

Project Ornithowave: Bird Species Identification

Avian diversity is a commonly used metric to assess environmental change and is growing in importance. In this project we use the MAX78000 to detect and classify birds based on their songs and calls. Thereby enabling remote sensing of avian diversity

Aldfly: Alder Flycatcher

- One song and 3 calls
- Males sing a burry rreeee-BEEP or free-beer!, reminiscent of but thinner than an Eastern Phoebe's song.
- CALLS
- The primary contact call is a low, flat pip.



https://www.allaboutbirds.org/guide/Alder_Flycatcher/sounds

1

Bkcchi: Black-capped Chickadee

- One song and 4+ calls
- SONGS
In most of North America, the song is a simple, pure 2 or 3-note whistled fee-bee or hey, sweetie. In the Pacific Northwest, the song is 3 or 4 notes on the same pitch; the song is also different on Martha's Vineyard in MA. In much of the range, males begin singing in mid-January, and the song increases in frequency as winter progresses. Females also sing occasionally.
- CALLS
Chickadees make their chickadee-dee-dee call using increasing numbers of dee notes when they are alarmed. They also have a garbling call, often given aggressively when a lower-ranking bird gets close to a higher-ranking one; also exchanged between members of a pair. Black-capped Chickadees make a high pitched see as a high-intensity alarm call, often when a fast-approaching predator is detected. When chickadees hear this call, they freeze in position until they hear a chickadee-dee call signifying "all clear". High see calls most often given by males.
- OTHER SOUNDS
Chickadee nestings make an explosive hiss and slap the inside of their nest cavity when an intruder looks in.



https://www.allaboutbirds.org/guide/Black-capped_Chickadee/sounds

2

Bushti: American Bushtit

- 3 Calls, 1 alarm call
- SONGS
Bushtits are active birds that use lots of contact calls, but they don't really have a song. Occasionally, several individuals gathered together make a long, drawn-out series of quiet twitters and chips.
- CALLS
Bushtits make several kinds of short, high, wispy contact calls or chip notes. These help group members know where flockmates are. They can intensify to indicate nesting activities or when mobbing predators or confronting opponents. Individuals that get separated from a group make a rapid series of high-pitched chip notes that carries well.



4

Gryfly: Grey Flycatcher

- One song, two calls
- SONGS
Males sing a sharp, quick chibit often followed by a higher-pitched, rising tweet. The use of these elements varies, with some (especially the chibit) repeated in quick sequences. Males also perform a flight song, a rapid series of whist calls notes followed by a jumble of other call types.
- CALLS
Both sexes give a loud whist call note.
- OTHER SOUNDS
Bill snapping serves as a threat display and is occasionally heard when birds capture an insect.



https://www.allaboutbirds.org/guide/Gray_Flycatcher/sounds

5

Motivation & Overview

Avian diversity is a commonly used metric to assess environmental change and is growing in importance due to climate change and its ecological effects. In the last few years, deep learning approaches have shown promise in reducing the expert labor intensive bird identification process. The 157 layer ResNet based BirdNet from Cornell University is currently the state of the art in this field and can identify almost a 1000 different bird species [1]. However implementations of this model are limited to areas where high bandwidth internet is available as the model is evaluated in the cloud.

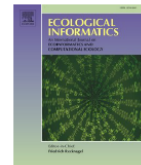
Ecological Informatics 61 (2021) 101236



Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Ecological Informatics

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



BirdNET: A deep learning solution for avian diversity monitoring

Stefan Kahl^{a,*}, Connor M. Wood^a, Maximilian Eibl^b, Holger Klinck^a

^a Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY 14850, USA

^b Technische Universität Chemnitz, D-09111 Chemnitz, Germany



We propose a simplified CNN based on the KWS-20 model to identify 20 different bird species based on their sounds and calls.

This project running on the ultra-low power MAX78000 can make bird monitoring possible in very remote locations without worrying about high bandwidth connectivity or big batteries. It can also be used by bird enthusiasts who want to keep track of bird species visiting their home or farm.

This Project Execution Comprised of 4 phases:

1. Literature Review & Open-Source Dataset Identification
2. Data Engineering
3. Model training & C-code synthesis using ai8x-training & ai8x-synthesis libraries
4. Testing on Hardware using audio samples from the test set.

Phase 1: Literature Review & Open-Source Dataset Identification

The Cornell Lab of Ornithology is an active research group in this space and regularly publish papers, and host competitions on Kaggle (BirdCLEF 2021: <https://www.kaggle.com/c/birdclef-2021>, CBL 2020: <https://www.kaggle.com/c/birdsong-recognition/overview>)

**Research Code Competition**

**BirdCLEF 2021 - Birdcall Identification**

Identify bird calls in soundscape recordings

Cornell Lab of Ornithology · 816 teams · a month ago



\$5,000
Prize Money

Their publications and datasets served as a solid foundation for this project. The networks that win these competition are very large and complex, for our project we used a subset of this dataset and retrained the KWS-20 model (ai85kws20) from maxim by replacing words with birds.

However, a major constraint is that all this data is weakly labelled and not standardized with varying lengths, quality and sampling frequency.

