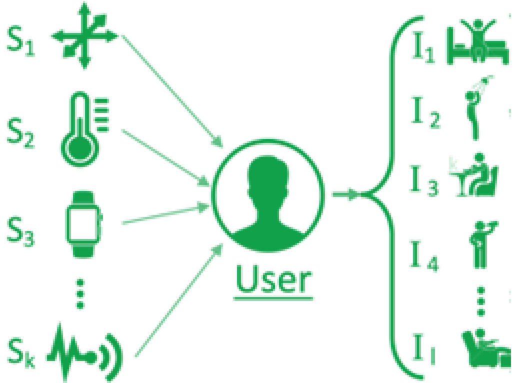


https://doi.org/10.1038/s41746-022-00629-2

Target information in the Data

$$\mathrm{X} \in \mathbb{R}^{t \times d \times c}$$

At each **time-point**, each **device**, generates samples coming from some **channels**.

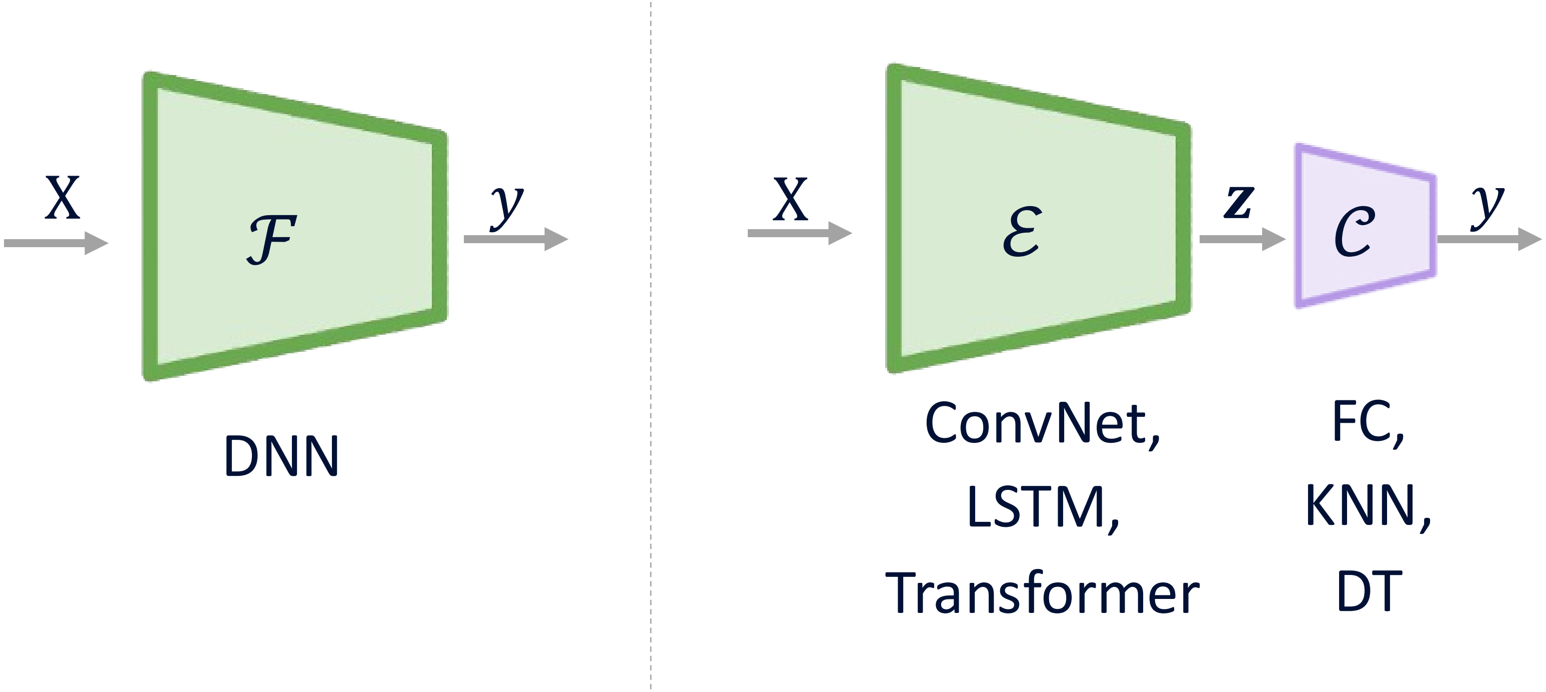


$$\mathrm{Y}^{\mathrm{i}} = \mathcal{F}^{i}(\mathrm{X})$$

Each **application** concerns only some specific **information**!

