

Towards Automatic Bird Detection: An Annotated and Segmented Acoustic Dataset of Seven *Picidae* Species

Ester Vidaña-Vila ¹, Joan Navarro ² and Rosa Ma Alsina-Pagès ^{1,*}

¹ GTM—Grup de recerca en Tecnologies Mèdia. La Salle—Universitat Ramon Llull. C/Quatre Camins, 30, 08022 Barcelona, Catalonia, Spain; esterv@salleurl.edu

² GRITS—Grup de Recerca en Internet Technologies & Storage. La Salle—Universitat Ramon Llull, C/Quatre Camins, 30, 08022 Barcelona, Catalonia, Spain; jnavarro@salleurl.edu

* Correspondence: ralsina@salleurl.edu; Tel.: +34-932-902-455

Academic Editor: Jamal Jokar Arsanjani

Received: 20 April 2017; Accepted: 13 May 2017; Published: 16 May 2017

Abstract: Analysing behavioural patterns of bird species in a certain region enables researchers to recognize forthcoming changes in environment, ecology, and population. Ornithologists spend many hours observing and recording birds in their natural habitat to compare different audio samples and extract valuable insights. This manual process is typically undertaken by highly-experienced birders that identify every species and its associated type of sound. In recent years, some public repositories hosting labelled acoustic samples from different bird species have emerged, which has resulted in appealing datasets that computer scientists can use to test the accuracy of their machine learning algorithms and assist ornithologists in the time-consuming process of analyzing audio data. Current limitations in the performance of these algorithms come from the fact that the acoustic samples of these datasets combine fragments with only environmental noise and fragments with the bird sound (i.e., the computer confuses environmental sound with the bird sound). Therefore, the purpose of this paper is to release a dataset lasting more than 4984 s that contains differentiated samples of (1) bird sounds and (2) environmental sounds. This data descriptor releases the processed audio samples—originally obtained from the Xeno-Canto repository—from the known seven families of the *Picidae* species inhabiting the Iberian Peninsula that are good indicators of the habitat quality and have significant value from the environment conservation point of view.

Data Set: <https://doi.org/10.5281/zenodo.574438>

Data Set License: CC BY-SA 4.0

Keywords: birdsong; acoustic bird recognition; audio sample; dataset; birdsong dataset

1. Introduction

Coordination and organization in any kingdom of life require some kind of communication. The most widespread ways to communicate between animal species and their related species rely on visible (i.e., changing the body position) and/or vocal (i.e., sequences of sounds) expression patterns [1]. Ranging from insects to mammals, including amphibians and songbirds, they all generate complex acoustic sequences [2] associated to different life events (e.g., mate attraction and selection, predator alerts, or food availability). In recent years, animal communication [1] has become a hot research topic since it enables researchers to obtain valuable information about the species population, their evolution, and the environments in which they typically inhabit [3]. This is especially interesting in

the domain of bird species since they use a complex acoustic communication—they distinguish songs from calls and include non-vocal sounds such as the drumming of woodpeckers and the “winnowing” of snipes’ wings in display flight [4]—which is known to be rich and meaningful [3,5]. Therefore, this work revolves around the analysis of this kind of acoustic sounds produced by birds.

One of the most time-consuming tasks that ornithologists and birders typically conduct is to record, annotate, and differentiate species in their natural habitat. Usually, this process requires the intervention of a human expert. That is, a specialist from the specific field of the species to be analyzed is asked to identify and accurately classify any individuals that live in a bio-diverse environment. Alternatively, some researchers from other areas have attempted to address this classification challenge by using signal processing techniques [6] or taking advantage of the latest advances in computer science and artificial intelligence to build learning classifier systems able to automatically identify each animal [7,8] with reasonable success. The accuracy of these automatic systems typically relies on (1) the acoustic properties of the animal sound; (2) the recording quality of the dataset; (3) the features selection extracted from each audio fragment; (4) the number of collected samples; and (5) the classification algorithm. Overall, the success of these computer-assisted strategies relies on the dataset used to train and test the classifier. For instance, samples recorded with background noise (Figure 1a), may negatively affect the accuracy of automatic classifiers.