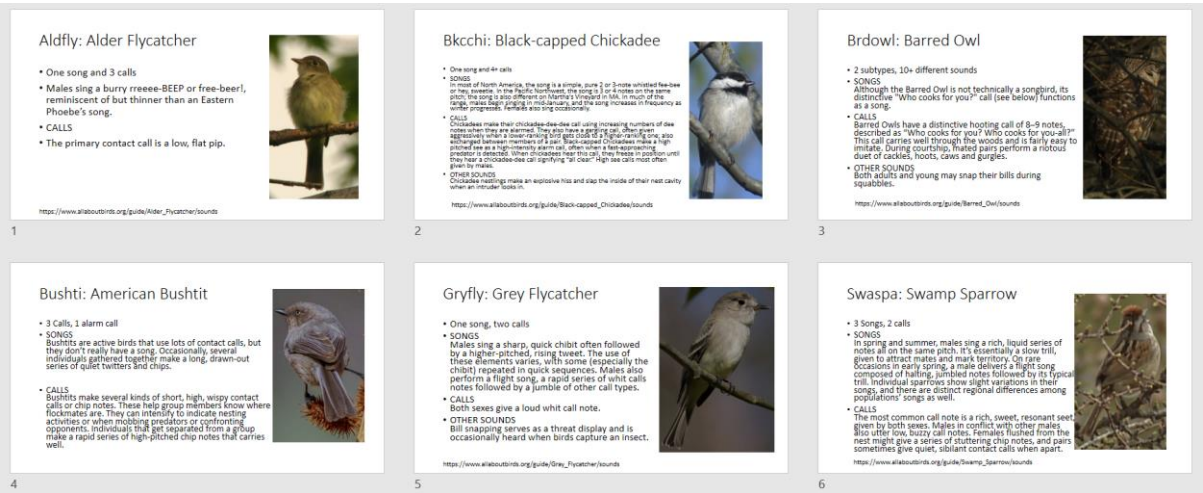
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI								
1	rating	play	bird	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id	_id							
78	4	no	aidfly	1	monos	Not spec	14	KC318891	Not spec	Alder Flyc	Not spec	KC318891	0	yes	Empidon	Draper T.	42.7646	44100	Hz	song	280	m				192000	0	mp3	Not spec	318899	https://	United St	Antonio X	Empidon	-83.116	Not spec	08.30	Antonio X	Creative C				
79	4	no	aidfly	1	monos	increasin	6	KC319021	level	Alder Flyc	01-Mar	KC319021	0	yes	Empidon	Bangor F.	44.8234	16000	Hz	call	40	m				319027	https://	United St	rob	acorn	Empidon	-88.765	0-361	09.30	rob	acorn	Creative C						
80	3.5	no	aidfly	2	stereos	Not spec	96	KC32481	level	Alder Flyc	01-Mar	KC32481	0	yes	Empidon	Jasper M.	52.8924	48000	Hz	song	1100	m				320000	0	mp3	level	Comm	324888	https://	Canada	Christoph	Empidon	-118.1	0-361	08.53	Christoph	Creative C			
81	3.5	no	aidfly	2	stereos	Not spec	178	KC33044	Not spec	Alder Flyc	Not spec	KC33044	0	yes	Empidon	For&ts M.	47.3197	44100	Hz	call, song	800	m				128000	0	mp3	Not spec	White	330449	https://	Canada	Martin St	Empidon	-71.146	Not spec	08.44	Martin St	Creative C			
82	3.5	no	aidfly	2	stereos	Not spec	64	KC33156	Not spec	Alder Flyc	Not spec	KC33156	0	yes	Empidon	Goldman	49.0889	44100	Hz	male, son	100	m				128000	0	mp3	Not spec	Song	331564	https://	Canada	Martin St	Empidon	-74.265	Not spec	06.19	Martin St	Creative C			
83	3.5	no	aidfly	2	stereos	Not spec	70	KC33468	level	Alder Flyc	Jul-20	KC33468	0	yes	Empidon	Chv Ave.	47.146	44100	Hz	song	400	m				160000	0	mp3	level	Comm	334685	https://	United St	Scott	Gray	Empidon	-92.717	0-361	08.00	Scott	Gray	Creative C	
84	3.5	yes	aidfly	1	monos	Not spec	103	KC37082	level	Alder Flyc	01-Mar	KC37082	0	yes	Empidon	Maine M.	44.5249	44100	Hz	song	50	m				224000	0	mp3	both	Amer	370826	https://	United St	Matt W	Empidon	-69.038	0-361	10.30	Matt W	Creative C			
85	4	no	aidfly	1	monos	Not spec	177	KC37088	Not spec	Alder Flyc	Not spec	KC37088	0	yes	Empidon	Elser's R.	41.8605	48000	Hz	call	240	m				320000	0	mp3	Not spec	ified	370881	https://	United St	Matt W	Empidon	-88.176	Not spec	08.32	Matt W	Creative C			
86	3.5	no	aidfly	1	monos	Not spec	5	KC37095	Not spec	Alder Flyc	Not spec	KC37095	0	yes	Empidon	Steven's I.	40.7457	44100	Hz	male, son	20	m				128000	0	mp3	Not spec	Will	370952	https://	United St	Dominic	Empidon	-74.024	Not spec	08.00	Dominic	Creative C			
87	4	no	aidfly	1	monos	Not spec	13	KC37141	Not spec	Alder Flyc	Not spec	KC37141	0	yes	Empidon	Elser's R.	41.8605	48000	Hz	wee	240	m				320000	0	mp3	Not spec	ified	371412	https://	United St	Matt W	Empidon	-88.176	Not spec	06.33	Matt W	Creative C			
88	4	no	aidfly	1	monos	Not spec	1	KC37388	level	Alder Flyc	01-Mar	KC37388	0	yes	Empidon	Massach	42.233	44100	Hz	song	170	m				96689	0	mp3	level	White	373885	https://	United St	Joseph B	Empidon	-71.764	0-361	13.00	Joseph B	Creative C			
89	3.5	yes	aidfly	1	monos	Not spec	10	KC37468	Not spec	Alder Flyc	Not spec	KC37468	0	yes	Empidon	St. Francis	39.7791	44100	Hz	male, son	1000	m				87990	0	mp3	Not spec	Hous	374626	https://	United St	Eric Houg	Empidon	-101.81	Not spec	14.30	Eric Houg	Creative C			
90	3.5	yes	aidfly	1	monos	Not spec	33	KC37468	Not spec	Alder Flyc	Not spec	KC37468	0	yes	Empidon	St. Francis	39.7791	44100	Hz	call, male	1000	m				87990	0	mp3	level	White	376468	https://	Canada	Michael I	Empidon	-114.61	0-361	07.30	Michael I	Creative C			
91	4.5	no	aidfly	1	monos	Not spec	48	KC38003	level	Alder Flyc	01-Mar	KC38003	0	yes	Empidon	Watts Fl.	42.033	44100	Hz	song	440	m				128000	0	mp3	decreasin	Seaw	380033	https://	United St	Jim Berry	Empidon	-79.422	0-361	09.00	Jim Berry	Creative C			
92	4	no	aidfly	1	monos	Not spec	21	KC38187	level	Alder Flyc	04-Jun	KC38187	0	yes	Empidon	Milbridge	44.5358	48000	Hz	song	10	m				320000	0	mp3	increasin	g	381871	https://	United St	Sue Riffe	Empidon	-67.841	0-361	09.30	Sue Riffe	Creative C			
93	4	no	aidfly	2	stereos	Not spec	45	KC40625	Not spec	Alder Flyc	Not spec	KC40625	0	yes	Empidon	Potter M.	61.0776	48000	Hz	song	0	m				320000	0	mp3	Not spec	ified	406251	https://	United St	Patric A.	Empidon	-149.83	Not spec	09.33	Patric A.	Creative C			
94	4	aidfly	2	stereos	Not spec	27	KC41451	Not spec	Alder Flyc	Not spec	KC41451	0	yes	Empidon	Maracy	4.8059	44100	Hz	call	1100	m				bird-seen	128000	0	mp3	Not spec	ified	414551	https://	Colombia	Jaime Su	Empidon	-75.87	Not spec	12.30	Jaime Su	Creative C			
95	3.5	no	aidfly	2	stereos	Not spec	51	KC41814	level	Alder Flyc	01-Mar	KC41814	0	yes	Empidon	Fraser-P.	53.6227	44100	Hz	song	850	m				bird call	128000	0	mp3	both	Ruby	418144	https://	Canada	Jeff Dyck	Empidon	-123.0	0-361	13.30	Jeff Dyck	Creative C		
96	4	no	aidfly	2	stereos	Not spec	25	KC42042	Not spec	Alder Flyc	Not spec	KC42042	0	yes	Empidon	Roan Mt.	36.0973	44100	Hz	song	1900	m				No modif	128000	0	mp3	Not spec	ified	420402	https://	United St	Scott	Gray	Empidon	-82.137	Not spec	11.00	Scott	Gray	Creative C
97	4	aidfly	2	stereos	Not spec	31	KC42090	Not spec	Alder Flyc	Not spec	KC42090	0	yes	Empidon	Finger La.	42.5053	48000	Hz	song	550	m				192000	0	mp3	Not spec	ified	420909	https://	United St	GABRIEL L	Empidon	-76.784	Not spec	08.30	GABRIEL L	Creative C				
98	4.5	no	aidfly	2	stereos	Not spec	43	KC47950	Not spec	Alder Flyc	Not spec	KC47950	0	yes	Empidon	Caroline	42.3306	44100	Hz	song	480	m				192000	0	mp3	Not spec	Tufha	479502	https://	United St	Meena H	Empidon	-76.335	Not spec	09.30	Meena H	Creative C			
99	4.5	no	aidfly	2	stereos	Not spec	98	KC47950	Not spec	Alder Flyc	Not spec	KC47950	0	yes	Empidon	Caroline	42.3306	44100	Hz	song	480	m				192000	0	mp3	Not spec	Tufha	479504	https://	United St	Meena H	Empidon	-76.335	Not spec	10.00	Meena H	Creative C			
100	4	aidfly	2	stereos	Not spec	10	KC78890	level	Alder Flyc	01-Mar	KC78890	0	yes	Empidon	Murphy-I	44.707	44100	Hz	song	280	m						196821	0	mp3	increasin	g	78890	https://	United St	Jonathan	Empidon	-63.319	0-361	12.20	Jonathan	Creative C		
101	4	no	blecchi	1	monos	Not spec	38	KC11407	both	Black-cap	04-Jun	KC11407	0	yes	Poeclie	at Envoronh	40.557	48000	Hz	call	1500	m				Three ind	48000	0	mp3	increasin	g	114073	https://	United St	Eric Defo	Poeclie	-105.02	0-361	11.46	Eric Defo	Creative C		
102	4.5	no	blecchi	1	monos	Not spec	83	KC11408	Not spec	Black-cap	Not spec	KC11408	0	yes	Poeclie	at Envoronh	44.942	44100	Hz	call	260	m				Callis	138000	0	mp3	Not spec	Amer	114086	https://	United St	Jonathan	Poeclie	-63.259	Not spec	14.26	Jonathan	Creative C		
103	4.5	no	blecchi	2	stereos	decreasin	33	KC12087	declearat	Black-cap	Jul-20	KC12087	0	yes	Poeclie	at Stoneh	47.0655	44100	Hz	call	380	m				One of the	128000	0	mp3	increasin	Red-b	120878	https://	Canada	Martin St	Poeclie	-71.415	0-361	08.51	Martin St	Creative C		
104	5	no	blecchi	1	monos	Not spec	51	KC12106	Not spec	Black-cap	Not spec	KC12106	0	yes	Poeclie	at Minnesot	44.84	44100	Hz	call	220	m				Two ind	192000	0	mp3	Not spec	ified	121068	https://	United St	Jonathan	Poeclie	-63.259	Not spec	09.30	Jonathan	Creative C		
105	5	no	blecchi	2	stereos	Not spec	59	KC12350	Not spec	Black-cap	Not spec	KC12350	0	yes	Poeclie	at St. Mary	44.9215	48000	Hz	call	260	m				Soft, high	192000	0	mp3	Not spec	Amer	123506	https://	United St	Jonathan	Poeclie	-63.259	Not spec	08.40	Jonathan	Creative C		
106	5	no	blecchi	2	stereos	Not spec	18	KC13195	level	Black-cap	01-Mar	KC13195	0	yes	Poeclie	at R&DServ	47.069	44100	Hz	call	0	m				One ind	128000	0	mp3	level	White	131952	https://	Canada	Martin St	Poeclie	-70.794	0-361	08.26	Martin St	Creative C		
107	4.5	no	blecchi	1	monos	Not spec	235	KC13286	Not spec	Black-cap	Not spec	KC13286	0	yes	Poeclie	at Lake Ben	47.5992	48000	Hz	song	420	m				Several b	192000	0	mp3	Not spec	Myrt	132861	https://	United St	Jonathan	Poeclie	-94.826	Not spec	14.25	Jonathan	Creative C		
108	4	no	blecchi	1	monos	Not spec	151	KC13402	Not spec	Black-cap	Not spec	KC13402	0	yes	Poeclie	at Maronell	41.32	44100	Hz	alarm	ca	20	m				Vocalizat	128000	0	mp3	Not spec	ified	134027	https://	United St	Dan Lane	Poeclie	-70.545	Not spec	0	Dan Lane	Creative C	
109	3.5	no	blecchi	2	stereos	decreasin	25	KC13547	level	Black-cap	01-Mar	KC13547	0	yes	Poeclie	at New Four	35.6109	44100	Hz	song	1500	m						128000	0	mp3	level	Gold	135477	https://	United St	Mike Nels	Poeclie	-83.425	0-361	09.30	Mike Nels	Creative C	
110	4	no	blecchi	2	stereos	decreasin	8	KC13547	level	Black-cap	01-Mar	KC13547	0	yes	Poeclie	at New Four	35.6109	44100	Hz	song	1500	m						128000	0	mp3	level	Gold	135478	https://	United St	Mike Nels	Poeclie	-83.425	0-361	09.30	Mike Nels	Creative C	
111	4	no	blecchi	2	stereos	Not spec	24	KC13905	both	Black-cap	04-Jun	KC13905	0	yes	Poeclie	at Noland	49.5727	44100	Hz	call	1700	m						128000	0	mp3	level	White	139055	https://	United St	Kim Allen	Poeclie	-83.48	0-361	07.30	Kim Allen	Creative C	
112	3.5	no	blecchi	1	monos	Not spec	49	KC13943	Not spec	Black-cap	Not spec	KC13943	0	yes	Poeclie	at Millbury	42.1752	12000	Hz	he	g	180	m				Flock of 8	140000	0	mp3	Not spec	Comm	139438	https://	United St	Kim Allen	Poeclie	-71.785	Not spec	08.45	Kim Allen	Creative C	
113	4	no	blecchi	2	stereos	Not spec	24	KC13960	level	Black-cap	01-Mar	KC13960	0	yes	Poeclie	at Noland	49.5727	44100	Hz	call	1700	m						128000	0	mp3	level	White	139603	https://	United St	Mike Nels	Poeclie	-83.48	0-361	07.30	Mike Nels	Creative C	
114	5	no	blecchi	2	stereos	Not spec	24	KC13960	level	Black-cap	01-Mar	KC13960	0	yes	Poeclie	at Noland	49.5727	44100	Hz	call	1700	m						128000	0	mp3	level	White	139605	https://	United St	Mike Nels	Poeclie	-83.48	0-361	07.30	Mike Nels	Creative C	

