Engineering and Physical Sciences Research Council

ORCHID

Adaptive acoustic species classification in the field

Davide Zilli[1,2], Lucas Joppa[2], Geoff Merrett[1] and Alex Rogers[1]

Agents, Interaction, and Complexity Research Group
School of Electronics and Computer Science
University of Southampton

Computational Ecology and Environmental Sciences
Computational Science
Microsoft Research Cambridge

1 Species Classification

In recent years, the field of **computational sustainability** has striven to apply artificial intelligence techniques to solve ecological and environmental problems. In ecology, a key issue for the safeguarding of our planet is the monitoring of biodiversity. Automated acoustic recognition of species aims to provide a cost-effective method for biodiversity monitoring.

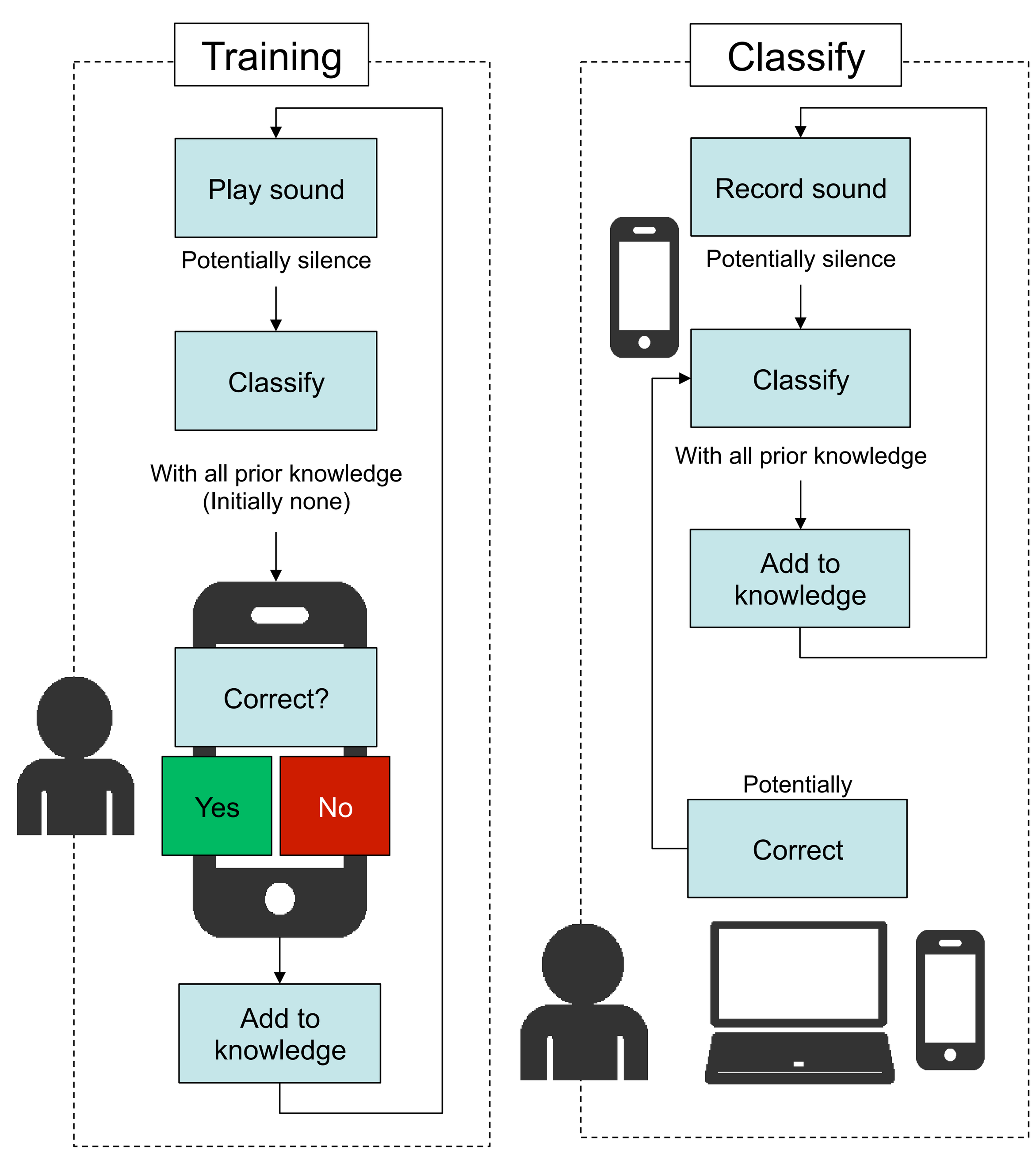
We investigate the development of a generic smartphone-based acoustic classification system that can be trained by ecologists with minimal effort and used in the field for the identification of virtually any species.

