Detection of the bird song – a study on the collared flycatcher with the help of deep neural networks

Sándor Zsebők[1], Máté Ferenc Nagy-Egri[2], Gergely Gábor Barnaföldi[2], Miklós Laczi[1], Gergely Nagy[1], Éva Vaskuti[1], László Zsolt Garamszegi[1,3]

[1]Department of Systematic Zoology and Ecology, Eötvös Loránd University

[2]Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, HAS

[3]Institute of Ecology and Botany, Centre for Ecological Research, Hungarian Academy of Sciences

SZÉCHENYI 2020

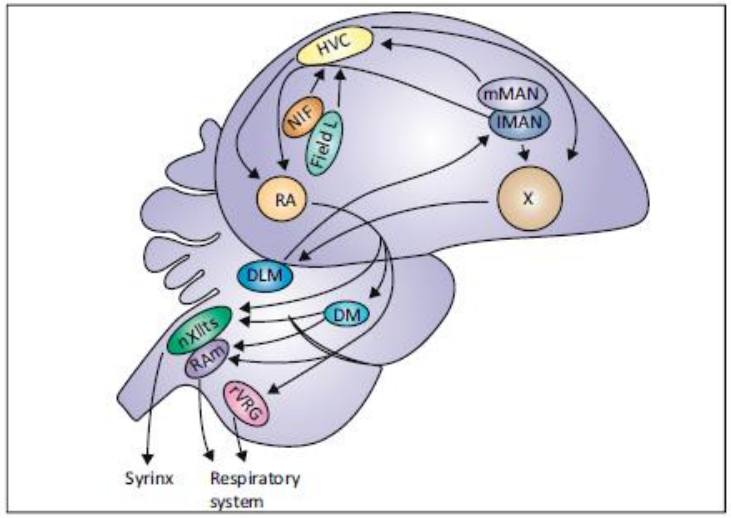
HUNGARIAN GOVERNMENT

European Union
European Social Fund

INVESTING IN YOUR FUTURE

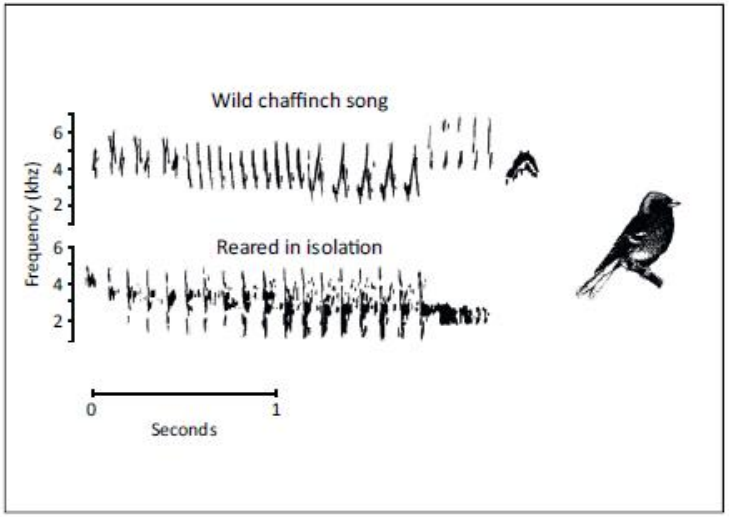
Main biological questions in bioacoustics