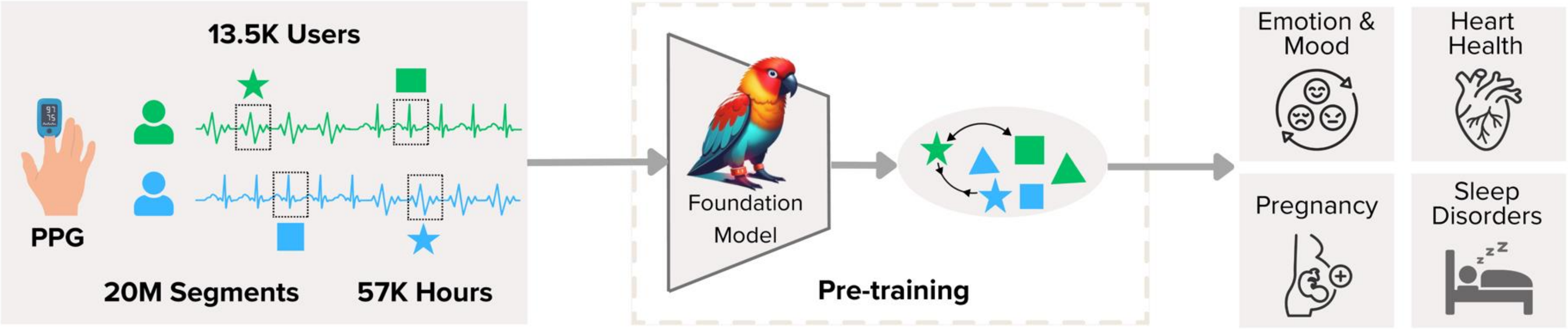
Deep Learning

an **encoder** to extract features and a **classifier** to make predictions



Learning from unlabeled PPG data

PaPaGei: Open Foundation Models for Optical Physiological Signals



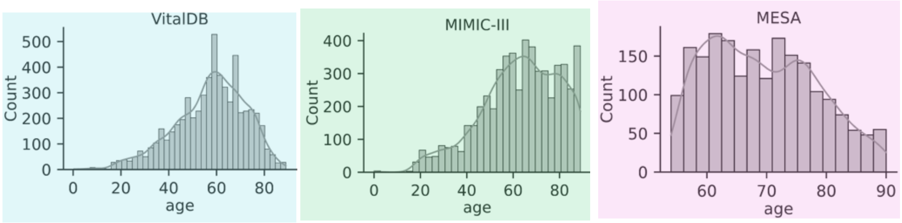
Arvind Pillai[2]*, Dimitris Spathis[1,3], Fahim Kawsar[1,4], Mohammad Malekzadeh[1]
[1]Nokia Bell Labs, UK, [2]Dartmouth College, USA,
[3]University of Cambridge, UK, [4]University of Glasgow, UK

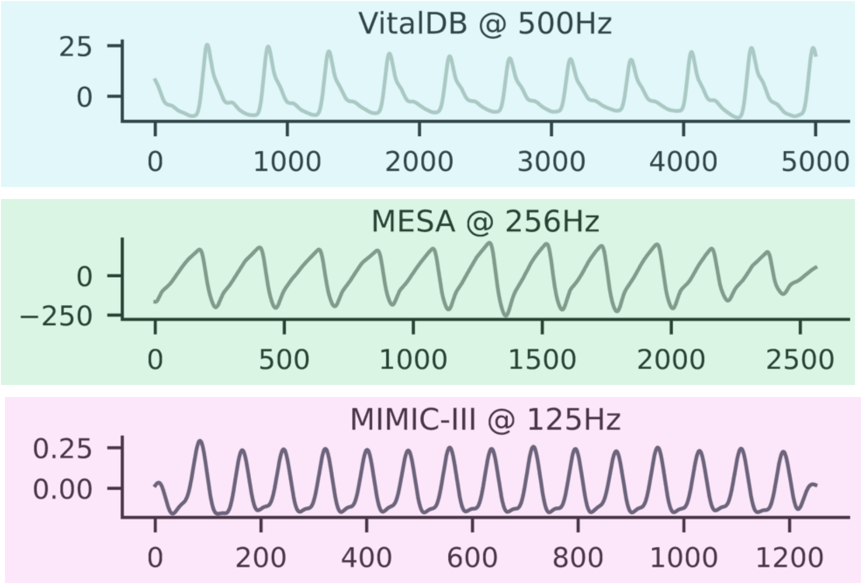
https://github.com/Nokia-Bell-Labs/papagei-foundation-model

PPG datasets

Publicly available!

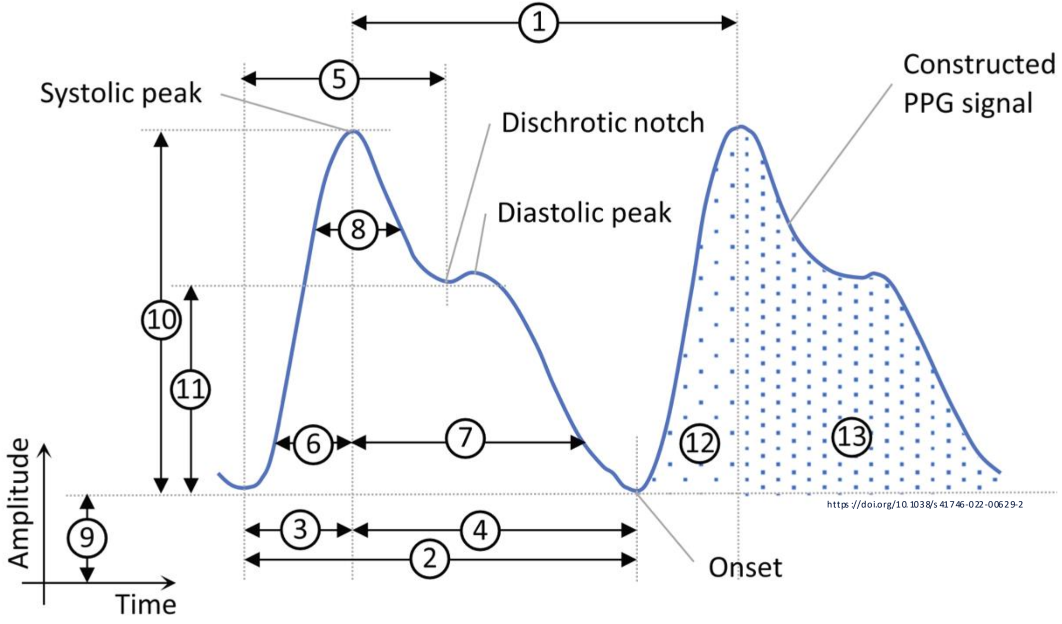
Dataset	#Subjects	#Segments	Hours
VitalDB	5,866	6,248,100	17,355
MIMIC-III	5,596	7,196,401	19,990
MESA	2,055	7,306,705	20,296
Total	13,517	20,751,206	57,641