Phase 2: Data Engineering

This the most challenging phase.
Upon closer inspection it was observed that all of the data in the open-source datasets come from this one website: <https://www.xeno-canto.org/>



Data on this website is collected by birders all over the world and hence the quality can vary from sample to sample. There is no standardization of sample rates and sometimes the bird sound is not very clear. Each bird can also make multiple different calls, sounds and alarms, adding to the complexity. A summary of a subset of the dataset (6 birds) can be found in image 6-birds.png, train-6.xls and dataset-meta.png



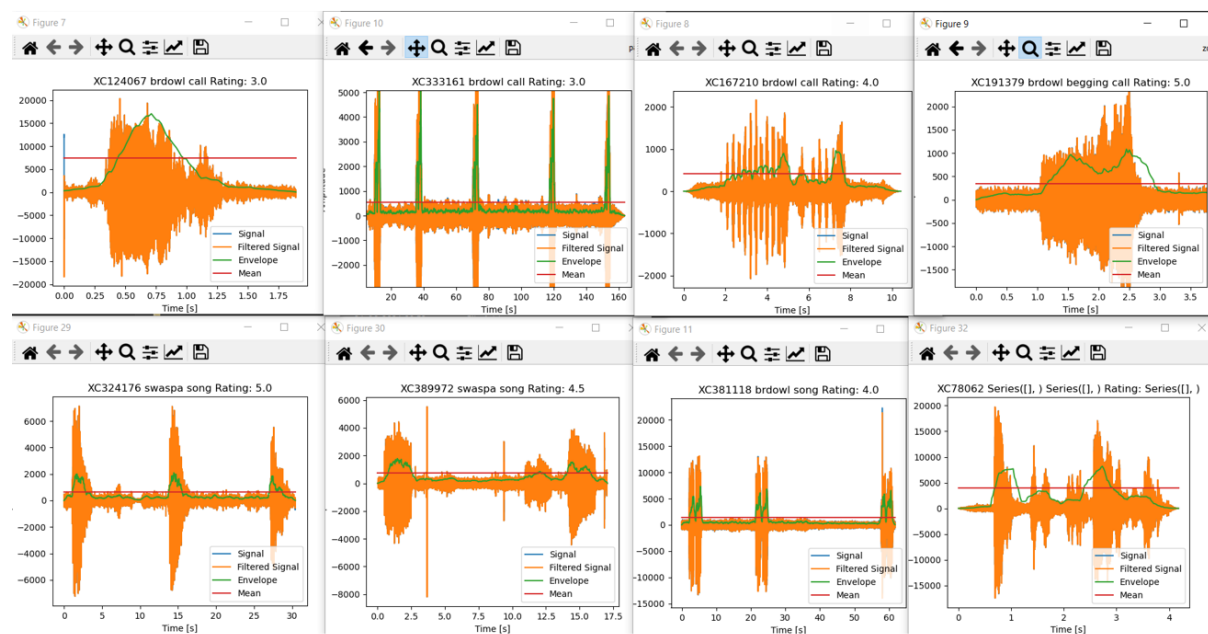
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI							
1	rating	play	chbird	cod	channels	date	pitch	duration	filename	speed	species	number	title	second	bird	s	v	o	p	q	r	s	t	u	v	w	x	y	z	aa	ab	ac	ad	ae	af	ag	ah	ai				
78	4	no	aidfly	1	monos	Not specifi	14	KC31899	Not specifi	Alder Flyc	01-Mar	KC31891	0	no	Empidonax	Draper	7	42.7646	44100	Hz	song	280	m		192000	0	mp3	Not specifi	318899	https://www.xeno-canto.org/sounds/318899	United	Sto	Antonia	X	Empidon	85.116	Not specifi	08-30	Antonio X Creative C			
79	4	no	aidfly	1	monos	Not specifi	6	KC31901	level	Alder Flyc	01-Mar	KC31901	0	yes	Empidonax	Jasper	FI	51.8234	16000	Hz	call	40	m	Recorded	48814	0	mp3	level	Comm	319017	https://www.xeno-canto.org/sounds/319017	United	Sto	Rob	acorn	Empidon	48.765	0-36s	08-30	Rob acorn Creative C		
80	3.5	no	aidfly	2	stereos	Not specifi	96	KC32348	level	Alder Flyc	01-Mar	KC32348	0	yes	Empidonax	Benjamin	FI	51.8234	16000	Hz	song	1100	m	Singing	320000	0	mp3	level	Comm	323488	https://www.xeno-canto.org/sounds/323488	Canada	Christoph	Empidon	118.1	0-36s	08-53	Christoph Creative C				
81	3.5	no	aidfly	2	stereos	Not specifi	178	KC33044	Not specifi	Alder Flyc	01-Mar	KC33044	0	yes	Empidonax	Forbes	FI	47.3197	44100	Hz	song	800	m	First ind	128000	0	mp3	Not specifi	White	330449	https://www.xeno-canto.org/sounds/330449	Canada	Martin	St	Empidon	71.146	Not specifi	08-44	Martin St Creative C			
82	3.5	no	aidfly	2	stereos	Not specifi	64	KC33156	Not specifi	Alder Flyc	01-Mar	KC33156	0	yes	Empidonax	Goldman	FI	45.0889	44100	Hz	male	son	100	m	Habitat	128000	0	mp3	Not specifi	Song	331564	https://www.xeno-canto.org/sounds/331564	Canada	Martin	St	Empidon	74.263	Not specifi	06-19	Martin St Creative C		
83	3.5	no	aidfly	2	stereos	Not specifi	70	KC33468	level	Alder Flyc	01-Mar	KC33468	0	yes	Empidonax	Owl Ave	FI	47.146	44100	Hz	song	400	m	Raw recor	160000	0	mp3	level	Comm	334685	https://www.xeno-canto.org/sounds/334685	United	Sto	Scott	Grav	Empidon	92.727	0-36s	08-00	Scott Grav Creative C		
84	3.5	no	aidfly	1	monos	Not specifi	103	KC37082	level	Alder Flyc	01-Mar	KC37082	0	yes	Empidonax	Marine	FI	44.3249	44100	Hz	song	50	m	I was wal	224000	0	mp3	both	Amer	370826	https://www.xeno-canto.org/sounds/370826	United	Sto	Watt	Watt	Empidon	69.038	0-36s	10-30	Watt W Creative C		
85	4	no	aidfly	1	monos	Not specifi	177	KC37088	Not specifi	Alder Flyc	01-Mar	KC37088	0	yes	Empidonax	Elson	FI	41.8605	48000	Hz	call	240	m		320000	0	mp3	Not specifi	White	370881	https://www.xeno-canto.org/sounds/370881	United	Sto	Watt	Watt	Empidon	88.176	Not specifi	08-53	Matt Watt Creative C		
86	3.5	no	aidfly	1	monos	Not specifi	5	KC37095	Not specifi	Alder Flyc	01-Mar	KC37095	0	yes	Empidonax	Stevens	FI	40.4757	44100	Hz	male	son	20	m		128000	0	mp3	Not specifi	White	370952	https://www.xeno-canto.org/sounds/370952	United	Sto	Dominic	Empidon	74.024	Not specifi	08-00	Dominic Creative C		
87	4	no	aidfly	1	monos	Not specifi	13	KC37141	Not specifi	Alder Flyc	01-Mar	KC37141	0	yes	Empidonax	Elson	FI	41.8605	48000	Hz	zevaco	240	m		320000	0	mp3	Not specifi	White	371412	https://www.xeno-canto.org/sounds/371412	United	Sto	Watt	Watt	Empidon	88.176	Not specifi	08-53	Matt Watt Creative C		
88	4	no	aidfly	1	monos	Not specifi	1	KC37388	level	Alder Flyc	01-Mar	KC37388	0	yes	Empidonax	Massach	FI	42.233	44100	Hz	song	170	m		96689	0	mp3	level	White	373885	https://www.xeno-canto.org/sounds/373885	United	Sto	Joseph	Rob	Empidon	71.794	0-36s	13-00	Joseph Rob Creative C		
89	3.5	no	aidfly	1	monos	Not specifi	10	KC37463	Not specifi	Alder Flyc	01-Mar	KC37463	0	yes	Empidonax	St. Francis	FI	39.7791	44100	Hz	male	son	1000	m	Rare this	97909	0	mp3	Not specifi	Hous	374636	https://www.xeno-canto.org/sounds/374636	United	Sto	Eric	Houg	Empidon	101.81	Not specifi	14-30	Eric Houg Creative C	
90	3.5	no	aidfly	1	monos	Not specifi	33	KC37463	Not specifi	Alder Flyc	01-Mar	KC37463	0	yes	Empidonax	St. Francis	FI	39.7791	44100	Hz	male	son	1000	m	Rare this	97909	0	mp3	Not specifi	Hous	374638	https://www.xeno-canto.org/sounds/374638	United	Sto	Eric	Houg	Empidon	101.81	Not specifi	14-30	Eric Houg Creative C	
91	4.5	no	aidfly	1	monos	Not specifi	20	KC37668	level	Alder Flyc	01-Mar	KC37668	0	yes	Empidonax	Winchell	FI	51.4555	48000	Hz	male	son	1300	m	Black	192000	0	mp3	level	White	376688	https://www.xeno-canto.org/sounds/376688	Canada	Michael	Empidon	114.61	0-36s	07-30	Michael Creative C			
92	3.5	no	aidfly	1	monos	Not specifi	48	KC38003	level	Alder Flyc	01-Mar	KC38003	0	yes	Empidonax	Watts	FI	42.038	44100	Hz	song	440	m		128000	0	mp3	decreas	Swat	380033	https://www.xeno-canto.org/sounds/380033	United	Sto	Jim	Berry	Empidon	79.412	0-36s	09-00	Jim Berry Creative C		
93	4	no	aidfly	1	monos	Not specifi	21	KC38187	level	Alder Flyc	01-Mar	KC38187	0	yes	Empidonax	Millbridge	FI	44.5358	48000	Hz	song	10	m	Natural v	320000	0	mp3	increas	Ing	381871	https://www.xeno-canto.org/sounds/381871	United	Sto	Sue	Riffe	Empidon	67.881	0-36s	05-30	Sue Riffe Creative C		