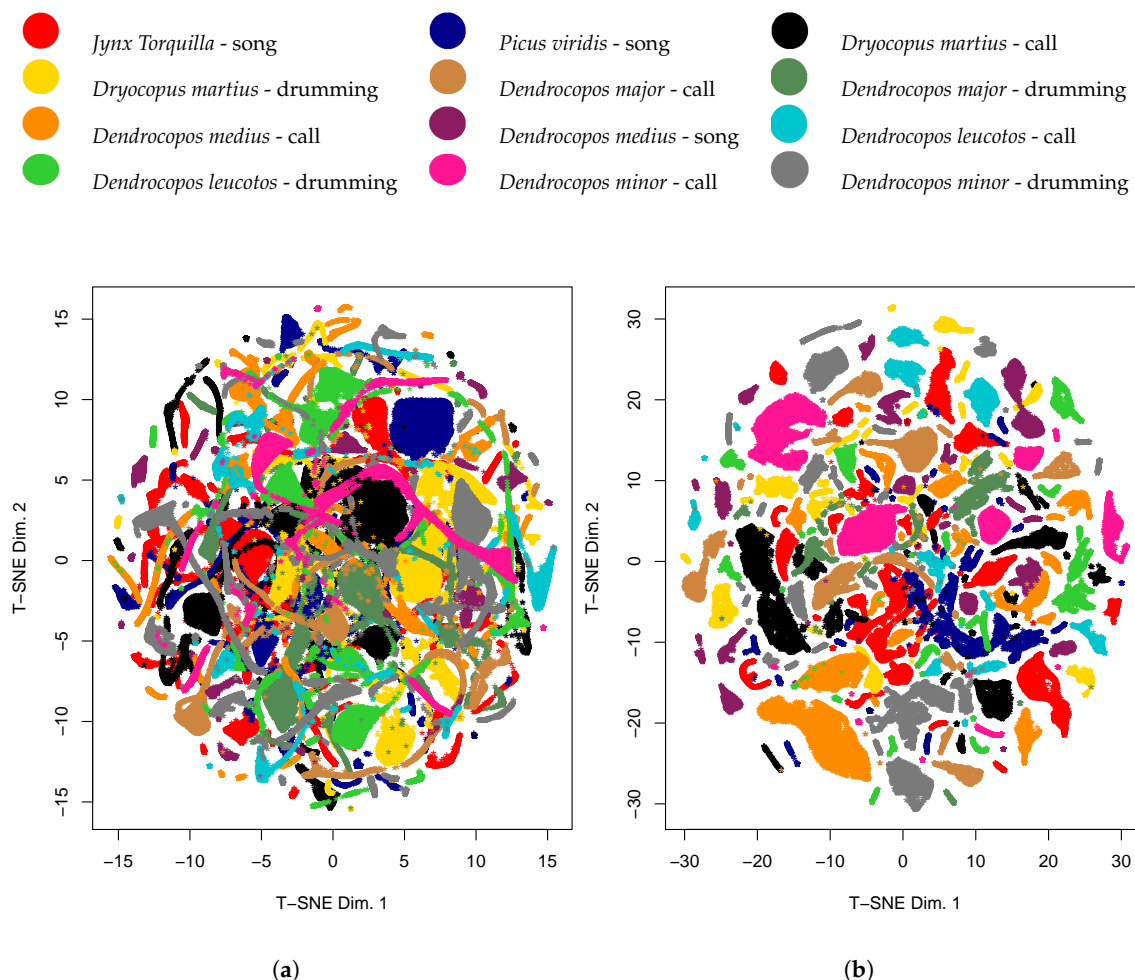


Figure 1. Effects of the background noise on the spatial distribution of a birdsong dataset. The selected features to model each dataset result from applying a 30 ms. sliding Hamming window to the spectral response of 12 types of sounds corresponding to seven *Picidae* bird species and computing the 13 Mel Frequency Cepstral Coefficients [9], as typically done in similar works [8]. (a) Spatial distribution of a dataset with background noise; (b) Spatial distribution of a dataset with no background noise.

Figure 1 emphasises this fact by depicting two t-Distributed Stochastic Neighbor Embedding (t-SNE) [10] plots (i.e., a space transformation that projects points from different classes in a bi-dimensional plane) of the same dataset. One with background noise (Figure 1a) and the other with the background noise removed (Figure 1b), which gives a visual perception of the class separability and the complexity of the classification problem. It can be seen that classes on the left plot are very overlapped while this effect is mitigated on the right plot. This means that if the dataset corresponding to the left plot were delivered to a learning classifier system, its performance would probably be very limited due to the fact that the system may confuse environmental noise with the actual bird sound. This stresses the need for a clean and reliable bird sound repository that researchers can safely use to train and tune machine learning systems.

So far, there exist several excellent databases containing general purpose recordings of birds and animal sounds. On the one hand, there are projects such as Xeno-Canto [11,12], Cornell's Macaulay Library [13], the Western Soundscape Archive [14] and Avibase [15], which usually share general sound recordings of animals and soundscapes of their habitat. On the other hand, there are repositories such as the British Library [16], the Berlin Museum für Naturkunde [17] or the Australian National Collection Sound Archive [18] that include metadata about the recordings (location, ecology, etc). However, these generic purpose databases suffer from two issues when used for automatic identification and classification of a closed set of bird species in their natural habitat: (1) label reliability and (2) environmental noise (also referred to as silence). First, these audio samples usually lack accurate and reliable annotations about the song types, which requires an expert to supervise the labelling process. Second, almost none of these repositories separates the silences (i.e., periods of time where the bird under interest is not emitting any sound but there might be environmental noise such as waterfalls or other species) between periods of bird sound, which also requires a further annotation in order to boost the classifier performance (see Figure 1). Additionally, the accuracy of automatic birdsong recognition and classification systems could be further improved if applied to taxonomic bird groups with songs of limited complexity, especially when these groups are good indicators of the habitat quality or its conservation status. For instance, this would be the case of woodpeckers (belonging to the *Picidae* family), which have simpler sound emissions and have significant value from the conservation point of view [19]—although some of their species are included in threat categories [20], they are still especially relevant for their role as indicators of quality and maturity of forest environments [19,21].

