Vocalization Detection

Team: Dat Bui, Amr Aboelnaga
Client: Dr. Michael Bowers
Course: CS4624 Multimedia, Hypertext and Information Access Capstone
Instructor: Dr. Edward Fox
Virginia Tech, Blacksburg 24061
4/29/20

Outline

- Technology used
- Timeline
- Overview
- Phases of Work
- Progress
- Deliverables
- Roles
- Future Work
- Lessons Learned
- Acknowledgements / References

Technology used

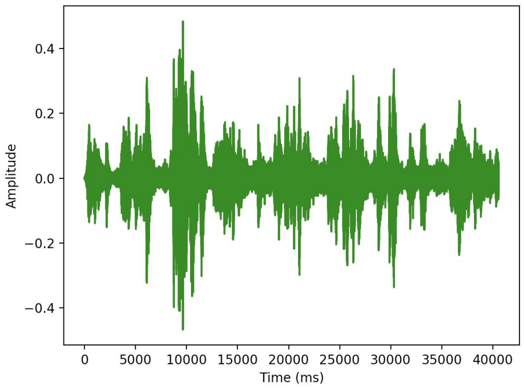
- Github
- Kaggle
- MATLAB
- PyWinAuto
- Ghidra
- Python
- NumPy
- Pandas

Timeline

- 2/3 - Milestone: Receive kaggle data
- 2/17 - Milestone: Process data
- 4/4 - Interim Report
- 4/7 - Milestone: Model refined
- 4/13 - Begin Data Matching
- 4/17 - Milestone: Application is developed and ready to use
- 4/26 - VTechWorks

Overview

- Goal is to facilitate Dr. Bower's research
- Attempts have been made to automate AviSoft
  - Automation with Python was not practical
  - Reverse engineering proved to be nearly impossible
- DeepSqueak
  - Transforming wav files into spectrograms for image processing
  - Detection
  - Classifications
    - Clustering
    - K-means clustering
    - Supervised classification
    - Wright Classifier Network



Phases of Work

- Pipelining and integration of DeepSqueak data with AviSoft
  - At the beginning
- Automation of AviSoft
  - Did not work
- Data matching between DeepSqueak and AviSoft
  - Current stage

Phases of Work

- Pipelining and Integration of AviSoft and DeepSqueak
  - Original plan
  - AWS, PostgresSQL, Python
  - Was not what Amr had in mind

Phases of Work

- Automation of AviSoft
  - Attempts were made to automate functionality of AviSoft
  - PyWinAuto
    - Would have been counterintuitive
  - Ghidra
    - Reverse engineering tool
    - Disassembly of source code
    - Not practical

Phases of Work

- Data matching
  - Python script that takes in AviSoft file and DeepSqueak classifications file
  - Syncs up classifications between the two
  - Checks for correlations between AviSoft data and DeepSqueak data

