

Contents

1	Motivation and Contributions	1
1.1	Avian ecology	1
1.2	Democratizing deep neural networks	3
1.3	Methodology and outline	4
1.4	Results	6
2	Theory on Bird Biology and Audio Signal Processing	11
2.1	On the relationship of humans and birds	12
2.2	Avian vocal behavior	13
2.2.1	Studying bird sounds	14
2.2.2	Production and perception of sounds in birds	17
2.2.3	Vocal development and repertoires	26
2.2.4	Song variation in space and time	31
2.2.5	Function of bird song and singing behavior	34
2.3	Avian ecology	36
2.3.1	Habitats, abundance, and migration	37
2.3.2	Conservation biology and citizen science	39
2.4	Sound digitization and representation	44
2.4.1	Digital sound recording	44
2.4.2	Spectrogram computation	47
2.4.3	Adaption to avian acoustic monitoring	55
2.5	Summary	58
3	Acoustic Events and Deep Learning	61
3.1	Acoustic event recognition	61
3.2	Deep artificial neural networks	64
3.2.1	Concepts and topologies	64

3.2.2	Evolution of CNN	73
3.3	Evaluation campaigns for avian acoustics	76
3.3.1	BirdCLEF	77
3.3.2	DCASE Bird Audio Detection	81
3.4	Summary	84
4	System Architecture	85
4.1	Design Decisions	86
4.1.1	Extensive functionality	87
4.1.2	Detailed configuration	90
4.1.3	Domain-agnostic workflow	91
4.1.4	Reproducibility and transparency	91
4.1.5	Distribution and applications	92
4.2	Concepts and workflows	94
4.2.1	Components	94
4.2.2	Training	96
4.2.3	Evaluation	106
4.2.4	Distribution	111
4.3	Implementation details	112
4.3.1	Audio and image processing	113
4.3.2	Deep learning frameworks	114
4.3.3	Web tools and services	116
4.4	Summary	117
5	Experiments	119
5.1	Goals and main focus	120
5.2	Experimental investigations	120
5.2.1	Acquisition and composition of data	121
5.2.2	Experimental setup	125
5.2.3	Spectrogram computation	129
5.2.4	Architecture design	133
5.2.5	Topologies and training regimes	141
5.2.6	Mobile architectures	152
5.2.7	Benchmark system	157
5.3	Results	166

5.4	Summary	181
6	Applications	183
6.1	Live stream analysis	183
6.1.1	Abundance and vocal activity	185
6.1.2	Temperature and vocal activity	186
6.2	SWIFT soundscape analysis	189
6.2.1	AMTiC	189
6.2.2	SWAMP	193
6.3	Smartphone app	197
6.3.1	Scope	198
6.3.2	Design	198
6.3.3	Distribution	200
6.3.4	Reception	201
6.3.5	Results	202
6.4	HaikuBox	208
6.5	Summary	209
7	Conclusion and Future Work	211
	Bibliography	XIX
A	Year-round Vocal Activity	A1
B	AMTiC Detection Plots	A5
C	SWAMP Spatial Maps	A10
D	Species-specific Results	A16