Therefore, the purpose of this paper is to present an annotated dataset of *Picidae* bird sounds lasting 4984 s in which bird sounds have been separated from the environmental noise. This dataset includes a collection of 1669 audio samples—derived from 161 audio files collected from the Xeno-Canto repository [11]—from the known seven families of the *Picidae* species inhabiting the Iberian Peninsula emitting up to three different sounds: call, drumming or song. Although the number of families included in this dataset is considerably smaller than other general purpose collections [22], the species herein incorporated have been specifically selected and neatly chosen according to their relevance for the habitat conservation and biodiversity quality domain [19,21]. Hence, researchers using this dataset might be able to train an automatic Learning Classifier System to (1) detect whether there is any *Picidae* bird in a given sound sample collected from a natural environment; (2) identify the type of *Picidae* bird; and (3) recognize the type of song. Automatic bird classification has emerged as a hot research topic in recent years. For instance, modern classification techniques based on deep neural networks and data augmentation have enabled researchers to reach an interesting accuracy (close to 60%) from a dataset of 33,203 audio recordings of 999 species collected in South America at the LifeCLEF Bird Task [22]. In comparison to the LifeCLEF Bird Task dataset (it has an average of 24 samples for each species) that covers a plethora of bird species, our proposal can be best seen as a reduced approach (it has an average of 96 samples for each species) targeted at analyzing a specific set of bird species with great biological interest [19–21].

The remainder of this paper is organized as follows. Section 2 describes the proposed dataset in terms of geographical distribution, number of files, duration, and spectral evolution over time. Section 3 details the segmentation and annotation process followed to obtain the proposed dataset. Section 4 specifies the files structure of the dataset and the instructions to download it. In Section 5, authors outline the conclusions of this work and present future research directions. Additionally, all the source audio files used from the Xeno-Canto project together with the names of the recorders are provided in the Acknowledgements section.

2. Dataset Description

The proposed *Picidae* annotated database has been built from 161 recordings obtained from the Xeno-Canto Project [11,12] (the detail and the names of the recorders for each original recording are provided in the Acknowledgements section of this paper). The geographic distribution of the original recordings, the composition of the resulting dataset, and the spectral analysis of the birdsongs are detailed in what follows.

Figure 2 depicts a map containing the location of every original audio from the Xeno-Canto [11] repository where each bird color represents a different *Picidae* species and a different type of sound (call, drumming or song). In those situations where different types of sound for the same species are available (e.g., *Dendrocopos leucotos*-call, *Dendrocopos leucotos*-drumming), each sound has been drawn with a different color to further improve the detail of the obtained map. As a result, this map provides a general overview of the geographical distribution of the collected samples that compose the dataset presented in this paper. It can be seen that although most of the recordings were taken in central Europe, some of them were registered in Africa and Asia.

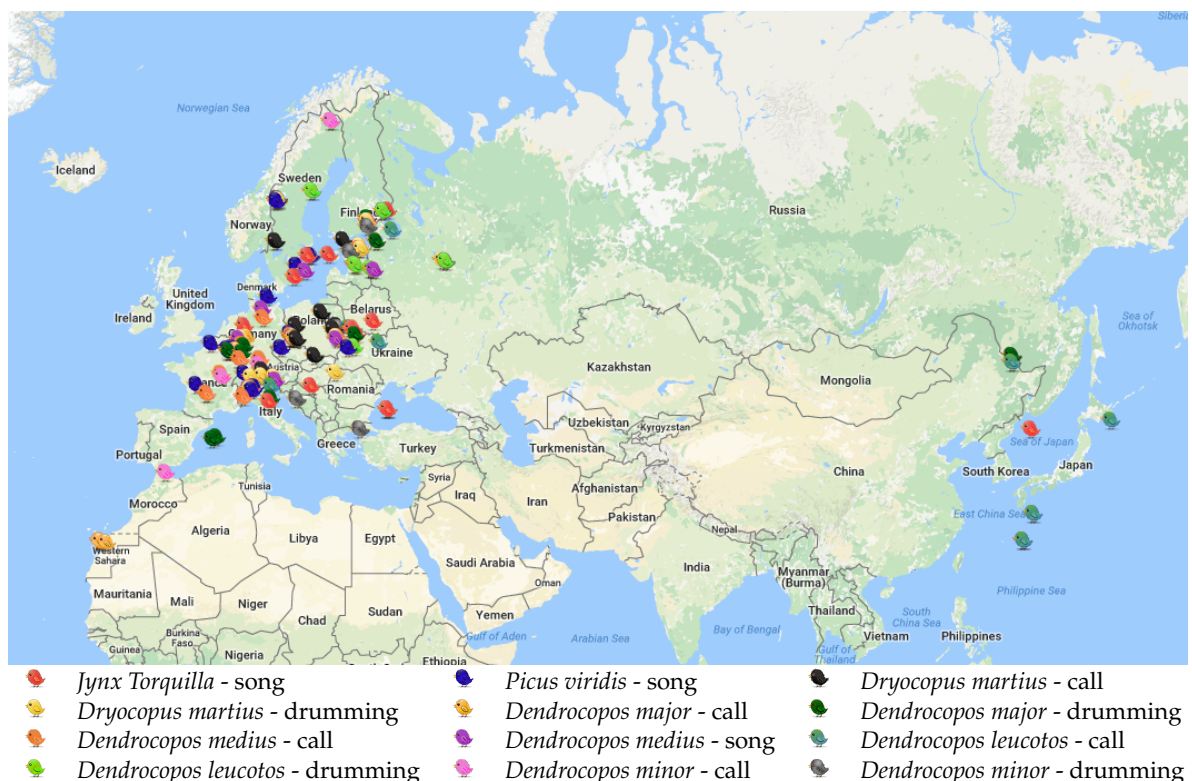


Figure 2. Location of the audio recordings obtained from the XenoCanto [11] repository.