AviSoft data

#	duration	interval	peak freq	peak amp	fundamer	min freq(s	max freq(	bandw(sta	entropy(s	peak freq	peak amp	fundamer	min freq(e	max freq(	bandw(en	entropy(e	peak freq	peak amp	fundamer	min freq(c	max freq(	bandw(ce	entropy(c	peak freq	peak amp	fundamer	min freq(r	max fi
mea	0.01009	0.09346	14382	-113.37	10300.4	2792.09	84052	81171.6	0.17796	13854.6	-117.22	8245.06	1289.25	92986.6	91602.6	0.17345	12857.3	-122.74	9422.05	2340.82	88389.2	85957	0.15529	15311.2	-101.85	14016.4	8515.3	5235
std	0.01543	0.18115	8526.7	9.30201	8756.6	5895.15	31048.7	35865.4	0.03518	8730.91	7.70777	8823.35	3743.47	21202.5	24409.9	0.03291	10384	25.3153	9771.96	6134.9	26914.1	31606.6	0.06743	9818.98	10.468	9587.86	9581.28	3933
cnt	3985	3984	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3985	3
1	0.0012		12500	-124.98	0	300	99200	98800	0.167	13200	-118.96	13200	300	99200	98800	0.159	12800	-118.96	12800	300	99200	98800	0.161	12800	-118.96	12800	300	99
2	0.0012	0.7644	12800	-118.96	12800	300	99200	98800	0.148	13200	-124.98	0	300	99200	98800	0.126	12800	-124.98	0	300	99200	98800	0.141	12800	-118.96	12800	300	99
3	0.0012	0.2787	11700	-108.08	11700	300	99200	98800	0.187	10900	-124.98	0	300	99200	98800	0.154	11300	-124.98	0	300	99200	98800	0.126	11700	-108.08	11700	300	99
4	0.0012	0.1481	12500	-115.44	12500	300	99200	98800	0.225	10900	-124.98	0	300	99200	98800	0.154	10900	-118.96	10900	300	99200	98800	0.173	12500	-115.44	12500	300	99
5	0.0016	0.2438	97200	-108.08	0	300	99200	98800	0.156	96800	-118.96	0	300	99200	98800	0.159	96800	-124.98	0	300	99200	98800	0.063	97200	-108.08	0	300	99
6	0.0025	0.2976	12800	-112.94	12800	300	99200	98800	0.202	12800	-124.98	0	300	99200	98800	0.26	13200	-99.4	13200	11700	16000	4200	0.172	13200	-99.4	13200	11700	16
7	0.0227	0.016	13200	-105.89	13200	300	99200	98800	0.189	12500	-124.98	0	300	99200	98800	0.089	11700	-109.42	11700	300	99200	98800	0.163	11700	-101.46	11700	10100	37
8	0.0012	0.0931	11300	-115.44	11300	300	99200	98800	0.155	11700	-111	11700	300	99200	98800	0.166	10900	-118.96	10900	300	99200	98800	0.173	11300	-115.44	11300	300	99
9	0.0073	0.0896	13200	-115.44	13200	300	99200	98800	0.167	12100	-106.92	12100	300	99200	98800	0.172	13200	-115.44	13200	300	99200	98800	0.171	11700	-99.4	11700	10500	27
10	0.0019	0.0633	12500	-124.98	0	300	99200	98800	0.199	12800	-115.44	12800	300	99200	98800	0.167	12800	-112.94	12800	300	99200	98800	0.165	12800	-111	12800	300	99
11	0.0022	0.0163	13200	-95.44	13200	12100	16000	3900	0.173	11300	-124.98	0	300	99200	98800	0.141	12800	-112.94	12800	300	99200	98800	0.165	13200	-95.44	13200	12100	16
12	0.0121	0.0131	12800	-106.92	12800	300	99200	98800	0.172	12500	-105.89	12500	300	99200	98800	0.183	0	-184.98	0	300	99200	98800	0	12800	-104.15	12800	11700	16
13	0.0182	0.0662	12800	-106.92	12800	300	99200	98800	0.172	12800	-124.98	0	300	99200	98800	0.154	12800	-112.94	12800	300	99200	98800	0.169	13200	-91.72	13200	12100	16
14	0.0016	0.0326	12800	-124.98	0	300	99200	98800	0.199	12500	-118.96	12500	300	99200	98800	0.159	12500	-124.98	0	300	99200	98800	0.089	12800	-118.96	12800	300	99