94	4	no	aidfly	2	stereos	Not specifi	45	KC40625	Not specifi	Alder Flyc	01-Mar	KC40625	0	yes	Empidonax	Potter	FI	61.0776	48000	Hz	song	0	m		320000	0	mp3	Not specifi	Ing	406251	https://www.xeno-canto.org/sounds/406251	United	Sto	Patric	A	Empidon	149.83	Not specifi	09-33	Patric A Creative C		
95	4	no	aidfly	2	stereos	Not specifi	27	KC41455	Not specifi	Alder Flyc	01-Mar	KC41455	0	yes	Empidonax	Maracey	FI	48.059	44100	Hz	call	1100	m	bird	son	128000	0	mp3	Not specifi	Ing	414551	https://www.xeno-canto.org/sounds/414551	Colombia	Jaime	Sul	Empidon	75.87	Not specifi	12-30	Jaime Sul Creative C		
96	3.5	no	aidfly	2	stereos	Not specifi	51	KC41814	level	Alder Flyc	01-Mar	KC41814	0	yes	Empidonax	Fraser	FI	55.6227	44100	Hz	song	850	m	Bird call	128000	0	mp3	both	Ruby	418144	https://www.xeno-canto.org/sounds/418144	Canada	Jeff Dyck	Empidon	112.0	0-36s	13-30	Jeff Dyck Creative C				
97	4	no	aidfly	2	stereos	Not specifi	25	KC42040	Not specifi	Alder Flyc	01-Mar	KC42040	0	yes	Empidonax	Roan Mt	FI	36.0973	44100	Hz	song	1900	m	No modifi	128000	0	mp3	Not specifi	Ing	420402	https://www.xeno-canto.org/sounds/420402	United	Sto	Scott	Grav	Empidon	82.137	Not specifi	11-00	Scott Grav Creative C		
98	4	no	aidfly	1	monos	Not specifi	31	KC42090	Not specifi	Alder Flyc	01-Mar	KC42090	0	yes	Empidonax	Finger	FI	42.5058	48000	Hz	song	550	m		192000	0	mp3	Not specifi	Ing	420909	https://www.xeno-canto.org/sounds/420909	United	Sto	GABRIEL	Empidon	78.794	Not specifi	08-30	GABRIEL Creative C			
99	4.5	no	aidfly	2	stereos	Not specifi	43	KC47950	Not specifi	Alder Flyc	01-Mar	KC47950	0	yes	Empidonax	Caroline	FI	42.3306	44100	Hz	song	460	m		192000	0	mp3	Not specifi	Tufes	479502	https://www.xeno-canto.org/sounds/479502	United	Sto	Meena	He	Empidon	76.335	Not specifi	09-30	Meena He Creative C		
100	4.5	no	aidfly	2	stereos	Not specifi	98	KC47950	Not specifi	Alder Flyc	01-Mar	KC47950	0	yes	Empidonax	Caroline	FI	42.3306	44100	Hz	song	460	m		192000	0	mp3	Not specifi	Tufes	479504	https://www.xeno-canto.org/sounds/479504	United	Sto	Meena	He	Empidon	76.335	Not specifi	10-00	Meena He Creative C		
101	4	no	aidfly	2	stereos	Not specifi	20	KC78890	level	Alder Flyc	01-Mar	KC78890	0	yes	Empidonax	Murphy	FI	44.707	44100	Hz	song	280	m		196821	0	mp3	increas	Ing	78890	https://www.xeno-canto.org/sounds/78890	United	Sto	Jonathan	Poclie	af	93.319	0-36s	12-20	Jonathan Creative C		
102	4	no	aidfly	2	stereos	Not specifi	18	KC11407	both	Black-cap	04-Jun	KC11407	0	yes	Poclie	at	Environ		40.557	48000	Hz	call	1500	m	Three ind	48000	0	mp3	increas	Ing	114073	https://www.xeno-canto.org/sounds/114073	United	Sto	Eric	Defa	Poclie	af	106.02	0-36s	11-46	Eric Defa Creative C
103	4.5	no	aidfly	1	monos	Not specifi	83	KC11408	Not specifi	Black-cap	04-Jun	KC11408	0	yes	Poclie	at	Environ		44.942	44100	Hz	call	260	m	Call	128000	0	mp3	Not specifi	Amer	114086	https://www.xeno-canto.org/sounds/114086	United	Sto	Jonathan	Poclie	af	93.259	Not specifi	14-26	Jonathan Creative C	
104	4.5	no	aidfly	1	monos	Not specifi	33	KC12087	decreas	Black-cap	01-Jul	KC12087	0	yes	Poclie	at	St. Maney		47.0655	44100	Hz	call	380	m	One of th	128000	0	mp3	increas	Ing	120878	https://www.xeno-canto.org/sounds/120878	Canada	Martin	St	Poclie	af	71.415	0-36s	08-51	Martin St Creative C	
105	5	no	aidfly	1	monos	Not specifi	51	KC12106	Not specifi	Black-cap	01-Jul	KC12106	0	yes	Poclie	at	St. Maney		44.84	44100	Hz	call	220	m	Two ind	192000	0	mp3	Not specifi	Ing	121068	https://www.xeno-canto.org/sounds/121068	United	Sto	Jonathan	Poclie	af	93.216	Not specifi	09-30	Jonathan Creative C	
106	3.5	no	aidfly	2	stereos	Not specifi	59	KC12350	Not specifi	Black-cap	01-Mar	KC12350	0	yes	Poclie	at	St. Maney		45.1155	48000	Hz	call	260	m	Soft, high	192000	0	mp3	Not specifi	Amer	123506	https://www.xeno-canto.org/sounds/123506	United	Sto	Jonathan	Poclie	af	93.259	Not specifi	06-40	Jonathan Creative C	
107	5	no	aidfly	2	stereos	Not specifi	18	KC13195	level	Black-cap	01-Mar	KC13195	0	yes	Poclie	at	St. Maney		47.069	44100	Hz	call	0	m	One ind	128000	0	mp3	level	White	131952	https://www.xeno-canto.org/sounds/131952	Canada	Martin	St	Poclie	af	70.794	0-36s	08-26	Martin St Creative C	
108	4.5	no	aidfly	1	monos	Not specifi	335	KC13281	Not specifi	Black-cap	01-Mar	KC13281	0	yes	Poclie	at	St. Maney		47.069	44100	Hz	call	420	m	Several bi	192000	0	mp3	Not specifi	Ing	132813	https://www.xeno-canto.org/sounds/132813	United	Sto	Jonathan	Poclie	af	94.826	Not specifi	14-25	Jonathan Creative C	
109	4	no	aidfly	1	monos	Not specifi	131	KC13401	Not specifi	Black-cap	01-Mar	KC13401	0	yes	Poclie	at	St. Maney		41.92	44100	Hz	call	20	m	Vocalizat	128000	0	mp3	Not specifi	Ing	134017	https://www.xeno-canto.org/sounds/134017	United	Sto	Dan	Lane	Poclie	af	70.543	Not specifi	08-30	Dan Lane Creative C
110	3.5	no	aidfly	2	stereos	Not specifi	25	KC13547	level	Black-cap	01-Mar	KC13547	0	yes	Poclie	at	St. Maney		35.6109	44100	Hz	song	1500	m		128000	0	mp3	level	White	135477	https://www.xeno-canto.org/sounds/135477	United	Sto	Mike	Nels	Poclie	af	83.425	0-36s	09-30	Mike Nels Creative C
111	4	no	aidfly	2	stereos	Not specifi	8	KC13547	level	Black-cap	01-Mar	KC13547	0	yes	Poclie	at	St. Maney		35.6109	44100	Hz	song	1500	m		128000	0	mp3	level	White	135478	https://www.xeno-canto.org/sounds/135478	United	Sto	Mike	Nels	Poclie	af	83.425	0-36s	09-30	Mike Nels Creative C
112	4	no	aidfly	2	stereos	Not specifi	24	KC13905	both	Black-cap	01-Mar	KC13905	0	yes	Poclie	at	Environ		35.5717	44100	Hz	call	1700	m		128000	0	mp3	Not specifi	Ing	139053	https://www.xeno-canto.org/sounds/139053	United	Sto	Mike	Nels	Poclie	af	88.48	0-36s	07-30	Mike Nels Creative C
113	4	no	aidfly	2	stereos	Not																																				