After the acoustic processing herein presented, these 161 recordings resulted in 1669 labeled files. As a result, the corpus contains three different sound types (call, drumming, and song) from the following *Picidae* species: *Jynx Torquilla*, *Picus viridis*, *Dryocopus martius*, *Dendrocopos major*, *Dendrocopos medius*, *Dendrocopos leucotos* and *Dendrocopos minor*. The list of the species (i.e., species name), the number of acoustic samples (i.e., file count), the overall length of all samples (i.e., total length), and the duration distribution of all the files that compose the proposed dataset are shown in Table 1. Note that the selected colors are consistent with the ones used in Figure 2. It can be seen that the longest individual sound is made by *Dendrocopos medius*, and corresponds to its song. Drumming in all species is usually a short and uniform sound, with a length ranging from 1 to 3 s. The songs of the species are also uniform in terms of duration, ranging from 1 to 8 s with median values around 3 s. Also, call signals are very variable in terms of length: on the one hand, *Dendrocopos minor* and *Dendrocopos medius* show high variability, with long samples, despite the median being around 2 s. On the other hand, *Dendrocopos leucotos*, *Dryocopus martius* and *Dendrocopos major* show calls from maximum 3 s.

Table 1. Files obtained after the segmentation and annotation processes. Proposed dataset composition.

Specie Name	File Count	Total Length (s)	Duration Distribution (s)
<i>Dendrocopos minor</i> -drumming	127	228	
<i>Dendrocopos minor</i> -call	105	209	
<i>Dendrocopos leucotos</i> -drumming	50	119	
<i>Dendrocopos leucotos</i> -call	146	128	
<i>Dendrocopos medius</i> -song	42	137	
<i>Dendrocopos medius</i> -call	124	299	
<i>Dendrocopos major</i> -drumming	62	88	
<i>Dendrocopos major</i> -call	146	183	
<i>Dryocopus martius</i> -drumming	53	133	
<i>Dryocopus martius</i> -call	140	198	
<i>Picus viridis</i> -song	49	143	
<i>Jynx Torquilla</i> -song	102	324	
Silence	523	2795	
Total	1669	4984	1 5.5 10

Additionally, a 1 s length sample spectrogram for every different type of sound contained in the dataset is shown in Figure 3. The first row shows the spectrograms of the song from three species: the *Picus viridis*, the *Jynx torquilla* and the *Dendrocopos medius*. Although the shape of the spectro-temporal representation of the signal is similar in the three cases with harmonics, the *Picus viridis* has five repetitions, the *Jynx torquilla* has four and the *Dendrocopos medius* has only two repetitions. The spectro-temporal representation of the call from five different species can be observed in Figure 3d–h. The longest call is shown in Figure 3h from the *Dryocopus martius* while the other four are considerably shorter. Figure 3e depicts a single short call (*Dendrocopos leucotos*) while Figure 3d (*Dendrocopos medius*), Figure 3f (*Dendrocopos minor*) and Figure 3g (*Dendrocopos major*) show three to four repetitions of a smaller period call. Finally, the spectro-temporal representation of the drumming—despite it being non-vocal, it is also considered a birdsong—from four species is shown in Figure 3i–l. The drumming shows several wide-spectrum temporal repetitions for all the species. The following section elaborates on the process to segment and annotate the proposed dataset.