2 Mobile Classification System for Ecologists

Automated detection and classification of animal sounds can be useful to determine the presence of particular species or individuals in a certain area. However, building an automated classifier can be time consuming. Therefore, we propose a system that:

Can be
- trained with minimal effort
- deployed in the field at low cost
- matches a sound to a training sample

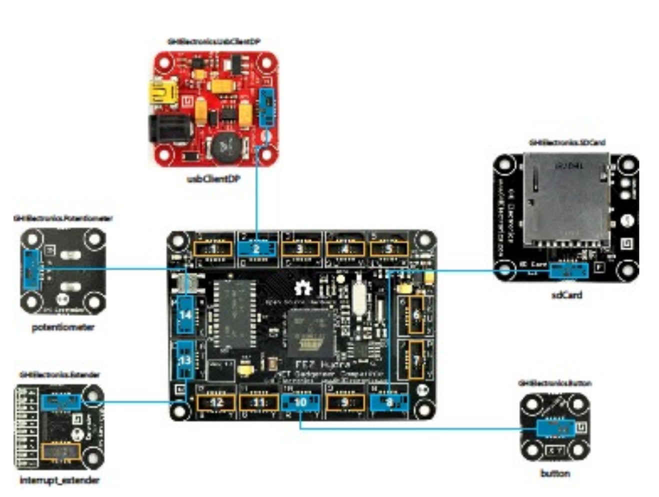
Doesn't require
- Extremely high accuracy
- Knowledge of the features to be extracted
- Knowledge of machine learning algorithms



3 .NET Gadgeteer for quick prototyping

The .NET Gadgeteer simplifies hardware prototyping by providing a high-level C#/VB interface to several hardware components.

Advantages	Disadvantages
Quick to build	Slow high-level VM layer
Easy to deploy	Not very cheap
Low power	Max interrupt ~ 2 kHz

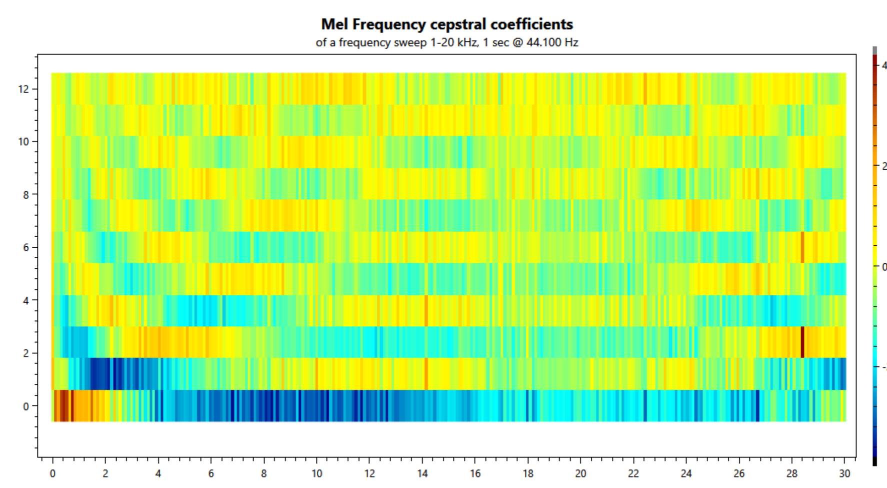


Prototype layout with FEZ Hydra motherboard

4 Feature extraction with MFCCs

Advantages	Disadvantages
Proven to match certain animals' hearing	Biased towards human hearing
Easy to cluster	Relatively expensive
Well established	

MFCCs pros and cons



Sample MFCCs from a frequency sweep.