(A) Mechanism

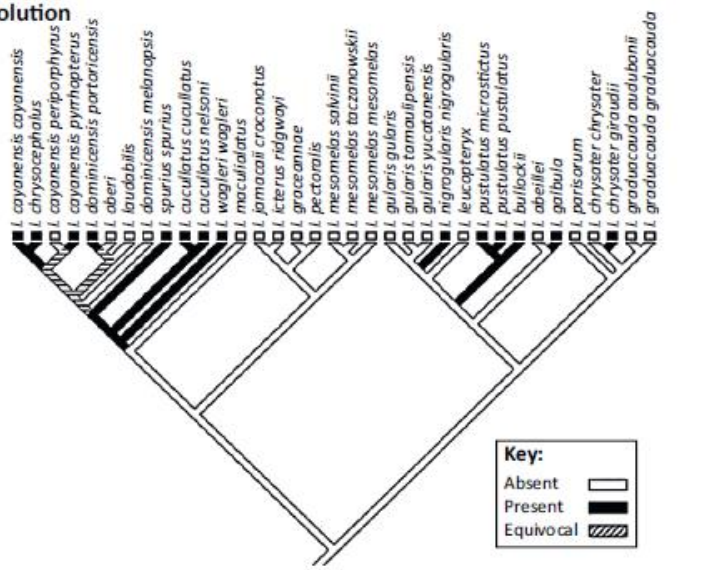


(B) Current utility

(C) Development



(D) Evolution

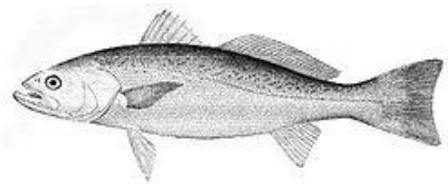


TRENDS in Ecology & Evolution

Applied research - Species recognition

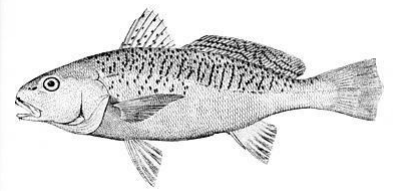
Rock Hind

bigeye scad



Weakfish

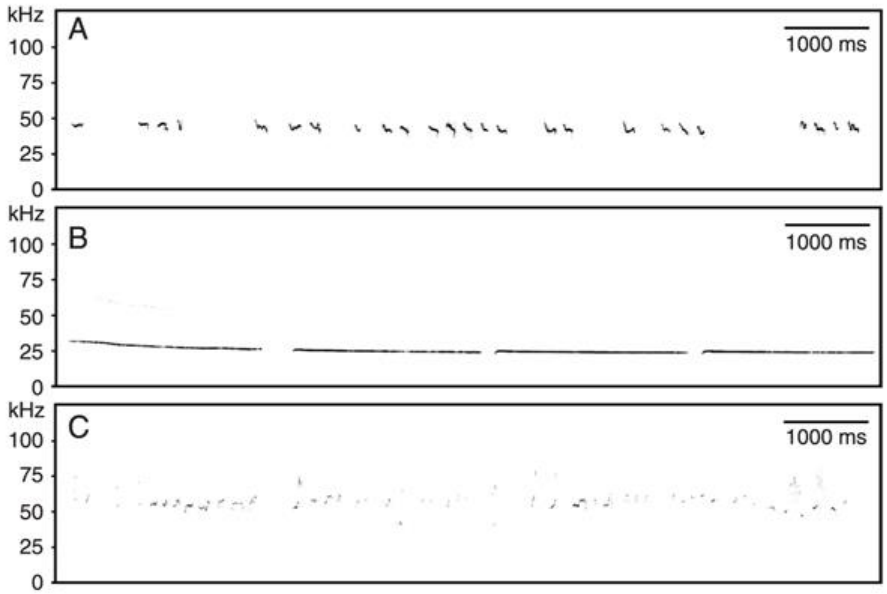
longhorn sculpin



Atlantic croaker

http://www.gso.uri.edu/fishsounds/CDindex.html

Applied research -
Quantifying the number of signals



27 signals

4 signals

>50 signals

Figure 1. Types of rat ultrasonic vocalizations (USV). *A*, Isolation-induced 40-kHz USV emitted by an 11-day-old male Wistar rat after separation from mother and littermates. *B*, Low-frequency 22-kHz USV emitted by a 3-month-old male Wistar rat during fear conditioning. *C*, High-frequency 50-kHz USV emitted by a 3-month-old male Wistar rat searching for conspecifics.

On the relationships between ultrasonic calling and anxiety-related behavior in rats. R.K.W. Schwarting and M. Wöhr. **Braz J Med Biol Res 012; 45: 337-348.**

General problems of acoustic processing

Easy recording -> large amount of data -> challenging to process

- Signal detection in noisy environment

PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Bat detective—Deep learning tools for bat acoustic signal detection

- Signal identification (individual/species recognition)

INTERFACE

royalsocietypublishing.org/journal/rsif

Automatic acoustic identification of individuals in multiple species: improving identification across recording conditions

RESEARCH ARTICLE

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

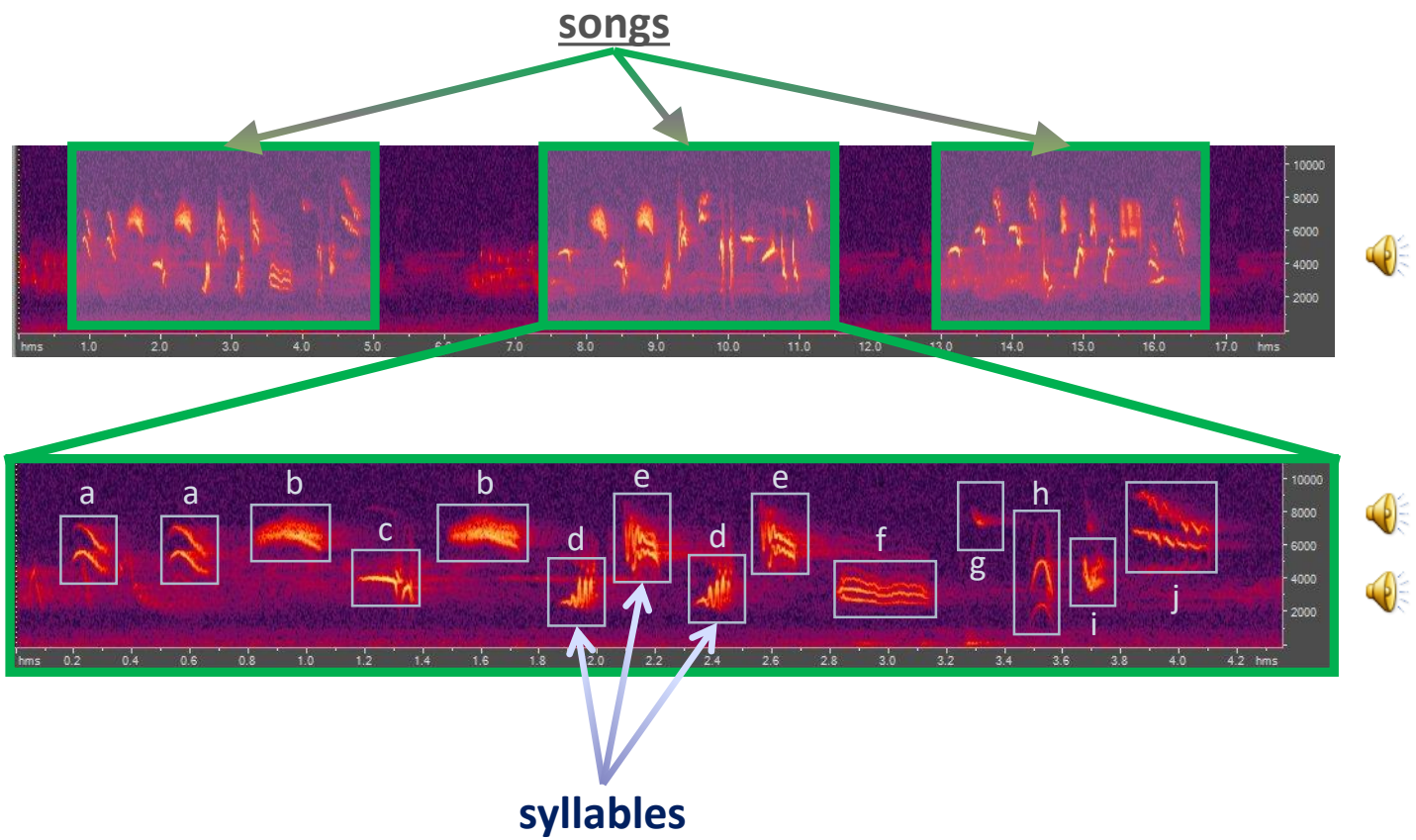
Automatic acoustic detection of birds through deep learning: The first Bird Audio Detection challenge