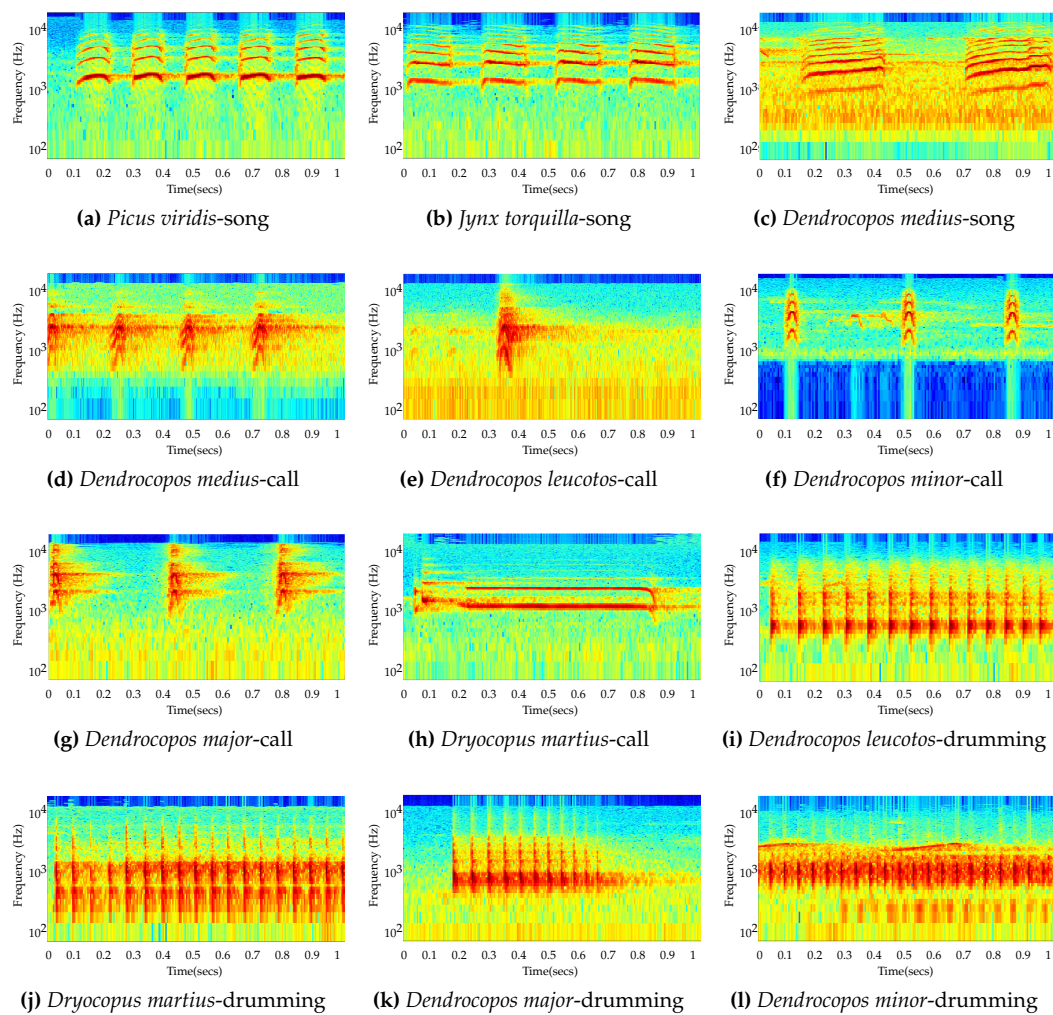


Figure 3. Spectrograms of the seven *Picidae* species with the available sounds (song, call and drumming).

3. Segmentation and Annotation Process

Figure 4 illustrates the segmentation and annotation process that has been conducted to separate the background noise from the sound of interest (i.e., bird sounds) for each of the 161 audio samples downloaded from the Xeno-Canto repository. Segmenting and annotating a dataset resulted to be successful in similar works [22]. For the sake of this paper, this process has been decomposed into six steps detailed in the following.

- Step I. Original audio sample acquisition.** It consists of downloading the audio samples from the Xeno-Canto repository [11,12] and running an initial manual pre-filtering. That is, listening to the downloaded samples and ensuring that they correspond to the species to which they are labeled. Low quality and too noisy samples are discarded.
- Step II. Energy threshold calculation.** It consists of establishing an energy threshold to distinguish between a sample containing sound and a sample containing silence. It is computed by squaring every sample value of the whole original audio and averaging the result (i.e., average power of the audio recording).
- Step III. Sample windowing.** It consists of applying a 500 ms. sliding window with no overlap to the original audio sample (i.e., splitting the file in 500 ms. segments). We have found that 500 ms. is a convenient size because it captures enough acoustic energy of the bird sound to discern it from the background noise of the acoustic sample. A larger window size may increase the chance of capturing different bird songs in the same

sample. Accordingly, a smaller window size may prevent the system from distinguishing background noise from the bird song.

- Step IV. Local energy calculation.** It consists of calculating the average power of the 500 ms. audio segment.
- Step V. Energy comparison.** It consists of deciding whether the 500 ms. audio segment is either sound or silence. If the result from Step IV is higher than the threshold established in Step III, the audio fragment is pre-labeled as a sample of interest (the bird is making some sound); otherwise, the fragment is labeled as background noise (i.e., silence).
- Step VI. Pre-labeling.** It consists of comparing the label obtained at Step V with the previous labels. If the label is the same (e.g., all the 500 ms. fragments so far have been labeled as silence), the system goes to Step III again with the following segment. If the label is different (e.g., the previous window was labeled as silence and the actual window was labeled as sound of interest), a new audio with all the previous same-label windows is generated and stored before going again to Step III with the next segment of the original audio sample. In the latter case, the segment will be included in the next sample collection. It is worth mentioning that acoustic samples—even from different 500-ms adjacent windows—that still contain subsequent repetitions from the same sound will be tied together as a single audio file at this step. This is done to capture the information of sound repetitions, which might be of great interest when identifying the bird species and its type of sound.
- Step VII. Fine tuning and final annotation.** It consists of conducting a manual review (i.e., removing false positives) and adjustment (i.e., moving background noises from actual sound samples to the Silence class) on the pre-labeled samples. For instance, there might be some audio samples that might introduce ambiguous information to the dataset and, thus, need to be deleted (e.g., a *Jynx Torquilla* and a *Picus viridis* singing at the same time and labeled as a *Jynx Torquilla* only). It is worth noting that the background noise (i.e., silence) segments might contain birdsongs from non-interesting bird species (i.e., birds that do not belong to any of the *Picidae* species that are being studied in this paper); the rationale behind this decision is to make sure that an automatic classifier will consider all the noises (even if they are birdsongs) that are not generated by the bird of interest as silence.