To standardize the dataset and make it similar to the KWS dataset, two steps were taken:

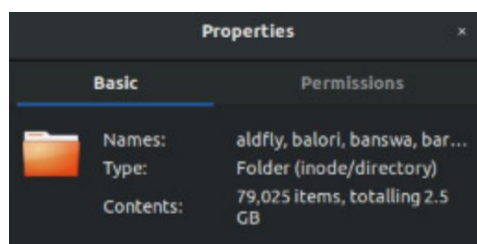
1. Convert mp3 to wav and resample all data to 16kHz (same as KWS dataset, and bird calls are typically between 1-8kHz in frequency so 16k is plenty): Refer code `convert.py` and `convert16k.py`

2. Segment the audio into 1 second clips: The recordings in the dataset vary from 1.2 seconds to 900 seconds in length, with anywhere from 1 to 100 relevant bird sounds within a recording. A processing script was written to first filter and then identify and extract useful 1 second long segments. Refer to code `bird_slicer.py` and `image slice_ops.png` to get an intuition on how this algorithm works.

Note: The code `convert16k.py` and `bird_slicer.py` can be found in folder `python-code.zip`



The outcome of this process is a dataset of 79025 unique samples (refer `data-properties.png`), with the following distribution. (refer `data-distribution.jpeg`). This data has not been manually cleaned after segmentation as the process is too time consuming and requires expert annotation

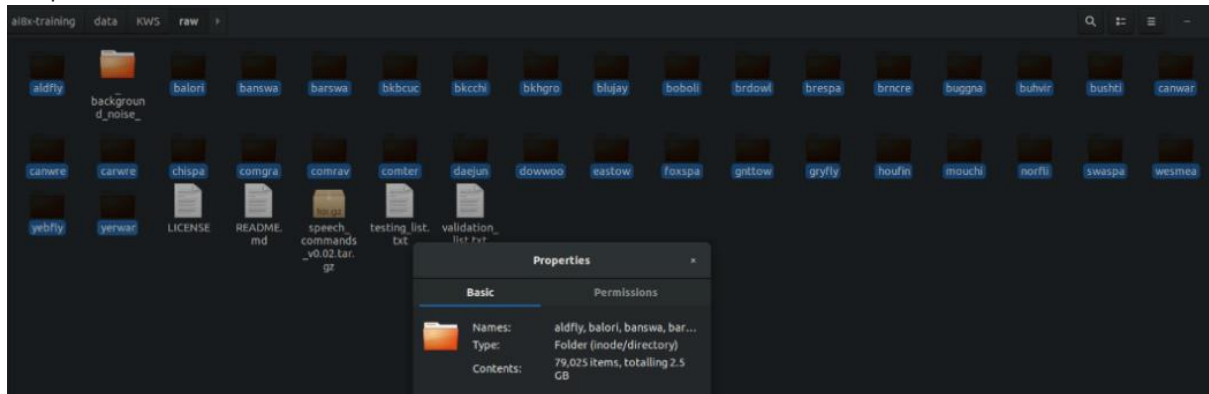


```
Warning: this process could take an hour!
data_len: 10384
..... Label Size .....
aldfly: 2207
balori: 1000
banswa: 3389
barswa: 3000
bkbuc: 607
bkcchl: 2180
bkhgro: 3062
blujay: 1423
boboll: 2400
brdowl: 1959
brespa: 3523
brncr: 1642
bupna: 3501
buhvir: 2917
bushti: 1970
camar: 1904
camres: 1771
carure: 2297
chispa: 2214
congra: 2290
convra: 3756
conter: 2372
daejun: 2338
downoo: 1971
eastov: 1741
foxspa: 2180
gnttow: 2420
gryfly: 2154
houlin: 2980
mouchi: 1868
norfll: 1684
swaspa: 1843
wesnea: 1679
yebfly: 1782
yernar: 1337
.....
Processing the label: aldfly, 1 of 10
```

Phase 3: Model training & C-code synthesis using ai8x-training & ai8x-synthesis libraries

Once the data was ready, only a few changes needed to be made to the training scripts to train on the new data, the steps were as follows:

- Update data folder with new data



- Remove data unzip operations from kws20.py (__extract_archive function)

```

234
235     def __extract_archive(self, from_path, # pylint: disable=no-self-use
236                             to_path=None, remove_finished=False):
237         #         if to_path is None:
238             #             to_path = os.path.dirname(from_path)
239
240         #         if from_path.endswith('.tar.gz'):
241             #             with tarfile.open(from_path, 'r:gz') as tar:
242                 #                 tar.extractall(path=to_path)
243
244         #         else:
245             #             raise ValueError("Extraction of {} not supported".format(from_path))
246
247         #         if remove_finished:
248             #             os.remove(from_path)
249         print("Not extracting")
250

```

- Update class arrays & dictionaries (multiple lines)

```

70     class_dict = {'aldfly': 0, 'balori': 1, 'banswa': 2, 'barswa': 3, 'bkbcuc': 4, 'bkccchi': 5,
71                  'bkhgro': 6, 'blujay': 7, 'boboli': 8, 'brdowl': 9, 'brespa': 10, 'brncre': 11,
72                  'buggna': 12, 'buhvir': 13, 'bushti': 14, 'canwar': 15, 'canwre': 16, 'carwre': 17,
73                  'chispa': 18, 'comgra': 19, 'comrav': 20, 'comter': 21, 'daejun': 22, 'dowwoo': 23,
74                  'eastow': 24, 'foxspa': 25, 'gnttow': 26, 'gryfly': 27, 'houfin': 28, 'mouchi': 29,
75                  'norfli': 30, 'swaspa': 31, 'wesmea': 32, 'yebfly': 33, 'yerwar': 34}
76
599     elif num_classes == 20:
600         classes = ['aldfly', 'banswa', 'barswa', 'bkccchi', 'bkhgro', 'boboli', 'brdowl', 'brespa', 'buggna', 'buhvir',
601                  'bushti', 'chispa', 'comgra', 'comrav', 'comter', 'daejun', 'gnttow', 'gryfly', 'houfin', 'swaspa']
602     elif num_classes == 35:
603
652     {
653         'name': 'KWS_20', # 20 keywords
654         'input': (128, 128, 1),
655         'output': (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20),
656         'weight': (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0.14),
657         'loader': KWS_20_get_datasets,
658     },

```

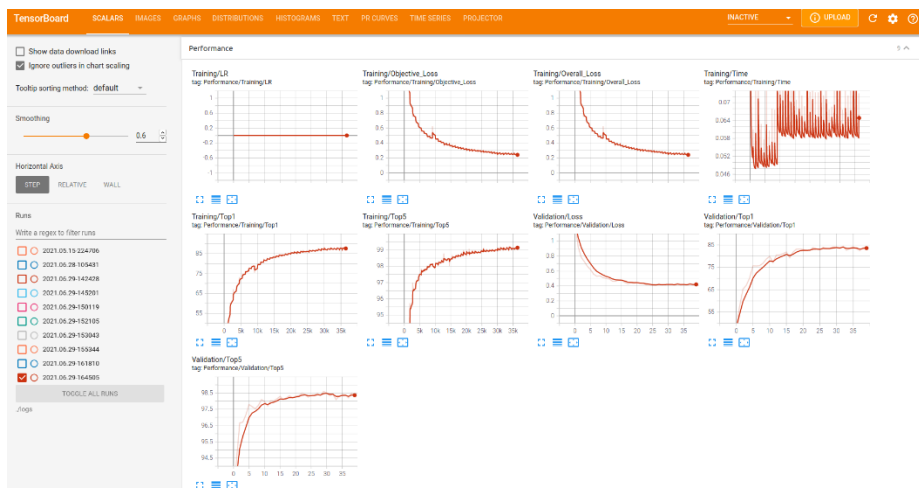
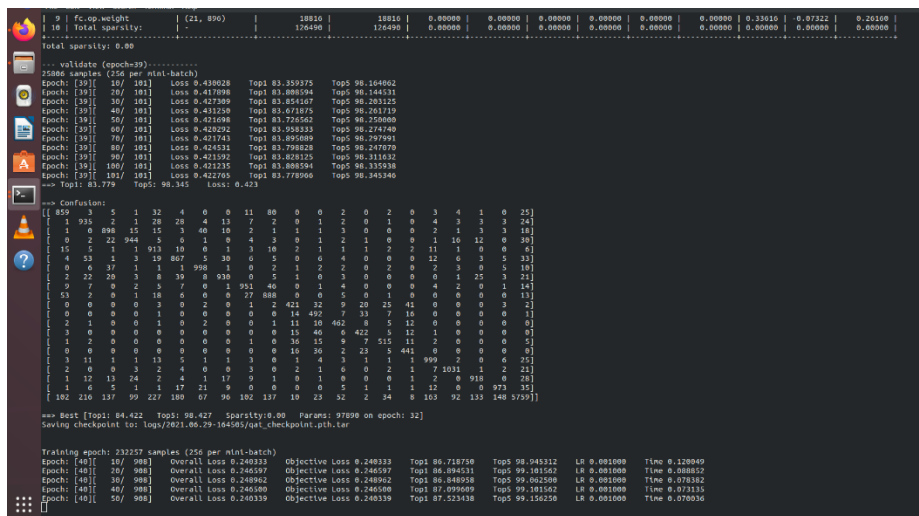
Run the script train.sh but reduce the number of epochs to 50.
The network: ai85kws20net that is found in ai85net-kws20.py was used.

We first tried to train a mix of birds and words to see how the network performs on the hybrid dataset.

Screenshots of the training process can be found in:

- Words&Birds-ep10-confusion
- Words&Birds-ep40-confusion
- Words&Birds-ep40-tensorboard

In the confusion matrix you can clearly identify the cluster of birds as they perform different from the words.



Satisfied with this PoC we trained the network on only birds.
We found that the network starts to overtrain after 50 epochs, so we stopped the training process then.