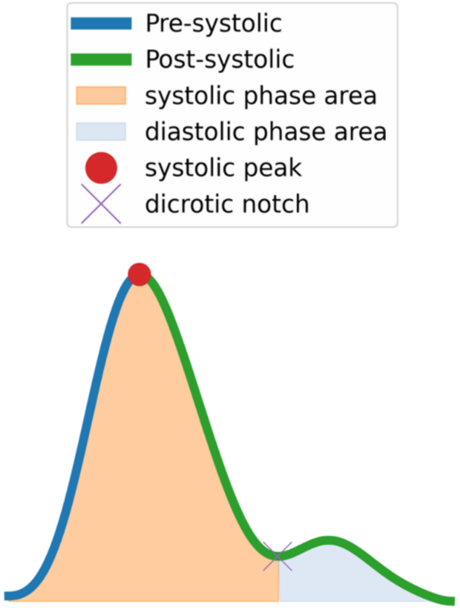




PPG characteristics

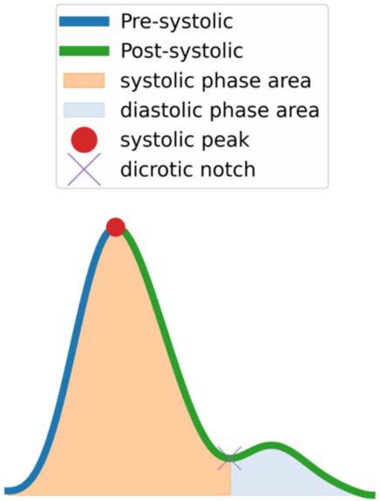
Better understanding of data for designing better models



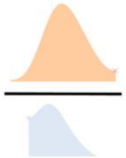


PPG characteristics

Better understanding of data for designing better models



IPA

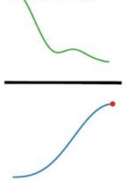


$$IPA(\mathbf{x}) = \frac{\int_0^{\hat{n}} \mathbf{x}\, dn}{\int_{\hat{n}}^{n} \mathbf{x}\, dn}$$

Inflection **P**oint **A**rea ratio:

the ratio of systolic to diastolic areas defined by the dicrotic notch

SVRI



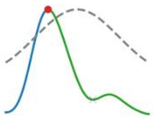
$$sVRI(\mathbf{x}) = \frac{sys \sum_{i=sys}^{n} x_i}{(n - sys) \sum_{i=1}^{sys} x_i}$$