Object detection on spectrograms

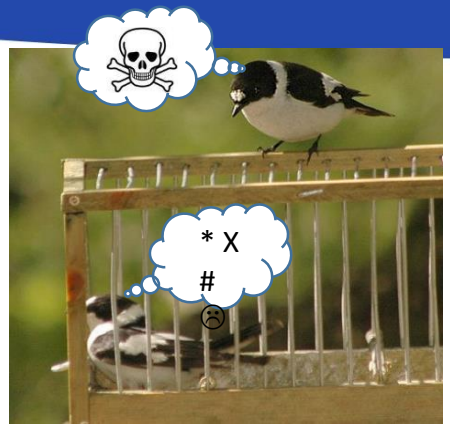
Deep Learning CNN

Study of collared flycatcher (*Ficedula albicollis*)

The song of collared flycatcher



Acoustic communication of collared flycatcher



Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

The relationship between syllable repertoire similarity and pairing success in a passerine bird species with complex song

Behav Ecol Sociobiol (2017) 71:154
DOI 10.1007/s00265-017-2379-0

ORIGINAL ARTICLE

Short- and long-term repeatability and pseudo-repeatability of bird song: sensitivity of signals to varying environments

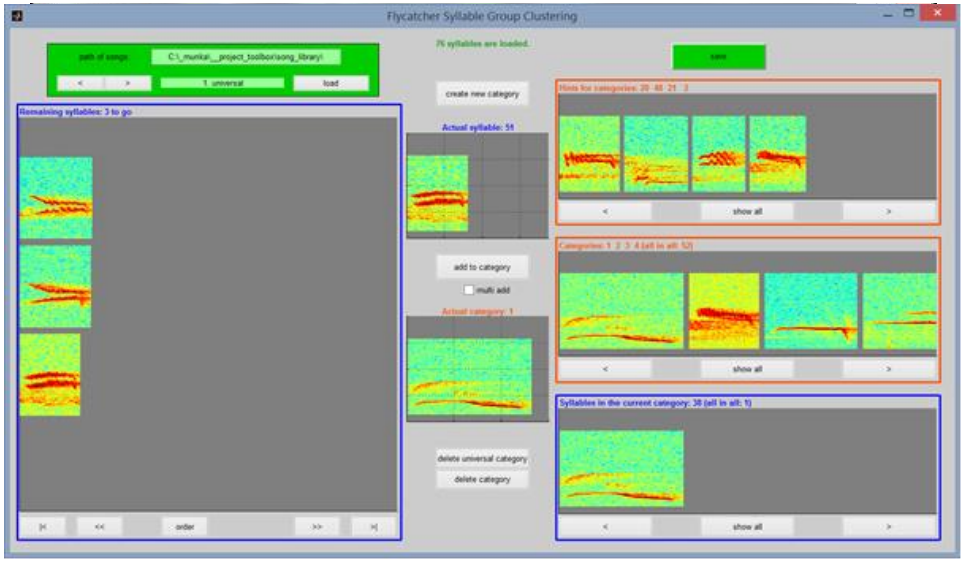
WILEY MOLECULAR ECOLOGY

ORIGINAL ARTICLE

MHC-mediated sexual selection on birdsong: Generic polymorphism, particular alleles and acoustic signals

Computer aided sound analysis with Ficedula

Matlab Toolbox (free, opensource script)



