EMBARGOED until July 17th 2014: 7 am EDT; 12 midday UK time

PeerJ

Your Peers, Your Science
Academic Publishing is Evolving

Birdsongs automatically decoded by computer scientists

Scientists from Queen Mary University of London have found a successful way of identifying bird sounds from large audio collections, which could be useful for expert and amateur bird-watchers alike.

The analysis used recordings of individual birds and of dawn choruses to identify characteristics of bird sounds. It took advantage of large datasets of sound recordings provided by the British Library Sound Archive, and online sources such as the Dutch archive called Xeno Canto.

Publishing in the journal *PeerJ*, the authors describe an approach that combines feature-learning – an automatic analysis technique - and a classification algorithm, to create a system that can distinguish between which birds are present in a large dataset.

“Automatic classification of bird sounds is useful when trying to understand how many and what type of birds you might have in one location,” commented lead author Dr Dan Stowell from QMUL’s School of Electronic Engineering and Computer Science and Centre for Digital Music.

Dr Stowell was recently awarded a prestigious five-year fellowship from the Engineering and Physical Sciences Research Council (EPSRC) to develop computerized processes to detect multiple bird sounds in large sets of audio recordings.

“Birdsong has a lot in common with human language, even though it evolved separately. For example, many songbirds go through similar stages of vocal learning as we do, as they grow up, which makes them interesting to study. From them we can understand more about how human language evolved and social organization in animal groups,” said Dr Stowell.

He added: “The attraction of fully automatic analysis is that we can create a really large evidence base to address these big questions.”

The classification system created by the authors performed well in a public contest using a set of thousands of recordings with over 500 bird species from Brazil. The system was regarded as the best-performing ‘audio only’ classifier, and placed second overall out of entries from 10 research groups in the competition.

The researchers hope to drill down into more detail for their next project.

Dr Stowell says, “I'm working on techniques that can transcribe all the bird sounds in an audio scene: not just who is talking, but when, in response to whom, and what relationships are reflected in the sound, for example who is dominating the conversation.”

###

EMBARGOED until July 17th 2014: 7 am EST; 12 midday UK time (i.e. the date of publication)

Link to the PDF of this Press Release: http://bit.ly/StowellPeerJRelease

Link to the Press Preview of the Original Article (this link should only be used **BEFORE** the embargo ends): http://static.peerj.com/press/previews/2014/07/488.pdf (note: this is a low-resolution author proof and so may change slightly before publication)

Link to the Published Version of the article (quote this link in your story – the link will ONLY work **after** the embargo lifts): https://peerj.com/articles/488 - your readers will be able to **freely** access this article at this URL.

Citation to the article: Stowell D, Plumbley MD. (2014) Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. **PeerJ 2:e488**
http://dx.doi.org/10.7717/peerj.488

###

About PeerJ

PeerJ is an Open Access publisher of peer reviewed articles, which offers researchers a lifetime publication plan, for a single low price, providing them with the ability to openly publish all future articles for free. PeerJ is based in San Francisco, CA and London, UK and can be accessed at https://peerj.com/. PeerJ's mission is to help the world efficiently publish its knowledge.

All works published in PeerJ are Open Access and published using a Creative Commons license (CC-BY 4.0). Everything is immediately available—to read, download, redistribute, include in databases and otherwise use—without cost to anyone, anywhere, subject only to the condition that the original authors and source are properly attributed.

PeerJ has an Editorial Board of almost 900 respected academics, including 5 Nobel Laureates. PeerJ was the recipient of the 2013 ALPSP Award for Publishing Innovation.

PeerJ Media Resources (including logos) can be found at: https://peerj.com/about/press/

Queen Mary University of London

Queen Mary University of London is one of the UK's leading research-focused higher education institutions with some 17,840 undergraduate and postgraduate students.

A member of the Russell Group, it is amongst the largest of the colleges of the University of London. Queen Mary's 4,000 staff deliver world class degree programmes and research across 21 academic departments and institutes, within three Faculties: Science and Engineering; Humanities and Social Sciences; and the School of Medicine and Dentistry.

Queen Mary is ranked 11th in the UK according to the *Guardian* analysis of the 2008 Research Assessment Exercise, and has been described as 'the biggest star among the research-intensive institutions' by the *Times Higher Education*.

The College has a strong international reputation, with around 20 per cent of students coming from over 100 countries. Queen Mary has an annual turnover of £300m, research income worth £90m, and generates employment and output worth £600m to the UK economy each year.

The College is unique amongst London's universities in being able to offer a completely integrated residential campus, with a 2,000-bed award-winning Student Village on its Mile End campus.

###

Media Contacts

Note: If you would like to join the PeerJ Press Release list, visit: http://bit.ly/PressList

For the authors: Neha Okhandiar, Public Relations Manager, Queen Mary University of London, +44 (0)20 7882 7927, n.okhandiar@qmul.ac.uk

For PeerJ: email: press@peerj.com , https://peerj.com/about/press/

###

Abstract (from the article):

Automatic species classification of birds from their sound is a computational tool of increasing importance in ecology, conservation monitoring and vocal communication studies. To make classification useful in practice, it is crucial to improve its accuracy while ensuring that it can run at big data scales. Many approaches use acoustic measures based on spectrogram-type data, such as the Mel-frequency cepstral coefficient (MFCC) features which represent a manually-designed summary of spectral information. However, recent work in machine learning has demonstrated that features learnt automatically from data can often outperform manually-designed feature transforms. Feature learning can be performed at large scale and "unsupervised", meaning it requires no manual data labelling, yet it can improve performance on "supervised" tasks such as classification. In this work we introduce a technique for feature learning from large volumes of bird sound recordings, inspired by techniques that have proven useful in other domains. We experimentally compare twelve different feature representations derived from the Mel spectrum (of which six use this technique), using four large and diverse databases of bird vocalisations, classified using a random forest classifier. We demonstrate that in our classification tasks, MFCCs can often lead to worse performance than the raw Mel spectral data from which they are derived. Conversely, we demonstrate that unsupervised feature learning provides a substantial boost over MFCCs and Mel spectra without adding computational complexity after the model has been trained. The boost is particularly notable for single-label classification tasks at large scale. The spectro-temporal activations learned through our procedure resemble spectro-temporal receptive fields calculated from avian primary auditory forebrain. However, for one of our datasets, which contains substantial audio data but few annotations, increased performance is not discernible.

We study the interaction between dataset characteristics and choice of feature representation through further empirical analysis.