15	0.0012	0.0348	18700	-118.96	18700	300	99200	98800	0.171	18300	-115.44	18300	300	99200	98800	0.168	18700	-115.44	18700	300	99200	98800	0.182	18700	-115.44	18700	300	99
16	0.0022	0.0348	11300	-124.98	0	300	99200	98800	0.154	11300	-124.98	0	300	99200	98800	0.227	11700	-104.98	11700	10500	16400	5800	0.18	11700	-104.98	11700	10500	16
17	0.0019	0.0208	13600	-108.08	13600	300	99200	98800	0.19	13600	-112.94	13600	300	99200	98800	0.173	13200	-124.98	0	300	99200	98800	0.141	13600	-108.08	13600	300	99
18	0.0288	0.0348	12500	-108.08	12500	300	99200	98800	0.169	12800	-118.96	12800	300	99200	98800	0.173	0	-184.98	0	300	99200	98800	0	13200	-89.13	13200	11700	15
19	0.0422	0.055	12500	-118.96	12500	300	99200	98800	0.171	10900	-118.96	10900	300	99200	98800	0.159	10900	-115.44	10900	300	99200	98800	0.229	13200	-96.04	13200	12100	16
20	0.0012	0.0752	13200	-115.44	14500	300	99200	98800	0.23	11300	-124.98	0	300	99200	98800	0.167	12500	-124.98	0	300	99200	98800	0.154	13200	-115.44	14500	300	99
21	0.0028	0.1091	13200	-108.08	13200	300	99200	98800	0.183	12800	-115.44	12800	300	99200	98800	0.167	12800	-111	12800	300	99200	98800	0.166	13200	-103.39	13200	12100	16
22	0.0012	0.0201	13200	-124.98	0	300	99200	98800	0.126	13200	-108.08	13200	300	99200	98800	0.171	13200	-124.98	0	300	99200	98800	0.141	13200	-115.44	13200	300	99
23	0.0022	0.0275	12800	-118.96	12800	300	99200	98800	0.161	13200	-124.98	0	300	99200	98800	0.167	12800	-115.44	12800	300	99200	98800	0.16	13200	-112.94	13200	300	99
24	0.0035	0.1328	12500	-118.96	12500	300	99200	98800	0.161	12100	-124.98	0	300	99200	98800	0.154	12100	-118.96	12100	300	99200	98800	0.164	12500	-104.98	12500	11300	15
25	0.0012	0.0156	13200	-101.46	13200	12100	16400	4200	0.174	12800	-115.44	12800	300	99200	98800	0.167	12800	-108.08	12800	300	99200	98800	0.181	13200	-101.46	13200	12100	16
26	0.0022	0.0428	12800	-94.88	12800	11700	19100	7400	0.189	12500	-118.96	12500	300	99200	98800	0.214	12500	-118.96	12500	300	99200	98800	0.161	12800	-94.88	12800	11700	19
27	0.0032	0.0297	12100	-104.15	12100	10500	14800	4200	0.172	10900	-124.98	0	300	99200	98800	0.154	12100	-103.39	12100	10900	14800	3900	0.17	12100	-99.87	12100	10900	14
28	0.0028	0.0544	12800	-94.88	12800	11700	15600	3900	0.175	11700	-124.98	0	300	99200	98800	0.126	12800	-102.7	12800	11700	16000	4200	0.172	13200	-86.92	13200	11700	15
29	0.0019	0.0217	13200	-115.44	13200	300	99200	98800	0.17	11700	-112.94	10800	300	99200	98800	0.23	11300	-124.98	0	300	99200	98800	0.109	11700	-109.42	11700	300	99
30	0.0012	0.0144	10900	-124.98	0	300	99200	98800	0.141	10900	-124.98	0	300	99200	98800	0.154	11300	-124.98	0	300	99200	98800	0.126	10900	-124.98	0	300	99
31	0.0016	0.4246	11300	-124.98	0	300	99200	98800	0.109	10500	-124.98	0	300	99200	98800	0.167	10900	-124.98	0	300	99200	98800	0.154	11300	-124.98	0	300	99
32	0.0016	0.0144	11300	-124.98	0	300	99200	98800	0.154	11700	-118.96	11700	300	99200	98800	0.159	11300	-124.98	0	300	99200	98800	0.154	11700	-112.94	11700	300	99