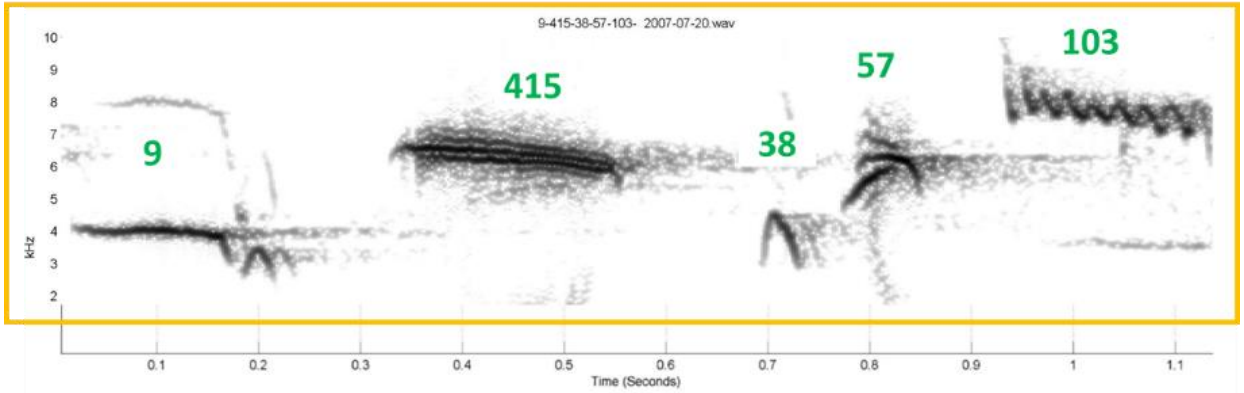
Sound recording

cutting

segmenting

clustering syllables within individual

clustering syllable types between individuals



The YOLO model

You Only Look Once: Unified, Real-Time Object Detection

Joseph Redmon*, Santosh Divvala*†, Ross Girshick¶, Ali Farhadi*
University of Washington*, Allen Institute for AI†, Facebook AI Research¶
http://pjreddie.com/yolo/

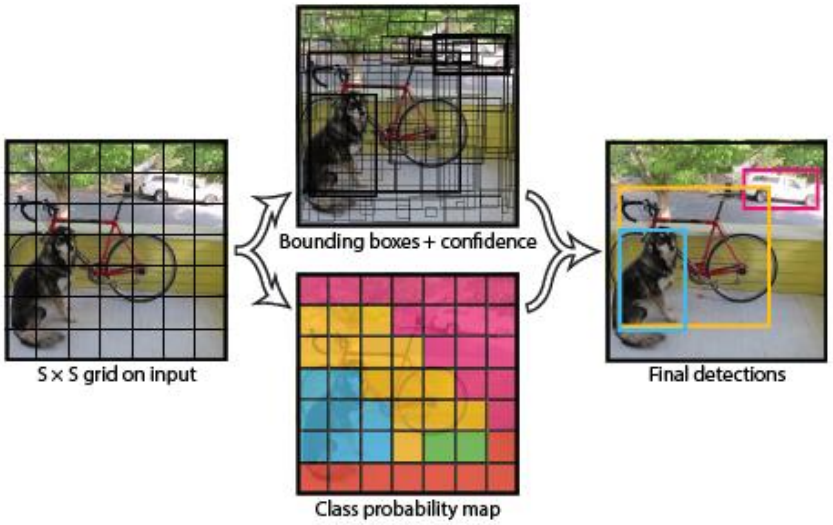


Figure 2: The Model. Our system models detection as a regression problem. It divides the image into an $S \times S$ grid and for each grid cell predicts B bounding boxes, confidence for those boxes, and C class probabilities. These predictions are encoded as an $S \times S \times (B * 5 + C)$ tensor.

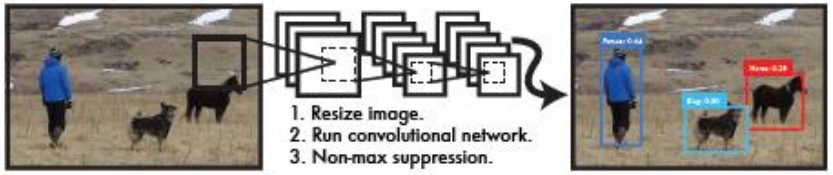
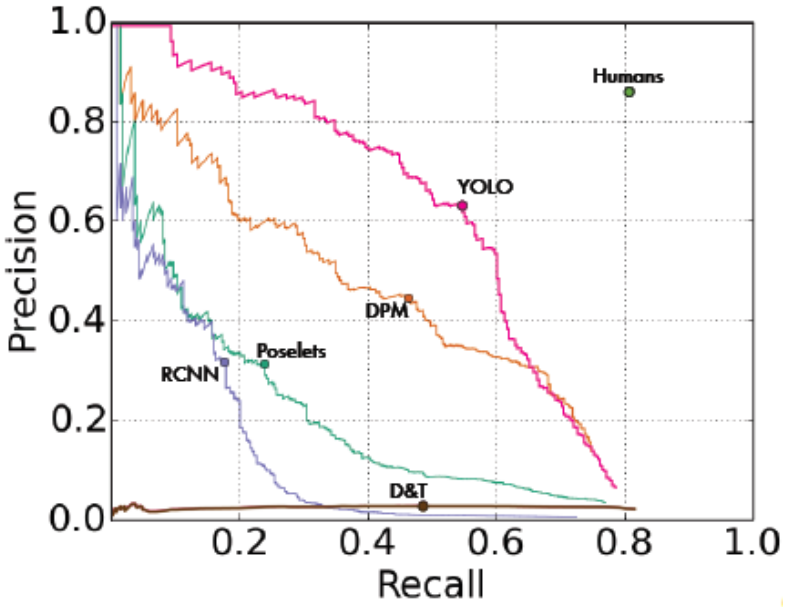


Figure 1: The YOLO Detection System. Processing images with YOLO is simple and straightforward. Our system (1) resizes the input image to 448 × 448, (2) runs a single convolutional network on the image, and (3) thresholds the resulting detections by the model's confidence.



(a) Picasso Dataset precision-recall curves.