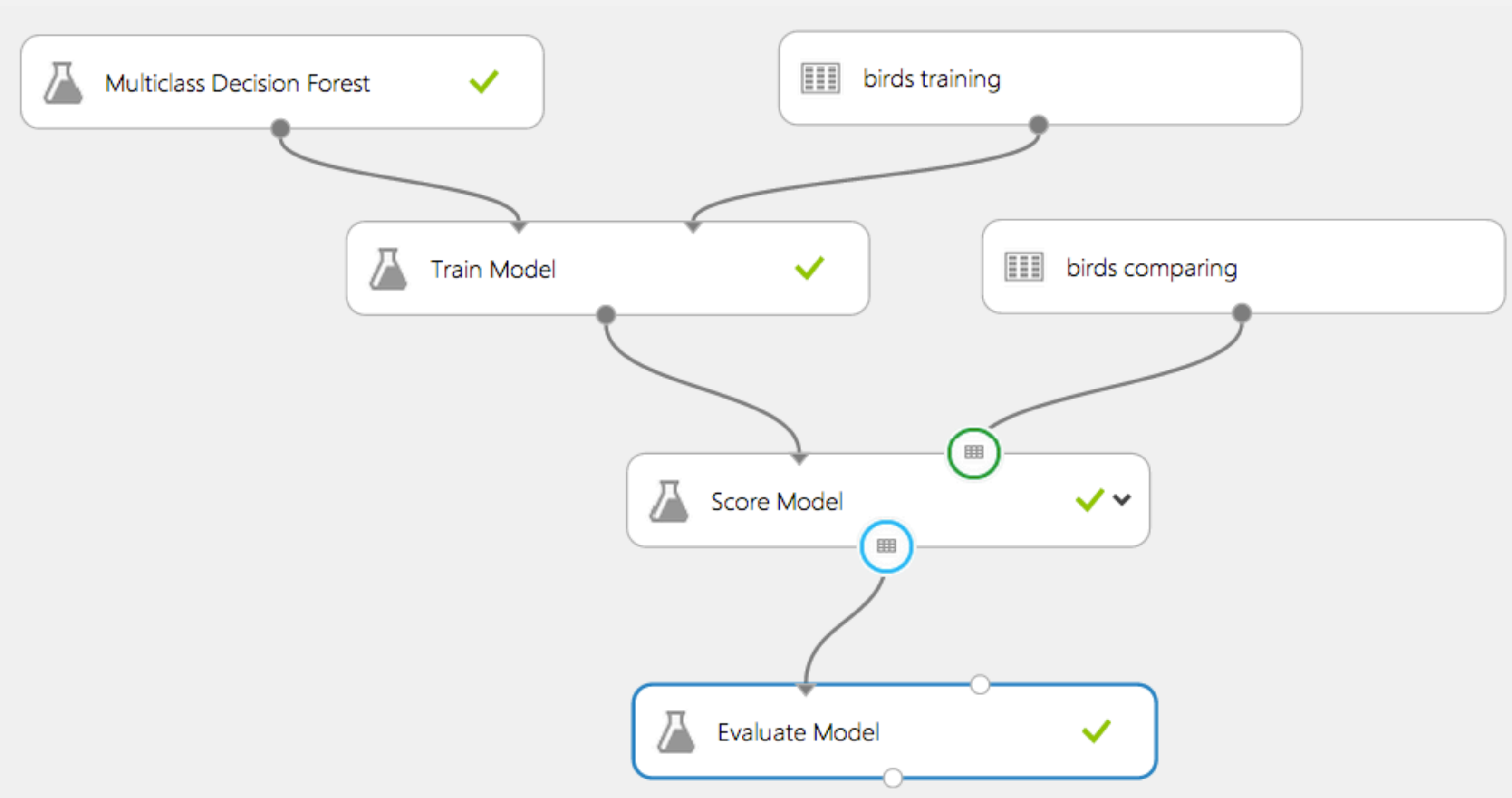
Firstly, we extract a number of features that are:

1. Unbiased to specific sounds
2. Sufficiently inexpensive to compute
3. Independent from volume

Mel Frequency Cepstral Coefficients (MFCCs) are an established technique from the speech recognition literature that represent the short-term power spectrum of a sound. They are calculated with the following steps:

- Compensate for any DC offset
- Frame and window the signal (with e.g. hamming window)
- Calculate the power density of the spectrum (DFT)
- Apply the mel filterbank to the spectrum and sum the energy
- Take the logarithm of mel filterbank energy
- Extract cepstral coefficients

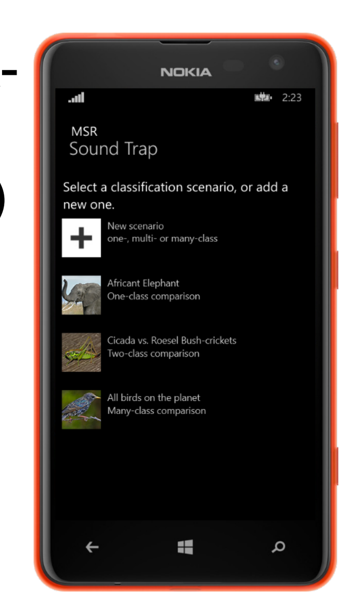
5 Classification on Azure ML



On Azure ML, a model can be trained with large datasets and on different classification algorithms. Results can then be easily evaluated in comparison to each other.

6 Deployment and future work

- System will be demonstrated on a sample elephant-recognition scenario.
- Elephants emit low-frequency calls (even subsonic)
- Deployment will be performed as a mobile phone application and on a .NET Gadgeteer board.
- System will be provided to other ecologists to train for different classes of sounds.
- Future applications could deviate from ecology.



BAE SYSTEMS · ACFR AUSTRALIAN CENTRE FOR FIELD ROBOTICS · SECURE · RESCUE GLOBAL Crisis and Disaster Response · UNIVERSITY OF Southampton