sys is the systolic peak

Stress-induced **V**ascular **R**esponse **I**ndex

the ratio of mean PPG signal between post- to pre-systolic phases

SQI



$$SQI(\mathbf{x}) = \frac{1}{W} \sum_{w} \frac{m_3}{m_2^{3/2}}$$

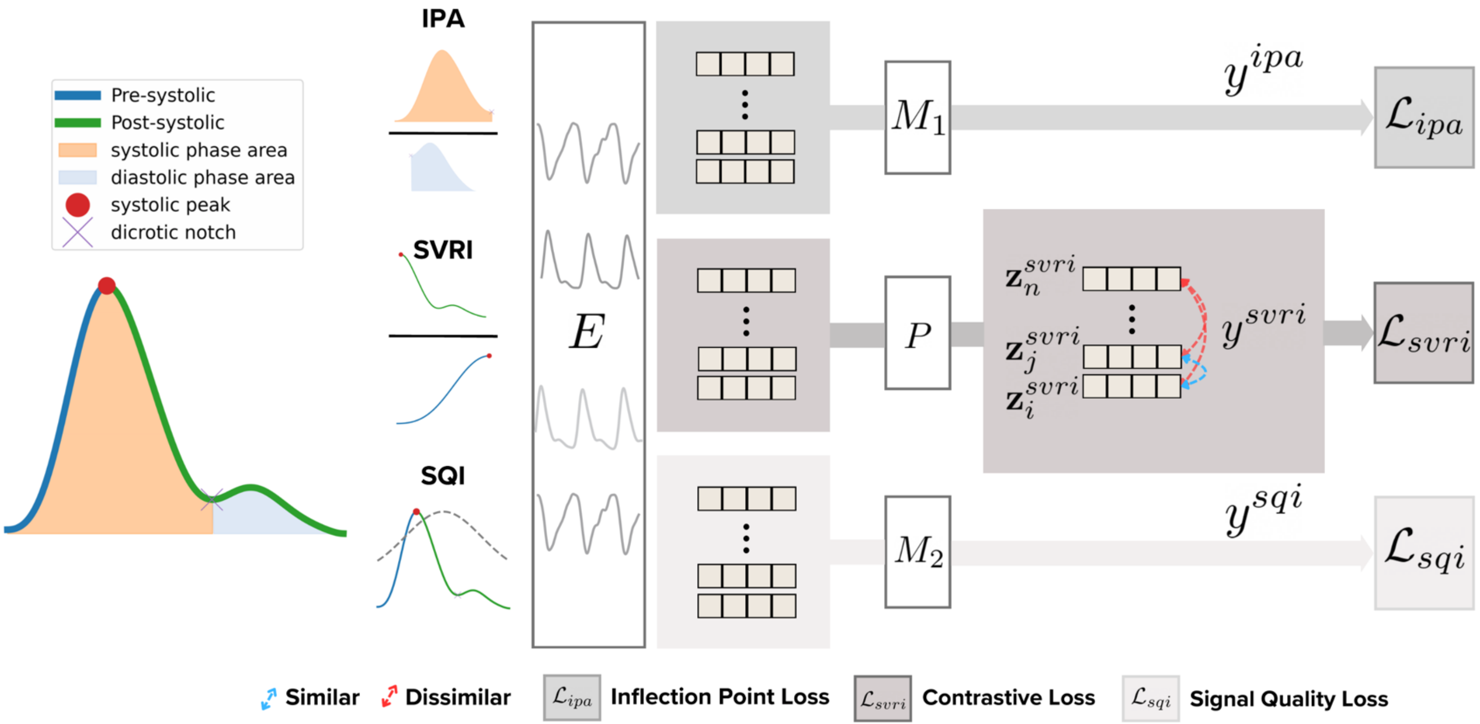
$$m_i = \frac{1}{5\times f} \sum_{j=1}^{5\times f} (x[j] - \mu_x[j])^i$$

Signal **Q**uality **I**ndex

skewness of the signal as an indicator of quality

PPG characteristics

Better understanding of data for designing better models

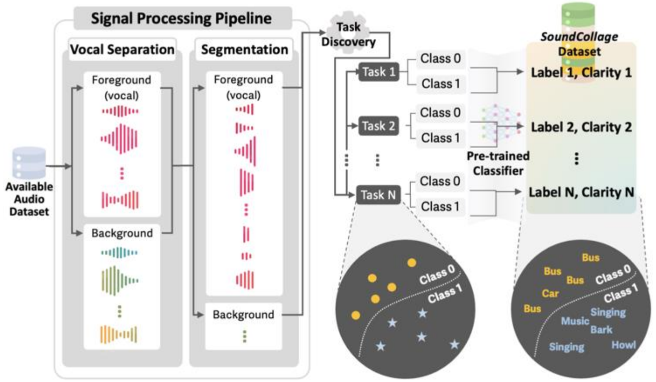


Linear Probing

Classification - AUROC (↑)	REGLE (0.07M) (Yun et al., 2024)	Chronos (200M) (Ansari et al., 2024)	Moment (385M) (Goswami et al., 2024)	PAPAGEI-P (5M)	PAPAGEI-S (5.7M)
ICU Admission	0.57 [0.52-0.62]	0.73 [0.68-0.80]	0.72 [0.70-0.80]	0.73 [0.67-0.78]	**0.79** [0.75-0.82]
Mortality	0.55 [0.52-0.59]	**0.68** [0.65-0.71]	0.67 [0.63-0.71]	0.67 [0.63-0.71]	0.67 [0.63-0.70]
Smoker	0.54 [0.47-0.59]	0.62 [0.57-0.67]	0.62 [0.56-0.67]	**0.64** [0.58-0.69]	0.61 [0.56-0.66]
Pregnancy stage	0.64 [0.57-0.63]	**0.81** [0.79-0.82]	0.76 [0.74-0.78]	0.74 [0.72-0.76]	0.78 [0.75-0.80]
Hypertension	0.47 [0.34-0.58]	0.57 [0.43-0.71]	0.75 [0.64-0.85]	0.74 [0.55-0.90]	**0.77** [0.68-0.87]
Sleep Disordered Breathing	0.45 [0.30-0.61]	0.58 [0.35-0.82]	0.45 [0.23-0.66]	0.54 [0.23-0.66]	**0.70** [0.57-0.84]
Mood Disturbance	0.41 [0.16-0.66]	0.43 [0.21-0.68]	0.55 [0.33-0.78]	0.53 [0.27-0.78]	**0.56** [0.33-0.77]
Valence	0.55 [0.52-0.57]	0.56 [0.53-0.59]	**0.57** [0.54-0.59]	0.53 [0.51-0.56]	0.56 [0.54-0.59]
Arousal	0.51 [0.52-0.58]	0.57 [0.54-0.60]	0.56 [0.53-0.58]	**0.58** [0.55-0.61]	0.55 [0.52-0.57]
Average	0.52 ± 0.06	0.62 ± 0.10	0.63 ± 0.09	0.63 ± 0.08	**0.67 ± 0.09**
Regression - MAE (↓)					
Apnea/Hypopnea Index > 3%	15.54 [14.20-16.69]	14.06 [13.05-15.16]	14.23 [13.04-15.42]	13.85 [12.43-15.49]	**12.97** [11.87-14.05]
Apnea/Hypopnea Index > 4%	12.64 [11.47-13.78]	11.57 [10.51-12.72]	11.80 [10.79-12.93]	11.24 [9.71-12.87]	**10.56** [9.59-11.62]
Gestation Age	7.28 [7.16-7.39]	**5.69** [5.54-5.85]	6.24 [6.10-6.37]	6.40 [6.21-6.59]	6.05 [5.91-6.17]
Systolic BP (VV)	15.88 [13.67-18.36]	17.24 [14.57-20.13]	14.71 [12.38-17.29]	19.11 [16.26-22.23]	**14.65** [12.50-16.78]
Diastolic BP (VV)	8.65 [7.16-10.27]	10.53 [8.91-12.19]	10.53 [8.91-12.19]	10.87 [9.10-12.98]	**8.29** [6.61-10.22]
Systolic BP (PPG-BP)	16.32 [13.87-19.13]	16.91 [13.31-19.34]	14.50 [11.98-17.31]	**13.60** [10.65-16.51]	14.39 [12.53-16.45]
Diastolic BP (PPG-BP)	9.30 [7.94-10.87]	10.26 [8.13-12.57]	9.53 [8.28-10.96]	8.88 [7.33-10.76]	**8.71** [7.18-10.01]
Average HR	6.88 [5.81-8.12]	8.51 [7.05-10.07]	4.41 [3.48-5.48]	**3.47** [2.74-4.32]	4.00 [3.34-4.67]
HR	16.35 [16.20-16.50]	9.65 [9.50-9.79]	**8.82** [8.68-8.96]	10.92 [10.80-11.04]	11.53 [11.40-11.66]
Average	12.09 ± 3.83	11.60 ± 3.60	10.43 ± 3.46	10.92 ± 4.25	**10.12 ± 3.47**