The YOLO model

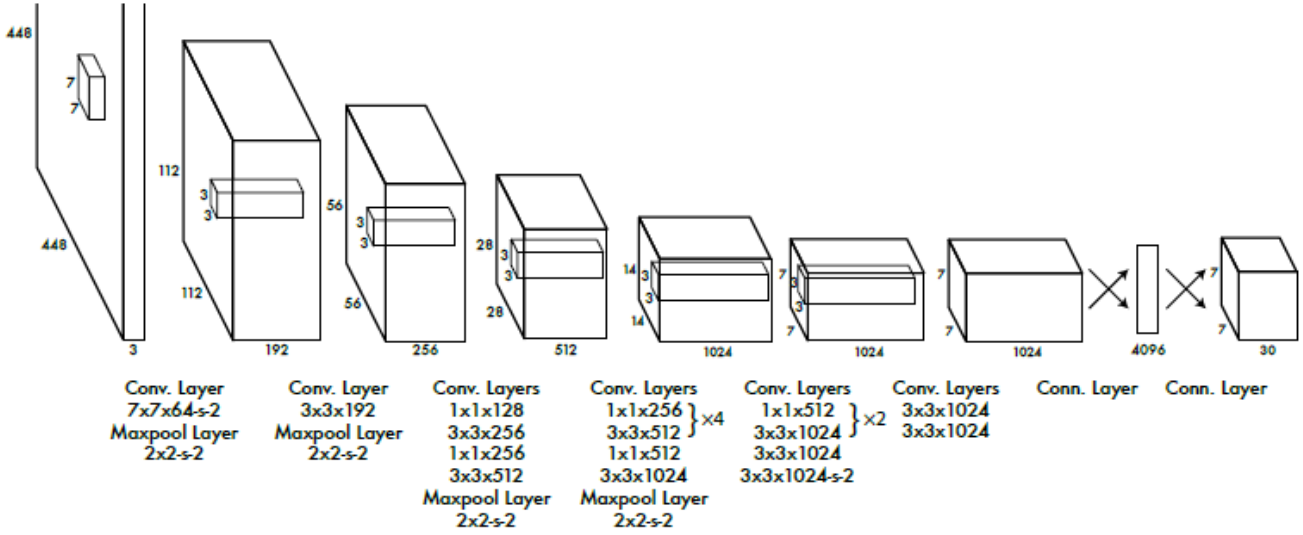
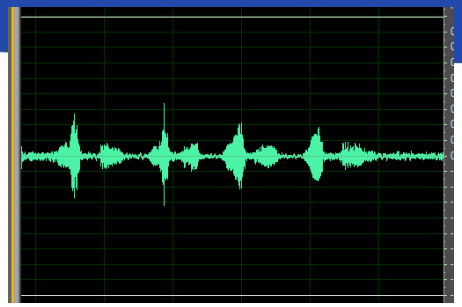


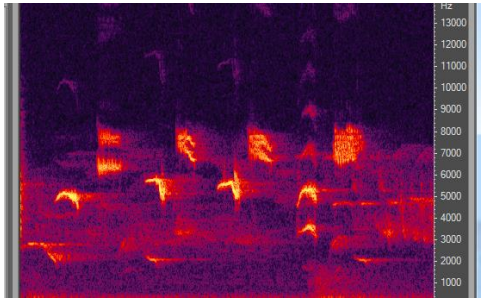
Figure 3: The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.

- Freeware, open source
- in C and Python
- OpenCV / CUDA
- CPU and GPU supported
- Changable input dimensions

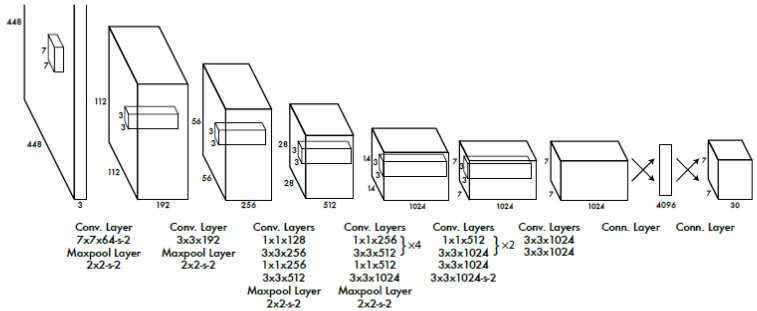
The YOLO model
on birdsong spectrogram



soundfile



spectrogram
(320 x 160 pixels)



YOLO

Detection of

1. **Song**
2. **Syllables**

xeno-canto

Sharing bird sounds from around the world

Search recordings... Search

About Explore Upload Sounds Forum Mysteries Articles











The teaching-testing database

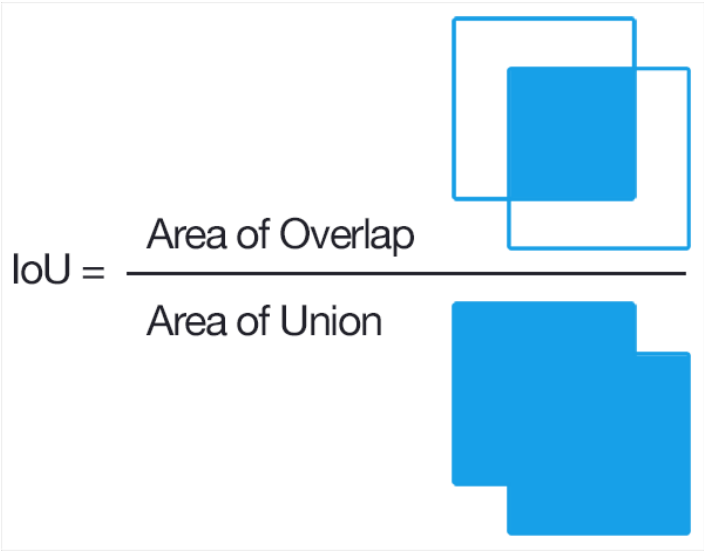
- Annotated collared flycatcher songs and syllables
- Diverse quality
  - From many years
  - Weather conditions
  - Early and late season
  - Different populations