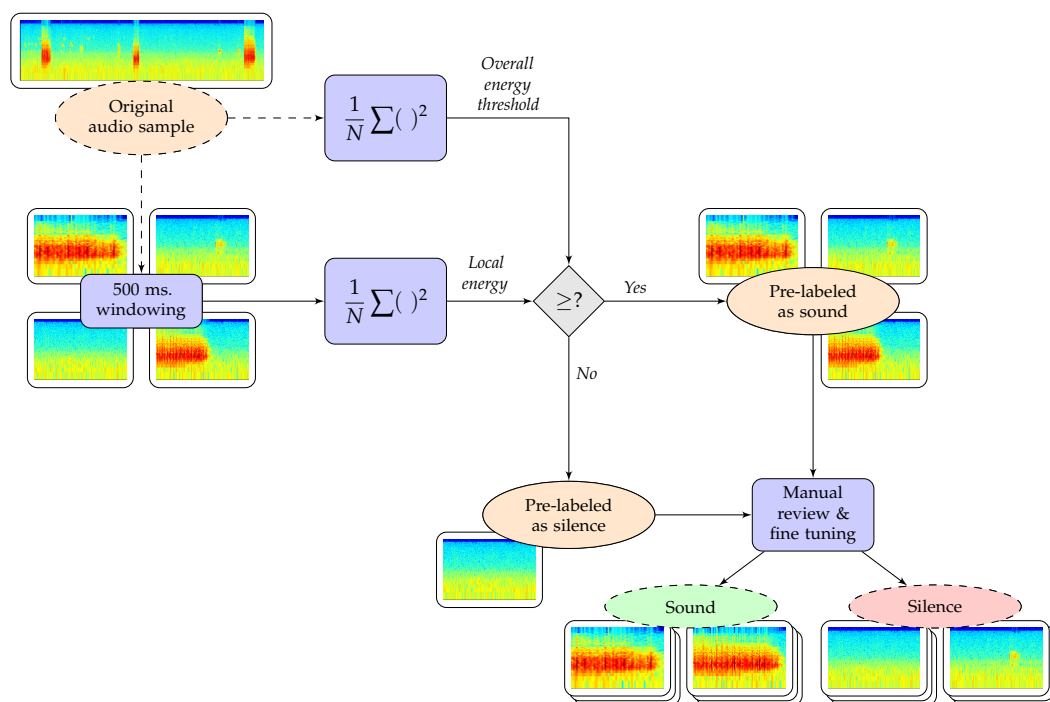


Figure 4. Annotation and segmentation process of the dataset generation.

Overall, Figure 5 shows the spectrogram and temporal waveform from an original audio sample. The audio belongs to a sample where an exemplar of the species *Dryocopus martius* drums. The parts where the bird is drumming have been highlighted with a red bounding box. After running the proposed algorithm with this sample, three sub audios of sound of interest labeled as *Dryocopus martius* drumming and four sub audios labeled as silence would be generated.

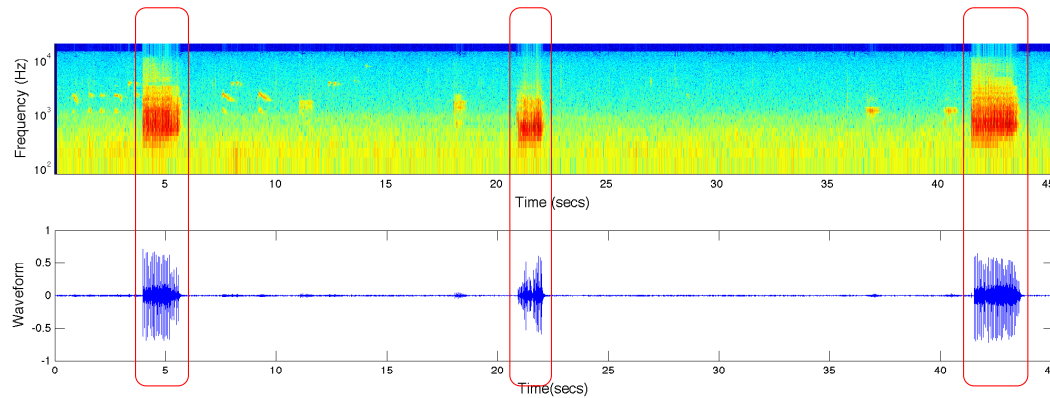


Figure 5. Spectrogram and waveform from one audio of the Xeno-Canto repository.