Learning from unlabeled Audio data

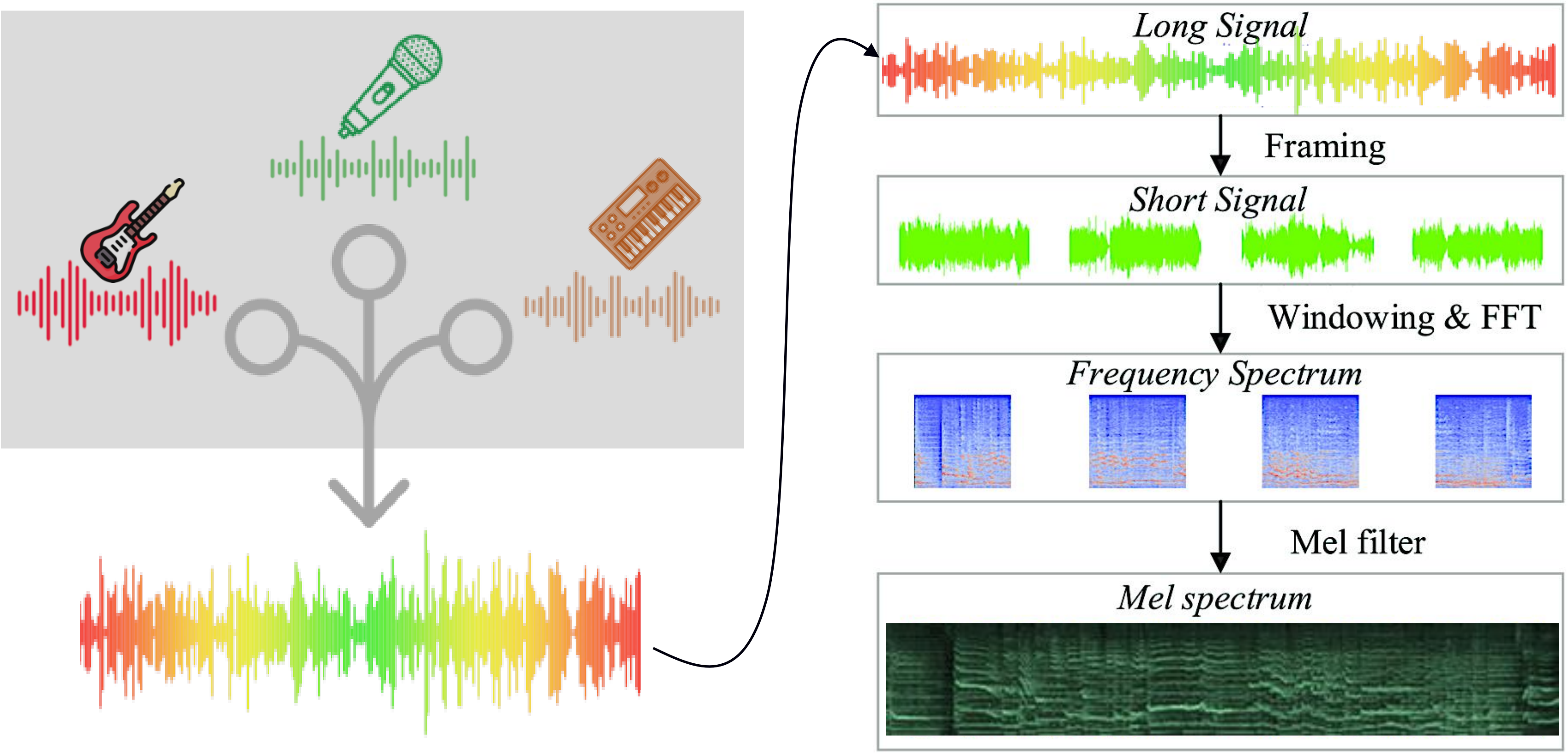
SoundCollage: Automated Discovery of New Classes in Audio Datasets