The following images document the training output:

- PNG files e1, e10, e20, e30, e40 and e50 show the progression of the confusion matrix during training

```
Activities Terminal Wed 12:28
mohan@mohan-TUF-FX505GCE: ~/ai/ix-training

File Edit View Search Terminal Help
Epoch: [9] 20/ 84 Loss 0.979634 Top1 59.511719 Top5 93.125000
Epoch: [9] 30/ 84 Loss 0.980195 Top1 59.674479 Top5 93.177003
Epoch: [9] 40/ 84 Loss 0.955198 Top1 60.029297 Top5 93.496094
Epoch: [9] 50/ 84 Loss 0.956701 Top1 59.953125 Top5 93.406250
Epoch: [9] 60/ 84 Loss 0.957902 Top1 60.071615 Top5 93.287760
Epoch: [9] 70/ 84 Loss 0.958440 Top1 60.122768 Top5 93.264589
Epoch: [9] 80/ 84 Loss 0.954308 Top1 60.048828 Top5 93.266602
Epoch: [9] 84/ 84 Loss 0.954295 Top1 59.953997 Top5 93.273248
==> Top1: 59.954 Top5: 93.273 Loss: 0.954

=== Confusion:
[[ 370 31 7 12 24 4 5 17 21 12 4 5 10 2 9 10 6 12 15 4 17]
 [ 3 736 20 11 2 2 12 6 1 1 4 10 0 23 5 3 7 0 5]
 [ 16 47 623 4 0 3 2 16 9 5 0 2 12 0 8 4 1 3 10 6 0]
 [ 8 17 6 494 1 1 7 7 16 2 11 10 7 8 7 13 3 3 1 4 12]
 [ 16 7 2 11 883 8 3 0 11 8 1 1 17 0 19 5 5 9 4 2 27]
 [ 20 4 40 14 27 288 2 16 25 3 0 7 33 0 22 11 25 24 41 5 30]
 [ 6 5 3 12 2 0 420 3 8 2 2 3 7 10 8 6 0 7 0 0 24]
 [ 11 19 4 10 1 1 6 714 5 0 3 45 1 3 4 46 15 4 16 4 14]
 [ 32 20 15 20 7 3 13 2 648 3 26 3 6 1 11 11 8 9 2 11 25]
 [ 24 14 6 15 7 0 6 2 9 647 10 0 9 1 17 2 9 1 12 2 40]
 [ 9 13 3 24 0 1 1 16 21 0 413 8 3 0 2 13 3 1 2 2 10]
 [ 5 8 5 20 0 1 1 29 2 2 14 418 2 2 0 30 3 2 1 3 7]
 [ 6 23 7 21 5 2 12 3 15 6 1 4 376 5 66 4 4 1 6 1 21]
 [ 0 8 0 2 0 1 3 0 0 0 0 0 1008 2 0 0 0 7 0 0 7]
 [ 8 21 7 7 2 5 1 2 7 1 0 1 25 7 561 0 1 3 4 0 10]
 [ 8 6 3 6 4 0 14 39 9 4 12 27 4 1 0 479 5 6 5 7 10]
 [ 39 21 6 11 11 8 18 80 7 6 3 16 10 3 3 27 260 20 45 0 40]
 [ 21 8 6 6 4 1 7 6 1 1 5 3 5 4 7 14 454 3 2 10]
 [ 29 15 17 10 10 3 2 8 5 11 2 0 5 0 7 3 15 10 608 0 48]
 [ 15 10 9 14 0 1 0 23 10 5 5 11 11 0 1 50 4 0 6 294 25]
 [ 427 254 178 599 405 115 120 257 225 316 141 47 240 105 264 264 167 210 445 141 2072]]

=== Best [Top1: 59.954 Top5: 93.273 Sparsity:0.00 Params: 97890 on epoch: 9]
Saving checkpoint to: logs/2021.06.30-114620/checkpoint.pth.tar

Training epoch: 191727 samples (256 per mini-batch)
Epoch: [10] 10/ 749 Overall Loss 1.065637 Objective Loss 1.065637 Top1 60.703125 Top5 93.945312 LR 0.001000 Time 0.110272
Epoch: [10] 20/ 749 Overall Loss 1.032516 Objective Loss 1.032516 Top1 58.437500 Top5 92.890625 LR 0.001000 Time 0.083720
Epoch: [10] 30/ 749 Overall Loss 1.022065 Objective Loss 1.022065 Top1 58.346354 Top5 92.877604 LR 0.001000 Time 0.074857
Epoch: [10] 40/ 749 Overall Loss 0.997872 Objective Loss 0.997872 Top1 58.300781 Top5 92.939453 LR 0.001000 Time 0.070511
Epoch: [10] 50/ 749 Overall Loss 0.992668 Objective Loss 0.992668 Top1 58.250000 Top5 92.937500 LR 0.001000 Time 0.068179
Epoch: [10] 60/ 749 Overall Loss 0.987534 Objective Loss 0.987534 Top1 58.235677 Top5 93.014323 LR 0.001000 Time 0.066407
Epoch: [10] 70/ 749 Overall Loss 0.981458 Objective Loss 0.981458 Top1 58.147321 Top5 92.991071 LR 0.001000 Time 0.065121
Epoch: [10] 80/ 749 Overall Loss 0.971540 Objective Loss 0.971540 Top1 58.505059 Top5 93.108945 LR 0.001000 Time 0.064311
Epoch: [10] 90/ 749 Overall Loss 0.972547 Objective Loss 0.972547 Top1 58.068576 Top5 93.098958 LR 0.001000 Time 0.063592
Epoch: [10] 100/ 749 Overall Loss 0.972555 Objective Loss 0.972555 Top1 58.109375 Top5 93.113281 LR 0.001000 Time 0.062961
Epoch: [10] 110/ 749 Overall Loss 0.972945 Objective Loss 0.972945 Top1 57.958097 Top5 93.004261 LR 0.001000 Time 0.062531
Epoch: [10] 120/ 749 Overall Loss 0.971213 Objective Loss 0.971213 Top1 58.011068 Top5 93.027344 LR 0.001000 Time 0.062083
Epoch: [10] 130/ 749 Overall Loss 0.970513 Objective Loss 0.970513 Top1 58.179087 Top5 92.996798 LR 0.001000 Time 0.061805
Epoch: [10] 140/ 749 Overall Loss 0.966716 Objective Loss 0.966716 Top1 58.127790 Top5 93.024554 LR 0.001000 Time 0.061463
Epoch: [10] 150/ 749 Overall Loss 0.962939 Objective Loss 0.962939 Top1 58.268229 Top5 93.046875 LR 0.001000 Time 0.061167

Total sparsity: 0.00

... validate (epoch=0)-----
21303 samples (256 per mini-batch)
Epoch: [0] 10/ 84 Loss 1.737101 Top1 34.062500 Top5 78.750000
Epoch: [0] 20/ 84 Loss 1.767183 Top1 34.101563 Top5 78.515625
Epoch: [0] 30/ 84 Loss 1.764736 Top1 34.583333 Top5 78.437500
Epoch: [0] 40/ 84 Loss 1.762329 Top1 34.501953 Top5 78.847656
Epoch: [0] 50/ 84 Loss 1.767567 Top1 34.359375 Top5 78.906250
Epoch: [0] 60/ 84 Loss 1.759341 Top1 34.485677 Top5 79.134115
Epoch: [0] 70/ 84 Loss 1.752764 Top1 34.726562 Top5 79.291295
Epoch: [0] 80/ 84 Loss 1.750233 Top1 34.951172 Top5 79.301758
Epoch: [0] 84/ 84 Loss 1.747855 Top1 34.957518 Top5 79.350326
==> Top1: 34.958 Top5: 79.350 Loss: 1.748

=== Confusion:
[[ 82 72 32 13 36 3 8 23 12 75 5 26 50 1 12 4 39 76 19 6 3]
 [ 3 589 42 18 10 4 2 23 3 24 6 3 31 1 18 3 24 14 15 8 6]
 [ 3 77 272 28 3 21 0 95 13 20 0 14 8 0 17 1 57 16 96 35 1]
 [ 0 60 16 263 3 6 15 5 21 46 56 17 38 7 12 0 22 22 10 13 6]
 [ 0 10 0 6 814 17 5 3 7 14 2 4 85 1 10 0 13 27 4 1 16]
 [ 4 32 26 15 72 187 1 11 9 7 0 10 41 0 54 1 46 42 50 2 27]
 [ 1 8 2 14 5 5 339 15 12 4 32 40 20 18 1 0 9 1 0 3 5]
 [ 1 183 51 9 1 5 6 430 8 0 20 198 6 5 9 11 21 9 19 21 1]
 [ 6 53 1 26 9 4 10 23 374 3 185 13 45 3 3 0 40 51 5 23 7]
 [ 10 49 9 15 21 2 0 6 7 543 19 9 59 1 5 1 14 16 35 3 17]
 [ 2 13 1 6 0 3 9 15 81 1 340 17 40 1 1 0 5 3 1 4 2]
 [ 2 11 3 8 0 1 9 44 10 0 29 398 14 0 1 6 6 0 0 7 0]
 [ 8 23 2 14 89 17 0 4 29 47 0 8 253 6 29 0 0 17 8 3 16]
 [ 0 2 0 1 0 0 1 0 0 0 1 1 14 1010 0 0 0 4 0 0 3]
 [ 4 31 36 16 44 71 1 5 1 31 3 1 96 5 215 0 18 25 56 2 12]
 [ 0 25 17 2 2 2 18 155 20 3 31 182 38 2 0 68 25 8 3 41 7]
 [ 5 64 30 25 11 17 15 66 10 34 14 23 27 1 3 6 189 26 37 15 18]
 [ 23 58 3 4 7 2 10 19 3 36 2 24 23 10 2 2 39 297 0 8 3]
 [ 11 50 104 15 23 47 0 9 1 44 0 4 29 0 29 2 36 31 359 4 10]
 [ 1 34 18 3 6 0 0 111 25 8 14 58 14 1 1 15 28 15 2 138 2]
 [ 46 552 273 365 679 391 110 333 214 718 422 255 687 151 196 53 371 327 311 251 287]]

=== Best [Top1: 34.958 Top5: 79.350 Sparsity:0.00 Params: 97890 on epoch: 0]
Saving checkpoint to: logs/2021.06.30-114620/checkpoint.pth.tar

Training epoch: 191727 samples (256 per mini-batch)
Epoch: [1] 10/ 749 Overall Loss 1.746780 Objective Loss 1.746780 Top1 35.625000 Top5 77.460938 LR 0.001000 Time 0.095479
Epoch: [1] 20/ 749 Overall Loss 1.731777 Objective Loss 1.731777 Top1 35.859375 Top5 79.101562 LR 0.001000 Time 0.071141
Epoch: [1] 30/ 749 Overall Loss 1.727034 Objective Loss 1.727034 Top1 35.716146 Top5 78.854167 LR 0.001000 Time 0.063049
Epoch: [1] 40/ 749 Overall Loss 1.734808 Objective Loss 1.734808 Top1 35.712891 Top5 79.280984 LR 0.001000 Time 0.058974
Epoch: [1] 50/ 749 Overall Loss 1.746245 Objective Loss 1.746245 Top1 35.515625 Top5 78.807188 LR 0.001000 Time 0.056552
Epoch: [1] 60/ 749 Overall Loss 1.742714 Objective Loss 1.742714 Top1 35.384115 Top5 78.880208 LR 0.001000 Time 0.054929
Epoch: [1] 70/ 749 Overall Loss 1.747758 Objective Loss 1.747758 Top1 35.262277 Top5 78.967634 LR 0.001000 Time 0.053786
Epoch: [1] 80/ 749 Overall Loss 1.747022 Objective Loss 1.747022 Top1 35.266600 Top5 78.896600 LR 0.001000 Time 0.052622
```


