

# What is my Dog trying to tell me?

## The Automatic Recognition of the Context and Perceived Emotion of Dog Barks

Simone Hantke<sup>1,2</sup>, Nicholas Cummins<sup>1</sup>, Björn Schuller<sup>1,3</sup>

<sup>1</sup> ZD.B Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg, Germany

<sup>2</sup> Machine Intelligence & Signal Processing Group, Technische Universität München, Germany

<sup>3</sup> GLAM - Group on Language, Audio & Music, Imperial College London, U.K.

### RESEARCH QUESTION

Can **acoustic feature representations**, designed and developed for **human-based affective computing** purposes, be used to recognise context and emotions in dog barks?



### TASKS

- (i) Classifying **context**
- (ii) Classifying perceived **emotion**
- (iii) Predicting perceived emotional **intensity**

### EXPERIMENTS

#### Features:

- *eGeMAPS* (88 dimensional features)
- *COMPARE* (6373 features):
  1. Prosodic features only
  2. Spectral and cepstral features only
- *Bag-of-Audio-Words* (BoAW) representations via *openXBOW* (6 different LLDs feature sets)

#### Setup:

- Features extracted using *openSMILE*
- Linear Support Vector Machine (*SVM*)
  - Prediction: Support Vector Regression (*SVR*)
- **Cost parameter (C) tuned** separately
  - Grid search to find **optimal C and epsilon**

#### Evaluation:

- **Leave-one-dog-out cross fold validation**
- Evaluation prediction tests: Root Mean Square Error (**RMSE**)
- Results: Unweighed Average Recall (**UAR**)

### CONCLUSIONS

**Acoustic feature** sets to capture human emotions can be used to classify **content, emotion and intensity** of dog barks accordingly.

→ Mammalian emotional changes produce similar acoustics effects across dogs and human.

#### Future Work:

- Testing alternative **feature representations**
- More advanced machine learning approaches, e.g. **LSTM based RNNs** and **transfer learning**

### DATA

#### Dataset:

- Emotional Dog Corpus (EmoDog)
- **12 Mudi dogs**, 226 bark sequences
- **7 situations**: Alone, Ball, Fight, Food, Play, Stranger, Walk

#### Annotations:

- Collected via crowdsourcing platform *iHEARu-PLAY*
- **6 annotators**
- **Emotion**: Aggression, Fear, Despair, Fun, or Happiness
  - 5-point likert scale
- **Content**: Alone, Ball, Fight, Food, Play, Stranger, or Walk
  - Classes

### RESULTS

#### Perceived emotional intensity:

RMSE	Agg.	Des.	Fear	Fun	Hap.
<i>eGeMAPS</i>	.888	.832	.891	<b>.876</b>	<b>.882</b>
<i>COMPARE</i>	.885	.834	.890	.928	.920
<i>COMPARE Pros.</i>	.855	.837	.902	.954	.932
<i>COMPARE Spec.</i>	.891	.833	.896	.937	.922
<i>BoAW MFCC</i>	<b>.847</b>	.814	.907	.925	.922
<i>BoAW MFCC + deltas</i>	.869	.819	.905	.931	.924
<i>BoAW eGeMAPS LLDs.</i>	.910	.853	.937	.957	.929
<i>BoAW COMPARE LLDs.</i>	.987	.818	.927	.982	.944
<i>BoAW COMPARE Pros. LLDs.</i>	.918	<b>.775</b>	<b>.877</b>	.977	.922
<i>BoAW COMPARE Spec. LLDs.</i>	.988	.815	.927	.982	.944

Scores ranged 1-5

#### Content and perceived emotion:

UAR [%]	Context	Emotion
<i>eGeMAPS</i>	24.4	<b>28.5</b>
<i>COMPARE</i>	31.3	25.7
<i>COMPARE Pros.</i>	30.0	27.5
<i>COMPARE Spec.</i>	<b>32.9</b>	25.8
<i>BoAW MFCC</i>	19.2	24.4
<i>BoAW MFCC + deltas</i>	21.1	23.8
<i>BoAW eGeMAPS</i>	19.1	25.1
<i>BoAW COMPARE</i>	16.7	21.9
<i>BoAW COMPARE Pros.</i>	15.9	22.9
<i>BoAW COMPARE Spec.</i>	16.6	21.8
<b>Human Performance</b>	23.7	--
<b>Chance</b>	14.3	20.0