4. Materials

All the materials are contained in a *zip* file that can be downloaded from <https://doi.org/10.5281/zenodo.574438>. Audio data are organized in thirteen folders: twelve containing the labeled audios of each birdsong type from every *Picidae* species listed in Table 1 and one folder containing the background sound samples. Every folder is named according to the following pattern: “*Picidae* species name—the type of sound” (e.g., *JynxTorquilla*-song). Analogously, all the audio files inside every folder have been named using the same strategy. More concretely, the structure of the audio file name is: Number of register of the audio at the Xeno-Canto repository—Name of the *Picidae* species—Name of the type of sound—Number of sub audio referenced to the original Xeno-Canto file (e.g., XC233286-JynxTorquilla-song-6.wav). For instance, if four sounds of interest were extracted from the original Xeno-Canto audio sample XCaaa, which was a recording from a *bbb* *Picidae* species doing a *ccc* type of sound, the generated audios will be the following: XCaaa-bbb-ccc-1.wav, XCaaa-bbb-ccc-2.wav, XCaaa-bbb-ccc-3.wav, and XCaaa-bbb-ccc-4.wav. Please, note that all the audio files have been converted to *.wav*. Finally, the *zip* file contains a *Read Me* document detailing the organization of the main folder.

5. Conclusions

The research conducted nowadays in automatic bird classification requires an extensive number of precisely labelled files of birdsongs recorded from real-life environments. Additionally, information about where the recordings were made, the type of sound and the animal species is very valuable. Unfortunately, annotating the pseudo-periodic songs, drumming or calls from each bird species, together with segmenting the silence recorded between samples is a delicate and challenging task. In this regard, we have constructed a segmented and annotated database of the seven *Picidae* bird species known to inhabit the Iberian Peninsula that enables such recordings to be shared for the community. We have also included the data structure to help researchers with the management and understanding of the dataset. Future work will be focused on increasing the information about all the files of the dataset, (e.g., the spectrogram for each file or embed the localization visualization).

Acknowledgments: The authors would like to thank project Xeno-Canto for their work gathering and publishing acoustic birdsong recordings. We would like to thank specially the recorders of all the birdsong data, without whom this work would not have been possible. We list all the recorders together with the Xeno-Canto code of the recording we used to generate the annotated database. Albert Lastukhin (XC267333), Alexander Kurthy (XC311605), Antero Lindholm (XC247633), Beatrix Saadi-Varchmin (XC356233), Bram Piot (XC170646), brickegickel (XC313131, XC356127), Buhl Johannes (XC177871), Concord Lexington (XC318086), Danuta Peplowska-Marcza (XC302722), Dare Sere (XC337994), Dawid Jablonski (XC176278), Dmitry Kulakov (XC309571, XC309576, XC309580), Dmitry Yakubovich (XC234609), Eetu Paljakka (XC207080, XC308855, XC308859, XC324852), Elias A. Ryberg (XC313173, XC338390), Francesco Sottile (XC307521, XC311566), Gunnar Fernqvist (XC221271), Hannu Jannes (XC214195), Hans Matheve (XC354870, XC354873), Jack Berteau (XC156677, XC156679), Jarek Matusiak (XC130202, XC153024, XC210380, XC234512), Jelmer Poelstra (XC315966), Jens Kirkeby (XC322593), Jerome Fischer (XC169322, XC187153, XC303432, XC308986), Joost van Bruggen (XC293003, XC321386), Jordi Calvet (XC306722, XC348419), José Carlos Sires (XC343460), Julien Rochefort (XC147107, XC304818), Justin Jansen (XC165371), Krzysztof Deonizak (XC310278, XC314203, XC314369, XC314609, XC314610, XC314985), Lars Adler Krogh (XC215636), Lars Buckx (XC179826), Lars Lachmann (XC127763, XC252476, XC331306, XC331317), Lauri Hallikainen (XC233286, XC234477, XC234544), Manuel Grosselet (XC298607), Marc Anderson (XC310360), Marco Dragonetti—www.birdsongs.it—(XC331113, XC331114, XC331115, XC331116), Martin Vlk Mrnous (XC355912), Mikael Litsgard (XC237259, XC239371), Miklos Heincz (XC313584), Niels Van Doninck (XC351824), Niels Van Doninck (XC351825), Nikolay Sariev (XC239446), Nils Agster (XC288966), Pascal Christe (XC302364), Patrik Aberg (XC26678, XC293037, XC293038, XC310710, XC343373, XC343374, XC349394, XC349395), PE Svahn (XC212569), Pepe Lehtikainen (XC277517), Peter Mazuryk (XC355288), Ruud van Beusekom (XC46219), Sonnenburg (XC166681, XC234155, XC313701, XC319646, XC342794, XC355052, XC355459, XC355460, XC356033, XC356034), Stanislas Wroza (XC354361), Stein O. Nilsen (XC278527, XC278528), Szymon Plawewski (XC335265, XC335267), Terje Kolaas (XC236456, XC236458, XC236460, XC236833, XC237245, XC302820, XC324357, XC324358), Tero Linjama (XC343654, XC343664, XC343667, XC343668, XC343669, XC343676, XC343683, XC343686, XC343803, XC343805, XC343807), Thomas Luthi (XC357122), Timo Roeke (XC346551), Timo Tschentscher (XC266489), Tom Wulf (XC329937, XC329942, XC329944, XC329955), Tomek Tumiel (XC282634, XC331530), Uku Paal (XC172697, XC215916, XC219334, XC219335, XC219336), Volker Arnold (XC130274). This research has been partially funded by the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (Generalitat de Catalunya) under grants ref. 2014-SGR-0590 and ref. 2014-SGR-589.

