



Building parsimonious SVM models for chewing detection and adapting them to the user



Iason Karakostas
Vasileios Papapanagiotou
Anastasios Delopoulos

Multimedia Understanding Group
Information Processing Laboratory
Dpt. Electrical and Computer Engineering
Aristotle University of Thessaloniki
Greece



Introduction

- Automatically monitoring eating activity has received significant attention in the research community
- Most of the proposed systems require proprietary/specialized sensors

We propose

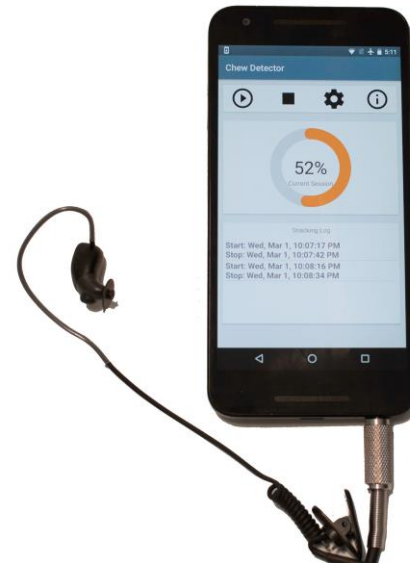
- A chewing detection system that captures audio from a commercial bone conduction microphone
- A method to build **efficient** and **effective** SVM models
- A method to **adapt** the SVM model to user requiring minimal user feedback



Chewing detection hardware



Commercial bone-conduction
microphone

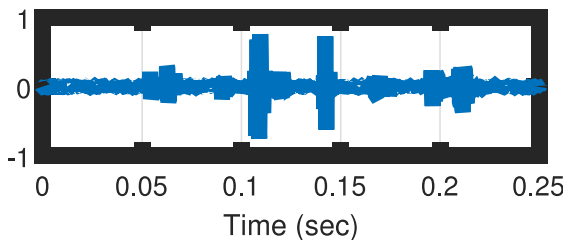


Android smart-phone

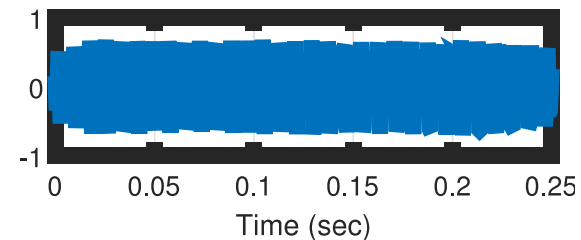


Audio signal pre-processing and feature extraction

- Sampling at 4 kHz
- High pass FIR cut-off frequency at 20 Hz
- Hamming filter of 3.72 sec
- Overlapping windows
 - step = 160 samples,
 - time domain features windows (length = 400 samples)
 - spectral features windows (length = 800 samples)



A single chew

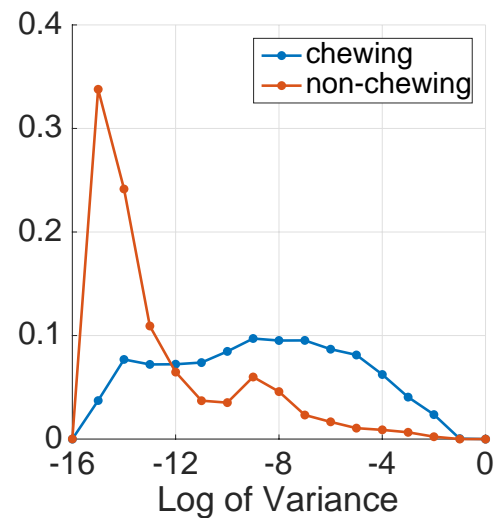
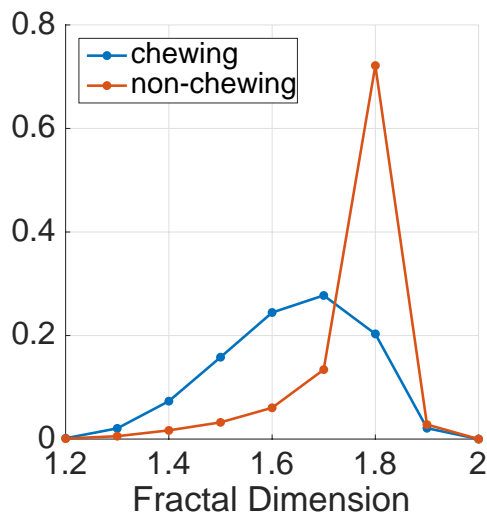