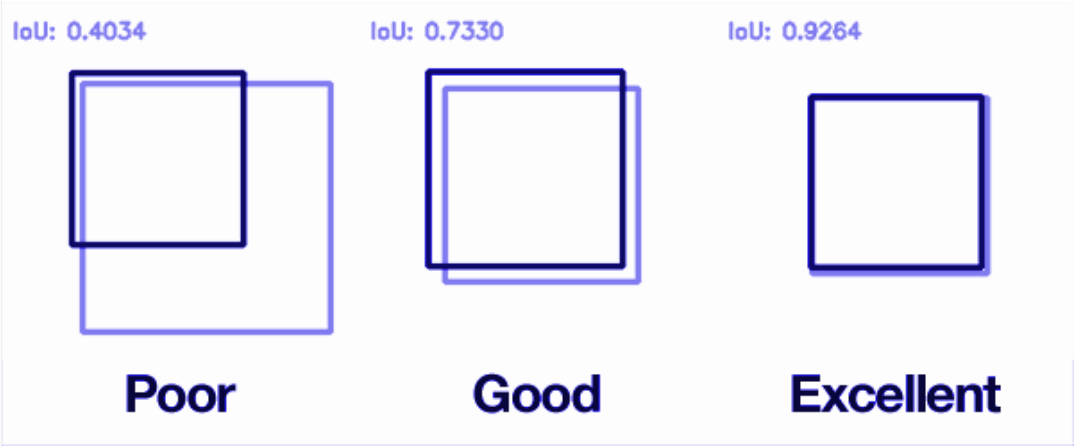
	Song database	Syllable database
Samples	6147	41229
Collared flycatcher	56 %	56%
Test %	10 %	10%

NVidia GeForce GTX 1080 Ti

Tracing:

- Loss function
- IOU (Intersection of Union):

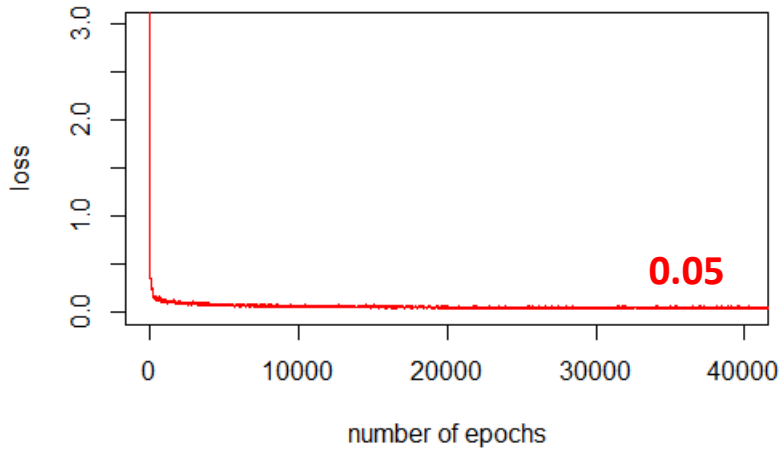




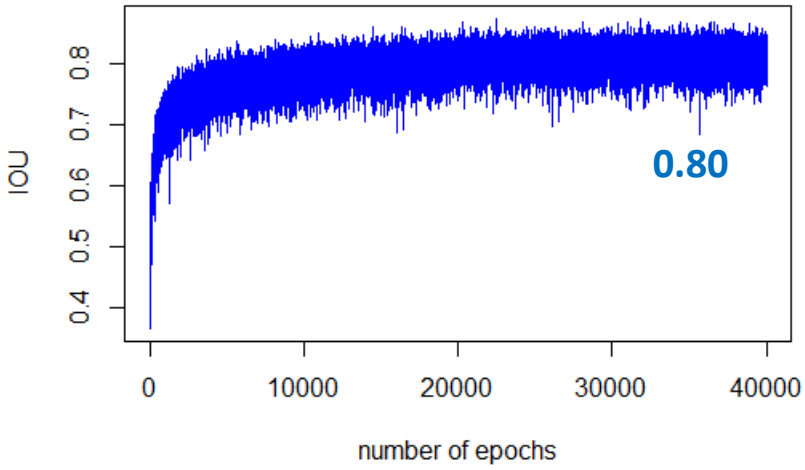
Results

Song detection

Average loss:

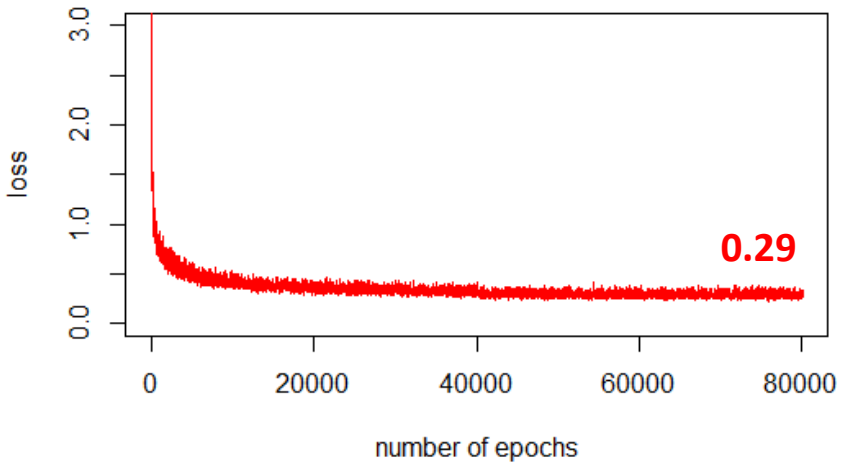


IOU:

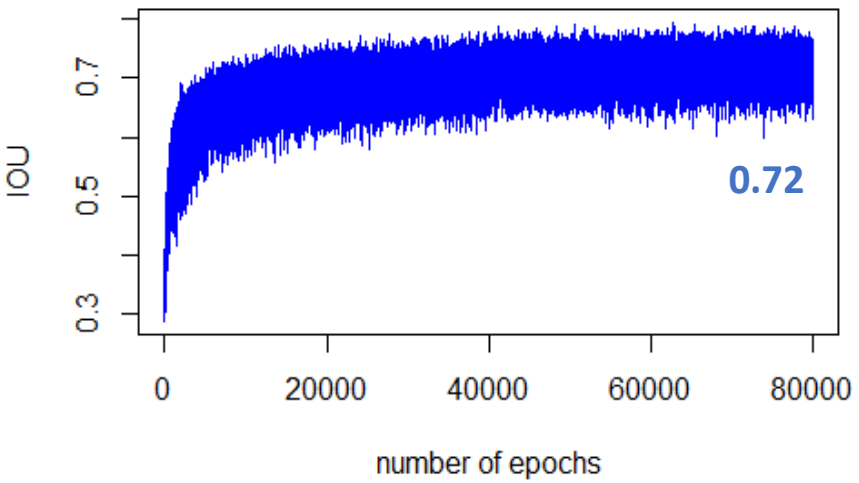


Syllable detection

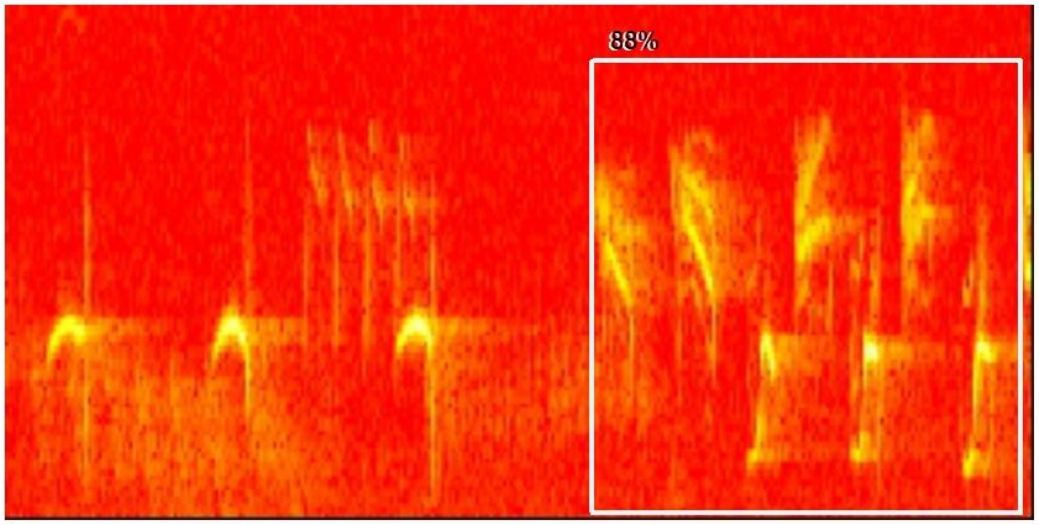
Average loss:

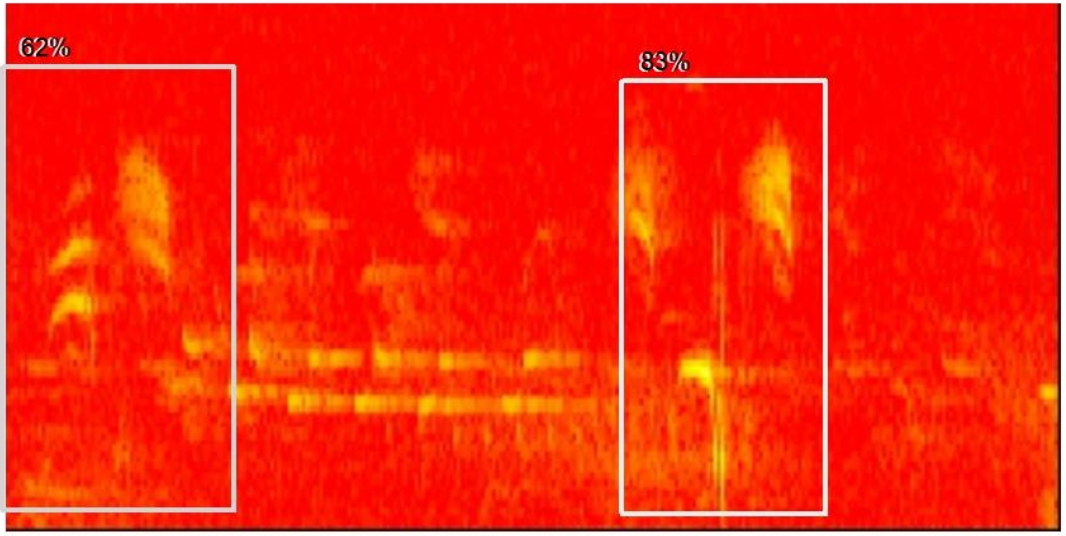


IOU:

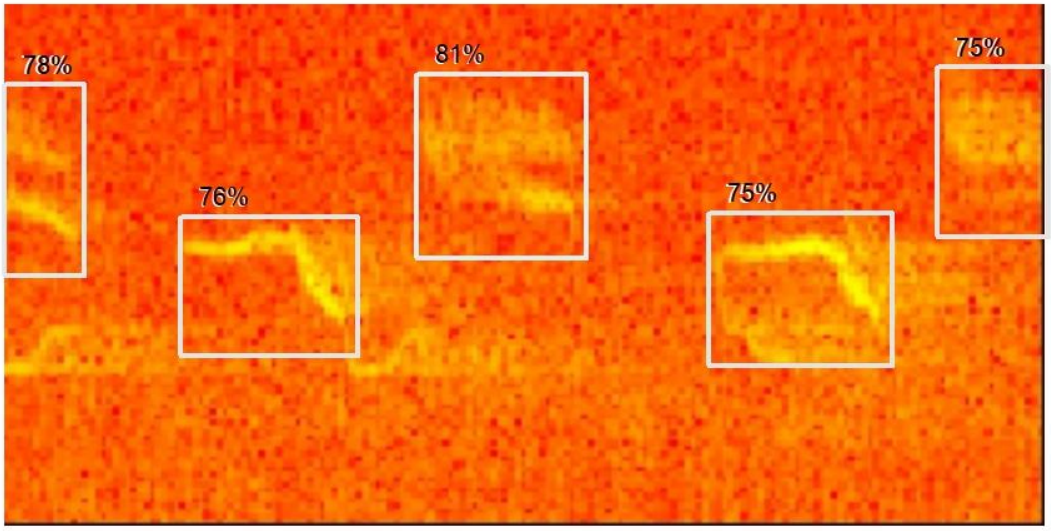


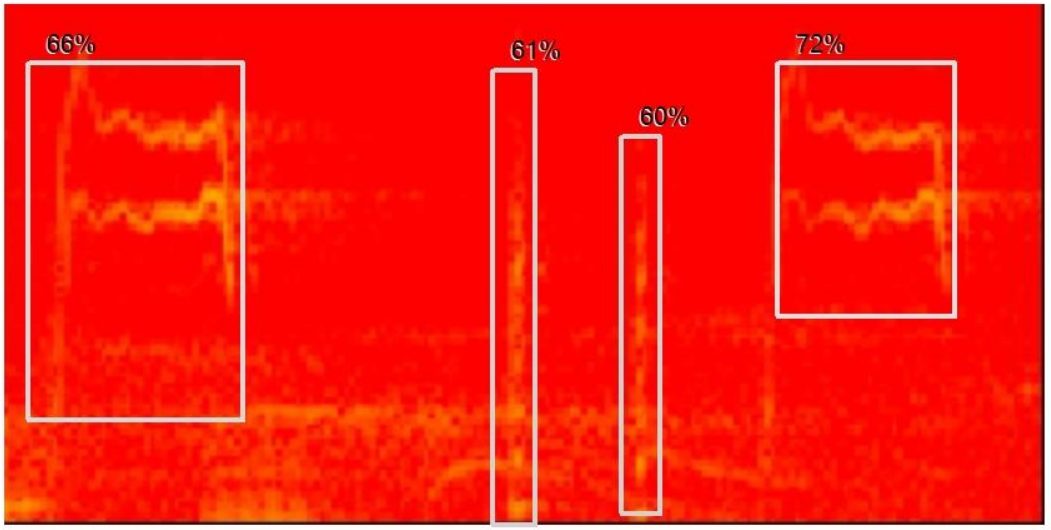
Examples for detection of songs





Examples for detection of syllables



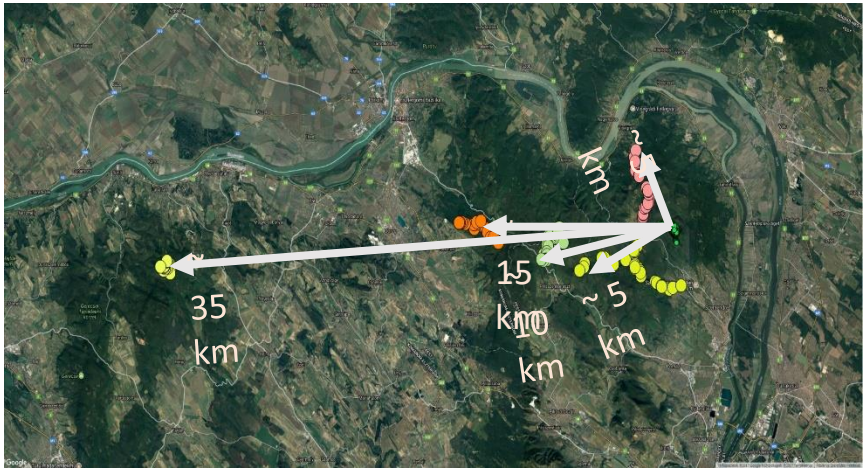


Application and further steps

- **First try on recordings from 2018:**

  Instead of 100 h manual segmentation, only 10 hours of checking!

- **Cultural evolution in time and space (>800 recordings):**



Acknowledgement

Éva Vaskuti

Miklós Laczi

László Zsolt Garamszegi

János Török

Gergely Nagy

György Blázi

Behavioural Ecology Group

WIGNER

NATIONAL RESEARCH, DEVELOPMENT AND INNOVATION OFFICE
(PD-117530, K-129215)

SZÉCHENYI 2020

EÖTVÖS LORÁND UNIVERSITY
Faculty of Science, Institute of Biology
Department of Systematic Zoology and Ecology
Pázmány Péter sétány 1/c.
H-1117 Budapest, Hungary

HUNGARIAN GOVERNMENT

European Union
European Social Fund

INVESTING IN YOUR FUTURE