DeepSqueak Output

	Label	Accepted	Score	Begin Time (s)	End Time (s)	Call Length (s)	Principal Frequency (kHz)	Low Freq (kHz)	High Freq (kHz)	Delta Freq (kHz)	Frequency Standard Deviation (kHz)	Slope (kHz/s)	Sinuosity	Mean Power (dB/Hz)	Tonality
1	Complex Trill	TRUE	0.372341454	11.398224	11.998224	0.6	18.04992006	18.04992006	18.04992006	0	0	0	1	65535	
2	Complex Trill	TRUE	0.372341454	27.366548	27.996548	0.63	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
3	Complex Trill	TRUE	0.9400440454	68.78697221	68.82217221	0.0352	64.07896971	57.63997744	66.82089419	9.180916756	3.129617728	133.4916442	1.448867017	-60.25709663	0.3471004786
4	Complex Trill	TRUE	0.372341454	130.32762	130.99762	0.67	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
5	Complex Trill	TRUE	0.372341454	160.015204	160.685204	0.67	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
6	Complex Trill	TRUE	0.372341454	168.019364	168.689364	0.67	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
7	Complex Trill	TRUE	0.372341454	171.969864	171.979864	0.01	18.2117207	18.2117207	18.2117207	0	0	0	1	65535	
8	Complex Trill	TRUE	0.372341454	173.024464	173.674464	0.65	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
9	Complex Trill	TRUE	0.372341454	176.023524	176.703524	0.68	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
0	Complex Trill	TRUE	0.372341454	176.954964	176.984964	0.03	18.2117207	18.2117207	18.2117207	0	0	0	1	65535	
1	Complex Trill	TRUE	0.372341454	178.019564	178.699564	0.68	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
2	Complex Trill	TRUE	0.372341454	184.017688	184.707688	0.69	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
3	Complex Trill	TRUE	0.372341454	186.023728	186.713728	0.69	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
4	Complex Trill	TRUE	0.372341454	190.015808	190.685808	0.67	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
5	Complex Trill	TRUE	0.372341454	199.022988	199.692988	0.67	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
6	Short	TRUE	0.6029146314	201.0664282	201.0704282	0.004	62.3146004	61.48334521	62.50014844	1.01680323	0.3575388996	254.5381676	1.026561312	-61.48384984	0.3124884095
7	Complex Trill	TRUE	0.372341454	203.015068	203.675068	0.66	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
8	Complex Trill	TRUE	0.372341454	206.024128	206.684128	0.66	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
9	Complex Trill	TRUE	0.372341454	214.018288	214.648288	0.63	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
0	Complex Trill	TRUE	0.372341454	216.024328	216.654328	0.63	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
1	Upward Ramp	TRUE	0.5572182536	230.4912117	230.5048117	0.0136	60.16669635	59.20184657	61.87373827	2.671891699	0.6794355162	137.9032494	1.087907012	-59.161353	0.3572105191
2	Complex Trill	TRUE	0.372341454	235.021712	235.641712	0.62	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
3	Complex	TRUE	0.4376208484	238.2625726	238.2877726	0.0252	65.65818082	63.14715566	67.16941385	4.022258182	1.20830709	76.03595235	1.135349788	-61.95113741	0.3254695543
4	Complex Trill	TRUE	0.372341454	247.017952	247.627952	0.61	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
5	Complex Trill	TRUE	0.372341454	250.017016	250.627016	0.61	17.99992006	17.99992006	17.99992006	0	0	0	1	65535	
6	Upward Ramp	TRUE	0.2947298586	286.3159384	286.3327384	0.0168	64.61817157	61.48832464	67.06142984	5.573105196	1.742282949	342.5759942	1.083234081	-62.95889878	0.2837938213
7	Short	TRUE	0.3929832578	290.0392189	290.0524189	0.0132	64.72242014	62.00912693	67.59470372	5.58557679	1.643883147	331.913059	1.057307651	-62.21832064	0.3057329997

Progress - Detection

- Faster R-CNN
- State of the art detection model
- We only want to detect Ultrasonic Vocalizations

Progress - Classifications

Wright Classifier Network

- Accuracy up to 75% from a baseline of ~30%
- Training with Dr. Bowers' data
- Data augmentation

Roles

- Dat Bui
  - Automation
  - Learning DeepSqueak MATLAB library
  - Data matching
  - Note-taking
- Amr Aboelnaga
  - Development and improvement of Machine Learning model
  - Graduate Student
- Mike Bowers, PhD
  - Collection and categorization of vocalization data from rats

Future Work

- Improved Classifier Network
- Standalone application

Lessons Learned

- Communication
- COVID-19
- Linear Progression

Acknowledgement

- Client
  - Dr. Mike Bowers
- Graduate Student Correspondent
  - Amr Aboelnaga
- Instructor
  - Edward Fox
- DeepSqueak
  - https://www.researchgate.net/publication/330143833_DeepSqueak_a_deep_learning-based_system_for_detection_and_analysis_of_ultrasonic_vocalizations