Voice



Extracted features

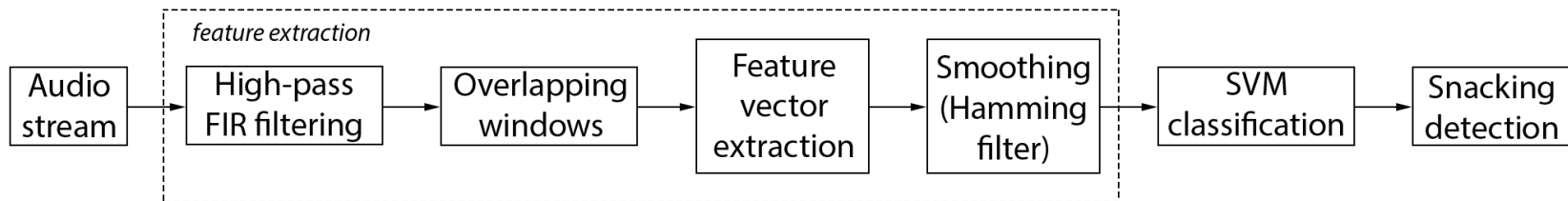
- 16 features
 - 7 time domain and 9 spectral features
- Fractal dimension and log of variance are quite discriminative individually





Classification and Active Learning

- Classification of feature vectors using the SVM classifier



- Active Learning

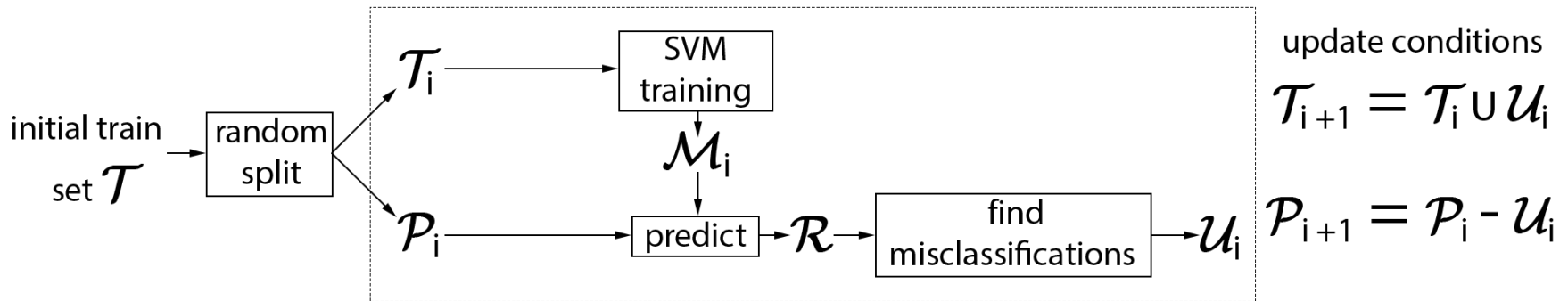
- A method of improving a classifier's effectiveness by enhancing the training set in "rounds"
 - Apply model on pool of feature vectors
 - Select feature vectors and request feedback (correct label)
- Active learning is used for:
 - Parsimonious Active Learning Training (PALT)
 - Inter-Active Learning Adaptation (IALA)



Parsimonious Active Learning Training (PALT)