[illegible]

```

k kus_conv1.op.weight (64, 64, 3, 3) 36864 36864 0.00000 6.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.26043 -0.62409 0.15979
l kus_conv1.op.weight (64, 64, 3, 3) 17280 17280 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.21811 -0.01154 0.16274
k kus_conv3.op.weight (7, 30, 3, 3) 1890 1890 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.26161 0.08223 0.20512
l kus_conv3.op.weight (7, 30, 3, 3) 1890 1890 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.26161 0.08223 0.20512
f fc.op.weight (21, 896) 18816 18816 0.00000 6.00000 0.00000 0.00000 0.00000 0.00000 0.34600 -0.65307 0.26733
Total sparsity: 126498 126498 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000
Total sparsity: 0.00

--- validate (epoch=39)-----
21383 samples (256 per mini-batch)
Epoch: [39] 16/ 84 Loss: 0.826379 Top1 66.796875 Top5 94.140625
Epoch: [39] 20/ 84 Loss: 0.805040 Top1 67.148438 Top5 94.335938
Epoch: [39] 30/ 84 Loss: 0.815455 Top1 66.536458 Top5 94.375000
Epoch: [39] 40/ 84 Loss: 0.826454 Top1 66.960797 Top5 94.531250
Epoch: [39] 50/ 84 Loss: 0.822546 Top1 66.031250 Top5 94.648438
Epoch: [39] 60/ 84 Loss: 0.819229 Top1 66.362083 Top5 94.641927
Epoch: [39] 70/ 84 Loss: 0.817263 Top1 66.385804 Top5 94.715402
Epoch: [39] 80/ 84 Loss: 0.824640 Top1 66.388594 Top5 94.677344
Epoch: [39] 84/ 84 Loss: 0.821101 Top1 66.474206 Top5 94.704971
== Top1: 66.474 Top5: 94.705 Loss: 0.821

=== Confusion:
[[ 484 12 5 4 10 7 5 12 21 8 9 4 9 1 5 5 5 11 21 18 26]
 [ 723 40 6 27 2 10 7 6 3 1 0 8 7 1 6 2 12 6 7 0 7]
 [ 6 27 659 1 0 7 1 6 16 3 2 2 1 0 3 0 6 3 13 0 13]
 [ 19 12 7 472 0 2 4 11 26 9 21 0 6 3 7 9 8 2 18]
 [ 10 4 3 5 860 9 4 1 7 14 1 11 0 0 0 7 20 12 18 1 54]
 [ 8 8 30 1 7 447 0 5 15 2 4 0 12 1 7 0 19 19 23 9 18]
 [ 3 3 4 12 1 1 417 8 12 7 3 1 10 0 4 3 3 10 2 2 20]
 [ 13 13 4 0 9 4 802 4 0 6 3 1 7 20 11 19 7 7 19 7 19]
 [ 17 8 12 0 14 5 5 719 7 11 3 4 4 2 6 15 12 9 9 19]
 [ 22 7 5 3 5 5 1 2 11 674 9 0 9 0 1 1 6 4 26 3 47]
 [ 11 4 0 1 11 0 1 16 454 1 0 0 1 1 1 2 19 7 10 8]
 [ 2 9 2 10 0 8 1 43 4 0 0 10 421 5 0 0 11 6 2 1 12 8]
 [ 3 21 13 5 4 14 3 9 15 13 1 4 422 3 20 3 4 2 8 2 26]
 [ 4 0 3 1 2 11 2 2 0 0 0 1 996 1 0 0 4 0 0 11]
 [ 4 24 5 2 2 11 3 5 9 1 1 2 19 2 546 1 9 0 4 1 22]
 [ 5 4 2 3 2 3 7 50 3 5 14 13 0 1 462 19 5 2 8 17]
 [ 13 10 1 3 16 9 38 0 6 3 11 0 0 6 393 27 44 8 20 1]
 [ 8 6 1 4 2 6 5 9 6 1 4 1 0 1 6 14 14 475 8 4 9]
 [ 14 3 16 4 5 7 0 3 4 3 1 0 4 0 2 0 17 6 696 4 45]
 [ 6 1 8 10 9 2 9 0 9 1 0 0 1 0 7 3 9 379 0 0]
 [ 239 165 118 313 239 236 75 277 322 265 151 58 186 52 102 99 318 177 601 256 2743]]

== Best Top1: 69.033 Top5: 95.522 Sparsity:0.00 Params: 97890 on epoch: 37]
Saving checkpoint to: logs/2021.06.30-114620/gst_checkpoint.pth.tar

Training epoch: 191727 samples (256 per mini-batch)
Epoch: [40] 10/ 7491 Overall Loss: 0.432847 Objective Loss: 0.432847 Top1 76.193062 Top5 98.007812 LR 0.001000 Time 0.118579
Epoch: [40] 20/ 7491 Overall Loss: 0.432841 Objective Loss: 0.432841 Top1 76.191468 Top5 98.027184 LR 0.001000 Time 0.083827
Epoch: [40] 30/ 7491 Overall Loss: 0.432841 Objective Loss: 0.432841 Top1 76.191468 Top5 98.027184 LR 0.001000 Time 0.071902