Author Contributions: Vidaña-Vila conceived the proposed data set, performed the experiments, and co-wrote the paper. Navarro analyzed resulting data and co-wrote the paper. Alsina-Pagès designed the experiments and participated in the writing and reviewing of the entire paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MFCC Mel Frequency Cepstral Coefficients

t-SNE t-Distributed Stochastic Neighbor Embedding

References

1. Witzany, G. *Biocommunication of Animals*; Springer: Dordrecht, The Netherlands, 2014; ISBN 978-94-007-7413-1.
2. Kershenbaum, A.A.; Blumstein, D.T.; Roch, M.A.; Akcay, G.; Backus, G.; Bee, M.A.; Bohn, K.; Cao, Y.; Carter, G.; Casar, C.; et al. Acoustic sequences in non-human animals: A tutorial review and prospectus. *Biol. Rev.* **2014**, doi:10.1111/brv.12160.
3. Slabbekoorn, H.; Smith, T.B. Bird song, ecology and speciation. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* **2002**, *357*, 493–503.
4. Howell, S.N.G.; Sophie, W. *A Guide to the Birds of Mexico and Northern Central America*; Oxford University Press: Oxford, UK, 1995; ISBN 0-19-854012-4.
5. Baker, M.C.; Cunningham, M.A. The biology of bird-song dialects. *Behav. Brain Sci.* **1985**, *8*, 85–100.
6. Stowell, D.; Wood, M.; Stylianou, Y.; Glotin, H. Bird detection in audio: A survey and a challenge. In Proceedings of the IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), Salerno, Italy, 13–16 September 2016; pp. 1–6. doi:10.1109/MLSP.2016.7738875
7. Sivarajan, S.; Schuller, B.; Coutinho, E. Bird Sound Classification. Available online: <http://www.imperial.ac.uk/media/imperial-college/faculty-of-engineering/computing/public/SINDURANSIVARAJAN.pdf> (accessed on 11 April 2017).

8. Noda, J.J.; Travieso, C.M.; Sánchez-Rodríguez, D. Automatic Taxonomic Classification of Fish Based on Their Acoustic Signals. *Appl. Sci.* **2016**, *6*, 443, doi:10.3390/app6120443.
9. Mermelstein, P. Distance measures for speech recognition—psychological and instrumental. *Pattern Recogn. Artif. Intell.* **1976**, *166*, 374–388.
10. Van der Maaten, L.J.P.; Hinton, G.E. Visualizing high-dimensional data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
11. Vellinga, W.P.; Planque, R. The Xeno-canto collection and its relation to sound recognition and classification. In Proceedings of the 2015 CLEF, Toulouse, France, 8–11 September 2015.
12. Xeno-Canto Foundation. Xeno-Canto: Sharing Bird Sounds from around the World. Available online: <http://www.xeno-canto.org/> (accessed on 15 April 2017).
13. Cornell Library of Ornithology Macaulay Library. Available online: <http://macaulaylibrary.org> (accessed on 11 April 2017).
14. University of Utah Western Soundscape Archive. Available online: <http://www.westernsoundscape.org> (accessed on 11 April 2017).
15. Lepage, D. Avibase—The World Bird Database. Available online: <http://avibase.bsc-eoc.org/avibase.jsp> (accessed on 11 April 2017).
16. British Library Sound Archive. Available online: <http://www.bl.uk/soundarchive> (accessed on 11 April 2017).
17. Tierstimmenarchiv, Animal Sound Archive (Tierstimmenarchiv) at the Museum für Naturkunde in Berlin. Available online: <http://www.tierstimmenarchiv.de> (accessed on 11 April 2017).
18. CSIRO. Australian National Wildlife Collection Sound Archive. Available online: <https://www.csiro.au/en/Research/Collections/ANWC/About-ANWC/Our-wildlife-sound-archive> (accessed on 11 April 2017).
19. Mikusiński, G.; Gromadzki, M.; Chylarecki, P. Woodpeckers as Indicators of Forest Bird Diversity. *Conserv. Biol.* **2001**, *15*, 208–217, doi:10.1111/j.1523-1739.2001.99236.x.
20. Real Decreto 139/2011, de 4 de Febrero, Para el Desarrollo del Listado de Especies Silvestres en Régimen de Protección Especial y del Catálogo Español de Especies Amenazadas. Available online: <http://www.boe.es/buscar/act.php?id=BOE-A-2011-3582> (accessed on 15 April 2017).
21. Pakkala, T.; Lindén, A.; Tiainen, J.; Tomppo, E.; Kouki, J. Indicators of forest biodiversity: Which Bird Species Predict High Breeding Bird Assemblage Diversity in Boreal Forests at Multiple Spatial Scales? In *Annales Zoologici Fennici*; Finnish Zoological and Botanical Publishing: Helsinki, Finland, 2014; Volume 51, pp. 457–476.
22. Sprengel, E.; Jaggi, M.; Kilcher, Y.; Hofmann, T. Audio Based Bird Species Identification using Deep Learning Techniques. In Proceedings of CEUR Workshop, Evora, Portugal, 5–8 September 2016; pp. 547–559.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).