Ryuhaerang Choi*†, Soumyajit Chatterjee‡, Dimitris Spathis‡, Sung-Ju Lee†, Fahim Kawsar‡§, Mohammad Malekzadeh‡
‡Nokia Bell Labs, Cambridge, UK †KAIST, South Korea §University of Glasgow, UK

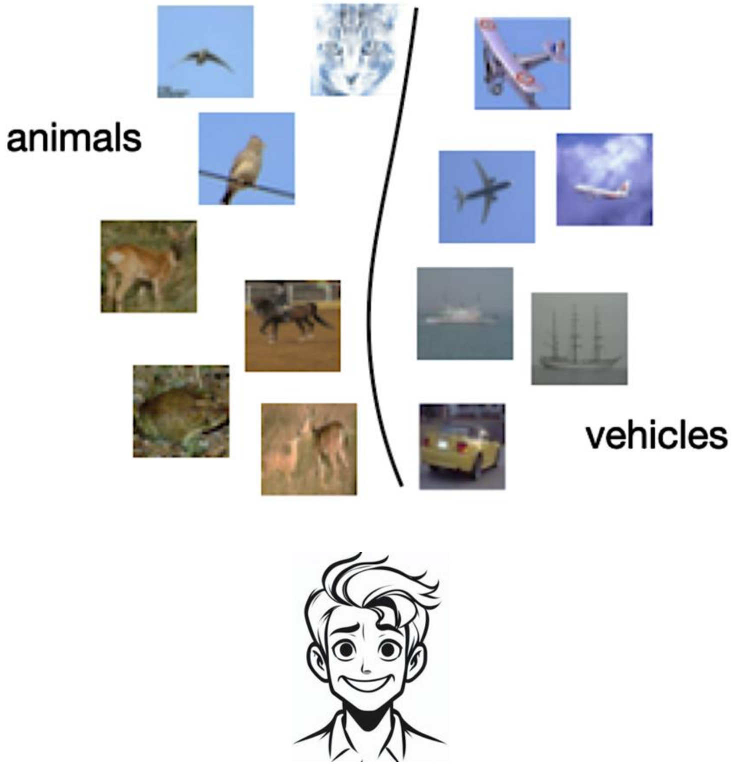
https://github.com/nokia-bell-labs/audio-class-discovery

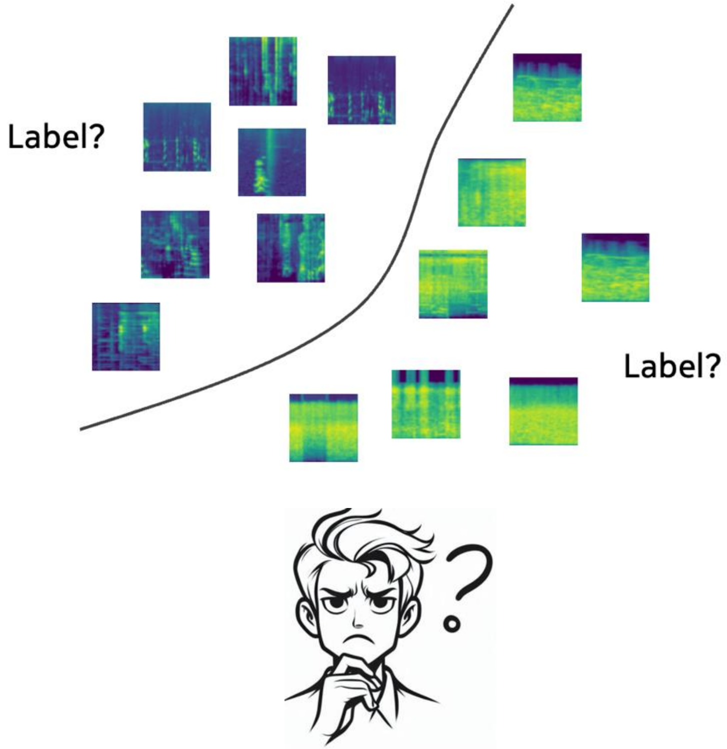
Audio to MFCC



https://doi.org/10.3390/su14052923

Annotating MFCC?!

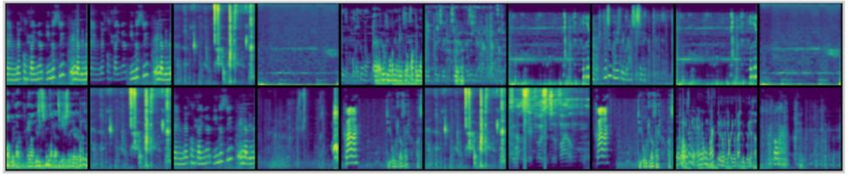




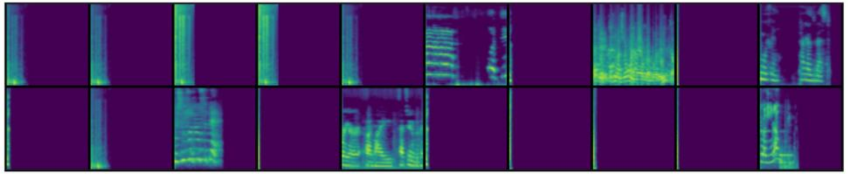
Annotating MFCC!