```

- PNG files eval and test show the results of the evaluation script and test run after training

```
Number of elements in class swaspa: 594
Number of elements in class unknown: 7632
[ 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20]

Dataset sizes:
training=191727
validation=21303
test=24045

--- test -----
24045 samples (256 per mini-batch)
Test: [ 10/ 94] Loss 41.182057 Top1 58.789062 Top5 91.718750
Test: [ 20/ 94] Loss 40.234916 Top1 59.121094 Top5 92.089844
Test: [ 30/ 94] Loss 38.800972 Top1 59.440104 Top5 92.174479
Test: [ 40/ 94] Loss 38.811320 Top1 59.482422 Top5 92.343750
Test: [ 50/ 94] Loss 38.468304 Top1 59.648437 Top5 92.343750
Test: [ 60/ 94] Loss 38.457561 Top1 59.479167 Top5 92.304688
Test: [ 70/ 94] Loss 38.368277 Top1 59.553571 Top5 92.433036
Test: [ 80/ 94] Loss 38.270075 Top1 59.663086 Top5 92.392578
Test: [ 90/ 94] Loss 38.492158 Top1 59.496528 Top5 92.434896
Test: [ 94/ 94] Loss 38.688944 Top1 59.401123 Top5 92.426700
==> Top1: 59.401 Top5: 92.427 Loss: 38.689

==> Confusion:
[[ 414 20 5 10 6 2 2 52 16 8 5 8 8 1 11 14 10 24 27 11 24]
 [ 4 965 41 17 2 1 1 22 14 2 0 1 14 3 35 5 12 6 13 1 8]
 [ 10 29 656 10 2 2 3 23 6 0 2 5 14 0 6 3 7 3 11 4 17]
 [ 9 12 7 478 0 0 2 23 23 1 40 5 8 3 4 11 10 4 4 4 19]
 [ 24 6 0 6 1001 12 4 2 7 10 0 1 19 0 9 8 1 17 11 1 46]
 [ 21 11 72 20 29 338 1 25 18 9 1 20 37 4 20 9 33 27 46 7 74]
 [ 5 12 1 18 1 0 459 4 11 2 12 3 10 14 12 20 1 2 0 5 20]
 [ 21 14 3 23 1 5 5 829 2 0 12 46 0 4 0 48 27 11 13 10 15]
 [ 60 15 38 50 7 3 5 14 697 5 47 5 10 1 13 16 22 21 6 13 38]
 [ 30 11 14 18 6 1 2 6 705 2 0 3 0 10 0 8 4 16 3 58]
 [ 4 5 1 31 0 0 0 19 7 0 437 3 1 2 0 9 0 1 0 1 4]
 [ 7 12 5 22 0 3 6 56 4 0 11 480 0 0 0 21 1 1 0 9 7]
 [ 9 21 16 27 4 6 6 1 24 9 0 3 430 6 98 6 2 4 1 4 25]
 [ 1 2 0 6 0 0 6 5 0 0 1 3 1 1146 5 0 0 7 1 0 7]
 [ 7 13 15 8 2 0 2 1 5 0 0 0 22 3 599 2 3 1 4 1 14]
 [ 16 4 5 20 3 0 16 56 8 1 16 17 2 3 1 502 9 7 6 9 16]
 [ 47 15 13 17 6 7 8 120 12 7 5 14 8 0 12 28 293 15 40 4 31]
 [ 21 7 3 13 6 2 14 23 7 8 3 17 8 10 8 15 15 518 4 3 21]
 [ 29 13 16 14 5 4 0 16 10 6 0 0 4 2 10 3 18 11 686 2 39]
 [ 13 13 23 26 0 2 1 33 16 3 8 12 12 0 3 57 14 6 6 318 28]
 [ 507 220 215 661 338 144 137 364 206 290 181 52 253 93 289 306 179 209 518 160 2310]]

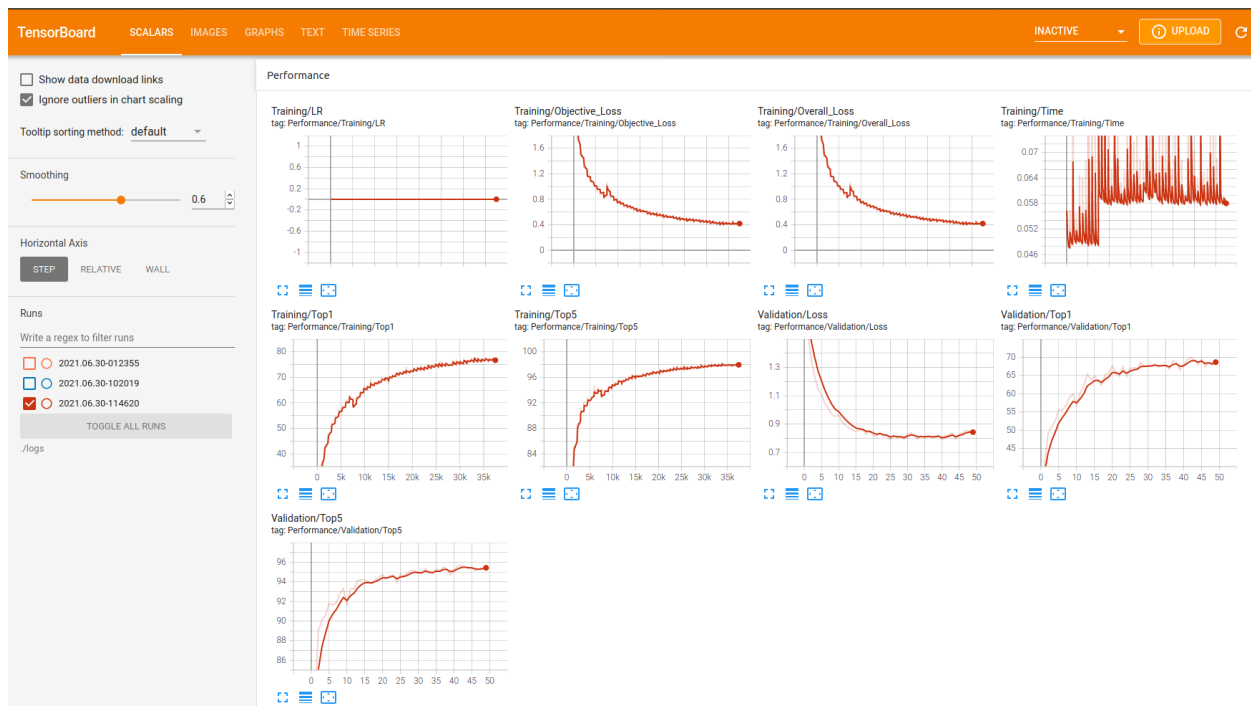
Log file for this run: /home/mohan/ai8x-training/logs/2021.06.30-130352/2021.06.30-130352.log
(AI ML) mohan@mohan-TUF-FX505GE:~/ai8x-training$
```

```
==> Best [Top1: 69.948 Top5: 95.691 Sparsity:0.00 Params: 97890 on epoch: 42]
Saving checkpoint to: logs/2021.06.30-114620/qat_checkpoint.pth.tar
--- test -----
24045 samples (256 per mini-batch)
Test: [ 10/ 94] Loss 0.908670 Top1 71.250000 Top5 94.765625
Test: [ 20/ 94] Loss 0.927943 Top1 69.140625 Top5 94.863281
Test: [ 30/ 94] Loss 0.944703 Top1 68.932292 Top5 94.830729
Test: [ 40/ 94] Loss 0.941977 Top1 68.984375 Top5 94.736328
Test: [ 50/ 94] Loss 0.935741 Top1 69.070125 Top5 94.812500
Test: [ 60/ 94] Loss 0.946427 Top1 68.847656 Top5 94.759115
Test: [ 70/ 94] Loss 0.940536 Top1 68.800223 Top5 94.765625
Test: [ 80/ 94] Loss 0.941018 Top1 68.750000 Top5 94.809570
Test: [ 90/ 94] Loss 0.940200 Top1 68.845486 Top5 94.782986
Test: [ 94/ 94] Loss 0.938936 Top1 68.800166 Top5 94.813891
==> Top1: 68.800 Top5: 94.814 Loss: 0.939

==> Confusion:
[[ 418 12 5 3 7 9 4 19 19 12 3 8 9 2 6 6 20 19 31 12 54]
 [ 3 954 34 11 1 4 1 20 25 5 0 3 12 4 38 1 14 9 8 1 19]
 [ 3 14 623 10 5 12 4 4 13 10 1 3 13 0 11 1 8 4 30 7 37]
 [ 7 5 3 479 0 8 5 15 32 4 17 7 5 5 3 7 13 0 5 10 36]
 [ 13 1 0 1 1021 17 5 0 1 11 0 0 8 2 8 1 4 13 13 1 65]
 [ 7 7 25 9 23 498 2 10 12 7 1 10 8 2 17 3 37 17 32 10 85]
 [ 0 4 0 5 2 0 504 0 7 5 2 2 7 20 11 6 6 1 4 3 23]
 [ 10 7 3 13 1 8 1 849 9 0 7 40 1 5 5 26 58 9 12 3 22]
 [ 17 8 8 31 4 17 3 8 793 8 21 5 5 0 13 9 34 10 16 8 68]
 [ 2 3 6 9 8 2 2 0 3 764 3 0 2 0 1 1 3 4 28 4 58]
 [ 0 1 0 23 1 2 7 11 18 10 420 6 1 5 2 1 3 1 1 1 11]
 [ 4 5 2 15 0 2 3 31 5 0 4 537 3 3 0 4 9 2 1 4 11]
 [ 6 11 15 12 13 12 1 17 10 0 2 444 1 67 0 3 3 10 3 59]
 [ 2 1 0 5 0 0 12 3 0 0 0 0 1 1152 6 0 0 4 0 5]
 [ 4 4 2 2 4 5 3 0 3 0 0 1 10 1 631 2 1 0 7 0 19]
 [ 8 4 3 7 15 4 14 35 7 1 9 40 1 4 5 482 23 4 4 11 36]
 [ 9 15 5 8 9 7 4 50 17 4 5 21 2 1 7 9 412 16 35 5 61]
 [ 1 3 1 6 7 9 12 8 7 1 1 14 2 14 9 3 15 565 9 2 37]
 [ 0 5 4 7 2 5 0 3 1 10 0 0 2 4 6 1 18 12 748 2 58]
 [ 7 2 15 17 0 4 1 12 23 3 6 13 6 1 8 15 13 11 10 368 59]
 [ 172 113 107 358 256 172 106 178 219 317 60 87 124 87 240 119 281 146 474 143 3873]]

Log file for this run: /home/mohan/ai8x-training/logs/2021.06.30-114620/2021.06.30-114620.log
(AI ML) mohan@mohan-TUF-FX505GE:~/ai8x-training$
```

- tensorboard-plots.png shows the output of the tensorboard console



The model was then synthesized using the ai8x-synthesis library, following standard synthesis procedure. Synthesized code can be found in Synthesized-C-Code.rar

Phase 4: Testing on Hardware using audio samples from the test set.

The code was then flashed onto the MAX78000 feather board using the KWS20 codebase as a reference. A demo of the same can be found in the video.

Test Setup

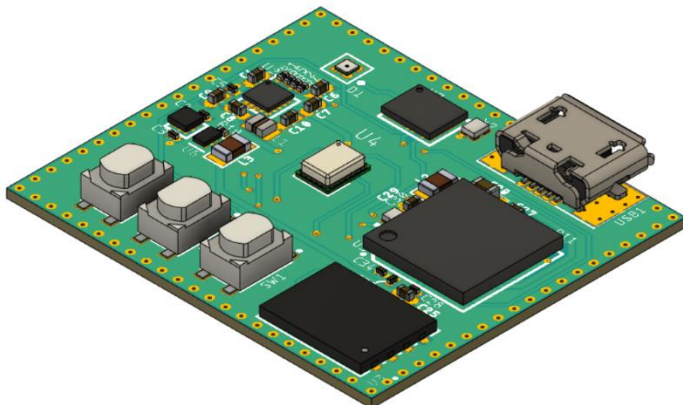


Please download the zip file brd20_demo.zip (based on kws_20) to be able to run the project on your own board. We have also included the test samples that were used in the video demo. No additional hardware is required to test this code as we have used the on board microphone.

Hardware Design

We have built a custom hardware based off of the MAX78000FTHR schematic. The major changes made to the design are

1. Removal of all video related hardware.
2. Addition of a 32MB flash memory to store collected audio data and improve detection accuracy over time.
3. Cost-effective BoM and PCB design to facilitate easy manufacturing and scale-up.



Future Scope:

- Clean dataset: The generated dataset is purely code based with no manual cleaning. As the xeno-canto data can sometimes be of questionable quality with a lot of background noise, manual cleaning and labelling can greatly improve accuracy.
- Enhance dataset: More birds, more samples, improve 'unknown' class by adding human and non-bird animal sounds as well.
- Larger input size: Increasing the input length from 1s to 2.5s is shown to improve accuracy.
- Neural Architecture Search: Did not play around with this much due to resource constraints. Hope to do this next
- Improve hardware: No additional hardware was required for this project as the on-board mic was sufficient. However it is reasonable to believe that a better mic can lead to better results. Since this a remote sensing solution, a battery pack and GSM module can also add value.

Other Uploads:

I have uploaded all the source code that I wrote and modified for this project and will be quite excited to work with someone who wants to build on this work. Unfortunately the dataset was too large to upload but I'm happy to share the same.

The demo video also walks through some of the execution steps and code modifications made.

Acknowledgements:

A big thank you to Elektor for hosting and Maxim for sponsoring this competition, it has been a great learning experience.

Many thanks to my friend & colleague Mr. Mohan Prabhakar for helping me by out training all these models on his gaming laptop (16GB Ram, i7-8th gen, GTX1050-Ti). He also helped me test the model on the MAX78000FTHR board and make the demo video.

References:

- [1] Stefan Kahl, Connor M. Wood, Maximilian Eibl, Holger Klinck, BirdNET: A deep learning solution for avian diversity monitoring, Ecological Informatics, Volume 61,2021,101236, ISSN 1574-9541, <https://doi.org/10.1016/j.ecoinf.2021.101236>. (<https://www.sciencedirect.com/science/article/pii/S1574954121000273>)
- [2] Audio Based Bird Species Identification using Deep Learning Techniques, ETH Zurich, <http://ceur-ws.org/Vol-1609/16090547.pdf>
- [3] KEYWORDS SPOTTING USING THE MAX78000 (AN7359) <https://www.maximintegrated.com/en/design/technical-documents/app-notes/7/7359.html>