A method to create SVM models with

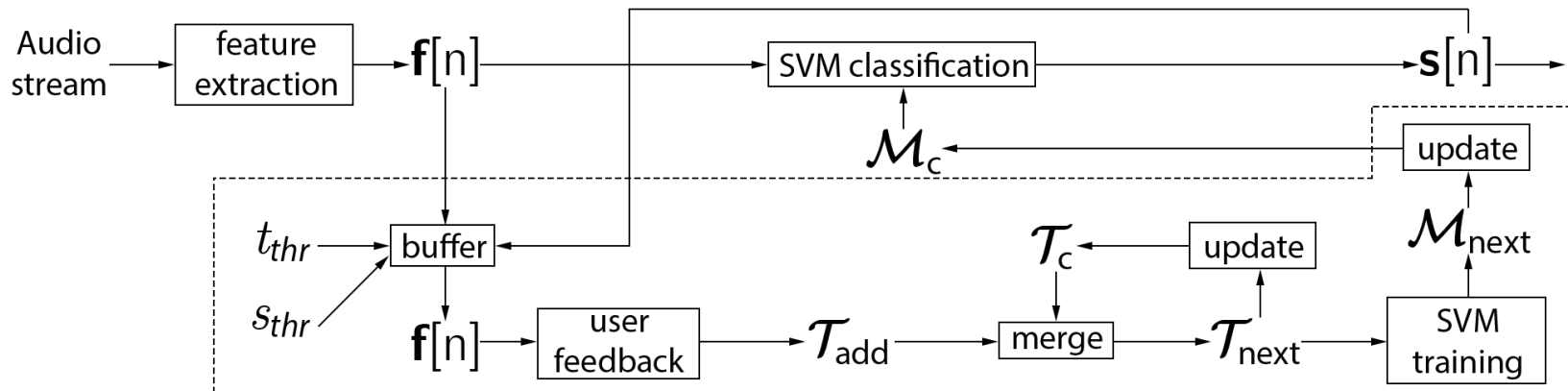
- much fewer support vectors
- without sacrificing much of the model's discriminative power





Inter-Active Learning Adaptation (IALA)

- Adaptation of a pre-trained SVM model to a single user based on inter-active feedback requests for ambiguous time intervals
- Can be applied both on directly trained models and on PALT models
- Based on time and SVM score thresholds





Dataset

- Recordings from 8 subjects using Invisio M3h microphone
- Recording protocol
 - 7 food types
 - Non-chewing activities
 - Both silent and noisy setups
- Ground truth labels assigned based on time-stamps and visual inspection of the captured signals
- Total duration: 90 minutes
- Prior probability: 0.45 for chewing class



Experimental evaluation

- Cross-Validation (CV) and Leave-One-Subject-Out (LOSO) experiment setups
- Baseline and PALT performance comparison
 - Recordings from 8 subjects
- IALA method evaluation using both directly trained and PALT as base models
 - Recordings from 6 subjects that recorded the protocol twice



PALT Evaluation Results

k-fold Cross-Validation (k=14)

initial train set has 40 feature vectors

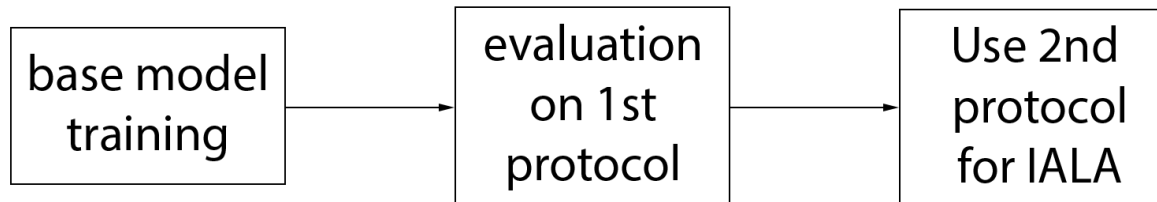
	Precision	Recall	F1 score	Accuracy	SVs
CV baseline	0.89	0.89	0.89	0.90	33552
CV PALT@100	0.83	0.89	0.86	0.87	232
CV PALT@800	0.85	0.90	0.87	0.88	1633

LOSO evaluation

	Precision	Recall	F1 score	Accuracy	SVs
LOSO baseline	0.84	0.81	0.81	0.83	31152
LOSO PALT@100	0.82	0.79	0.79	0.83	233
LOSO PALT@800	0.81	0.82	0.8	0.83	1632



IALA Evaluation Results



	Precision	Recall	F1 score	Accuracy	SVs
LOSO baseline	0.84	0.82	0.81	0.82	25043
LOSO PALT	0.87	0.66	0.72	0.82	1633
LOSO base + IALA	0.84	0.83	0.82	0.83	25038
LOSO PALT + IALA	0.88	0.80	0.83	0.85	1652

LOSO evaluation on 6 subjects that recorded the protocol twice



Conclusions

- We use active learning techniques for two tasks
 - Create and deploy classification models with fewer SVs that require reduced computational resources
 - Per-user adaptation of the deployed model, requiring minimal user feedback, and leading to increased accuracy
- User adaptation with IALA has better performance when used with a PALT base model
- Validation on an experimental dataset recorded in lab conditions shows inter-subject accuracy of 0.85 using user-adapted models and parsimonious initial SVM models
- Future work: Evaluation of the proposed system on a larger dataset under free-living conditions



Thank you