Singing

label= 1

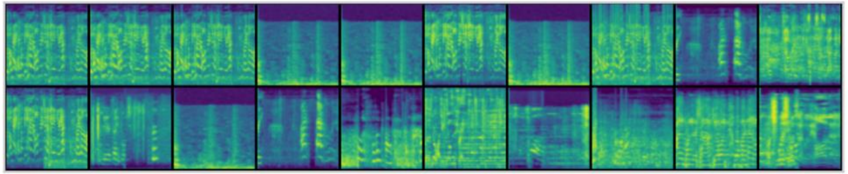


label= 0

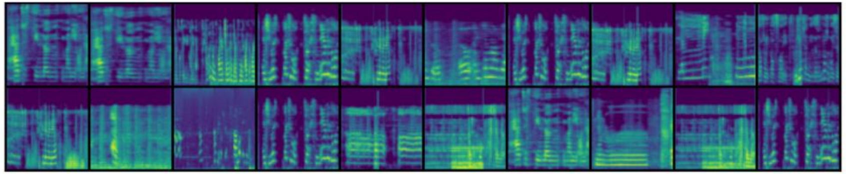


Inside a hall

label= 1

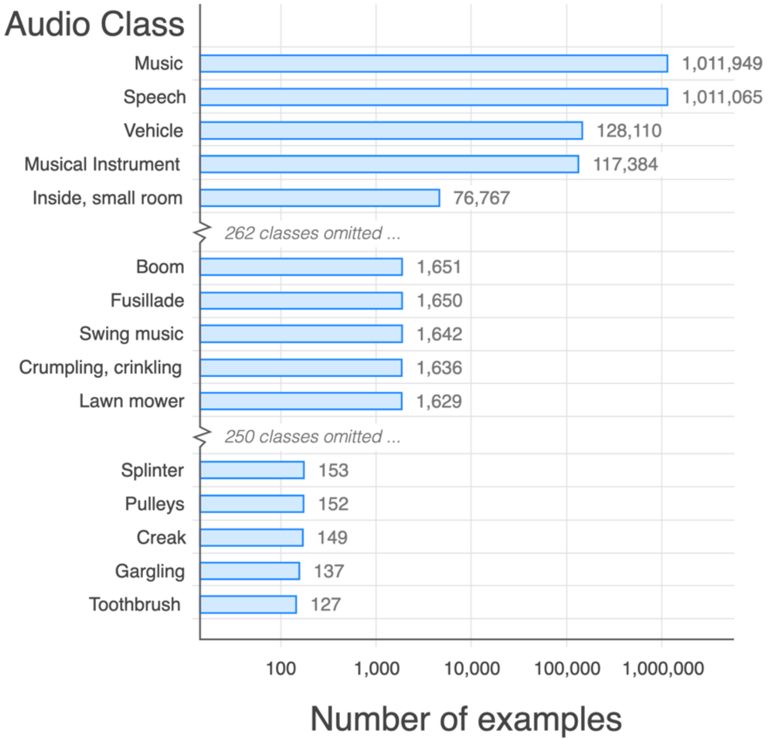


label= 0



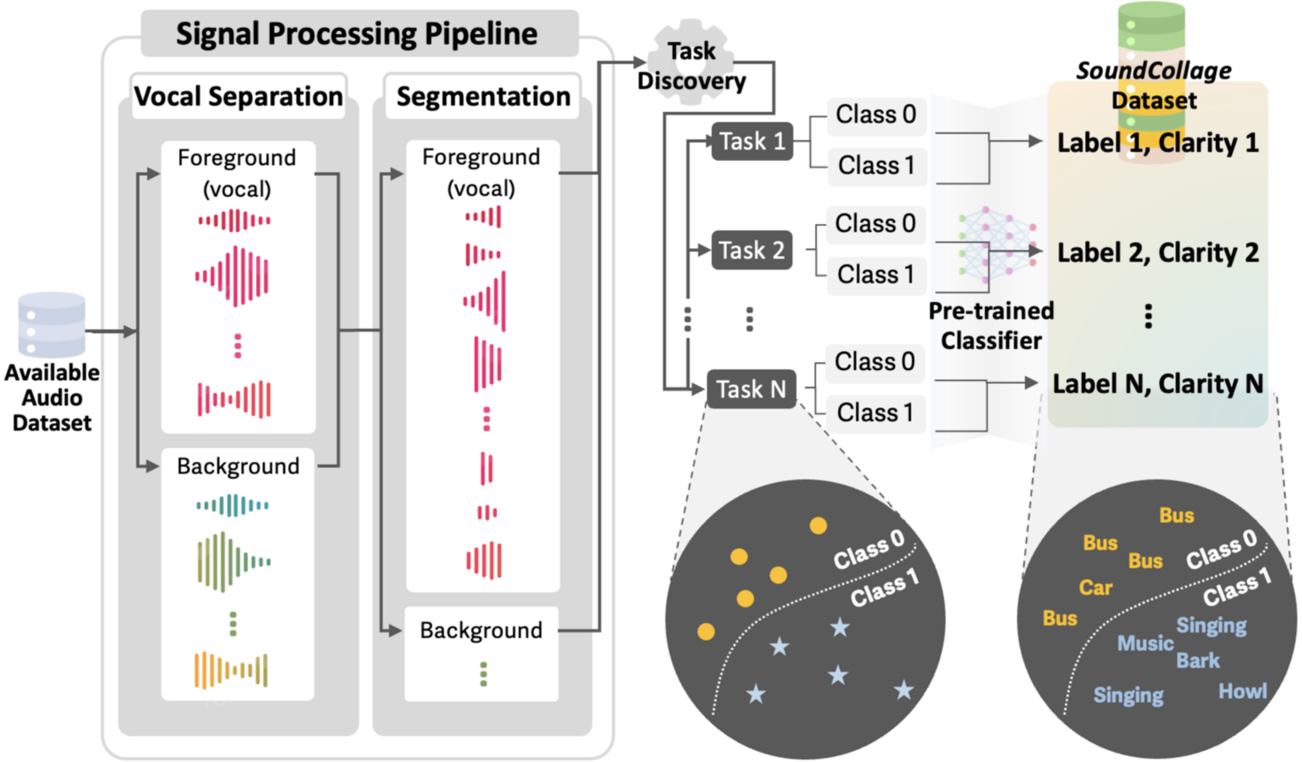
In-the-wild Audio

https://research.google.com/audioset/



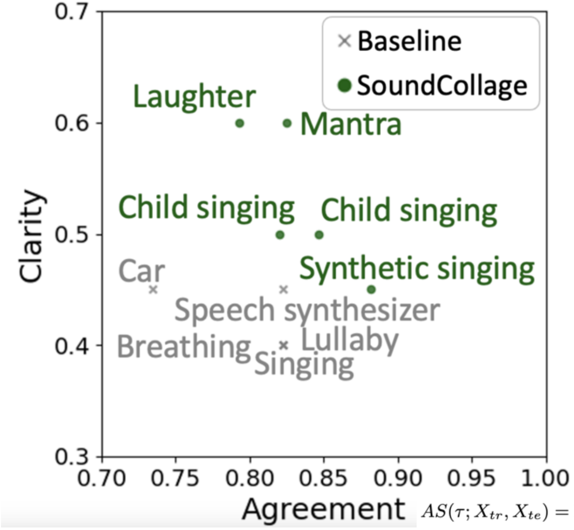
SoundCollage

Automated Discovery of New Classes in Audio Datasets



Class Discovery

$$C = \max\left(\frac{|N^0_{ij} - N^1_{ij}| - min\{N^0_{ij}, N^1_{ij}\}}{max\{N^0_i, N^1_i\}}, 0\right).$$



$$AS(\tau; X_{tr}, X_{te}) = \mathbb{E}_{w_1,w_2\sim\mathcal{P}(\mathcal{D}(X_{tr},\tau))}\mathbb{E}_{x\sim X_{te}}[(f(x;w_1) = f(x;w_2))].$$

Original	System	Acc (%)	Prec (%)	Rec (%)	F1 (%)	Label
Speech	Baseline	51.8 ± 4.6	51.7 ± 4.8	51.8 ± 4.6	51.4 ± 5.0	Speech synthesizer
	Ours	**75.6 ± 13.0**	**77.7 ± 14.1**	**75.6 ± 13.0**	**74.0 ± 13.8**	Mantra
Domestic sounds, home sounds	Baseline	56.2 ± 10.8	53.7 ± 16.0	56.2 ± 10.8	54.1 ± 13.4	Laughter
	Ours	**90.9 ± 5.3**	**91.1 ± 5.2**	**90.9 ± 5.3**	**90.9 ± 5.3**	Inside, large room or hall
Outside rural or natural	Baseline	62.2 ± 19.7	63.1 ± 21.4	62.2 ± 19.7	59.3 ± 21.1	Silence
	Ours	**91.6 ± 4.2**	**92.1 ± 4.8**	**91.6 ± 4.2**	**91.2 ± 4.6**	Musical instrument

Takeaways!

- ✓ Balanced Approach: Effective solutions in human sensing require both model-centric and data-centric approaches for optimal compute efficiency and model accuracy.
- ✓ Challenges of Data Labeling: Due to the sensitivity and complexity of health data, labeling is often impractical or cost-prohibitive.
- ✓ Unlabelled Data Learning: Learning from unlabelled data is essential, as it enables us to build meaningful feature representations without needing extensive labeled datasets. Robust feature extractors allow limited labeled data to support accurate classifiers, enabling effective model training with minimal labeling.
- ✓ Automated Pipelines for New Insights: Combining domain expertise with machine learning in automated workflows helps unlock new insights from existing unlabelled datasets. Large-scale, unlabelled data sources (like PPG and Audio) can be leveraged to develop innovative solutions.

We are building multi-sensory device platforms that learn, infer, and augment human behavior and health awareness.

For **Research Scientists** and **PhD interns**

we have opportunities

Mohammad **Malekzadeh**
Senior Research Scientist & Tech Lead
mmalekzadeh.github.io

NOKIA BELL LABS

Device Software Research