ᛦ **pariajm** / **sharif-emotional-speech-database**   `Public`

forked from mansourehk/ShEMO

<> Code   ᛦ Pull requests   ⊙ Actions   ⊞ Projects   📖 Wiki   ⊘ Security   📈 Insights

---

ᛦ master ⌄

sharif-emotional-speech-database / README.md

🖼 **pariajm** Update README.md   ⟲ **History**

👥 **2 contributors**   🖼 🟣

≣   63 lines (52 sloc)   │   3.81 KB   ···

`Open Source` 🧡

ShEMO: a large-scale validated database for Persian speech emotion detection

Abstract

This paper introduces a large-scale, validated database for Persian called Sharif Emotional Speech Database (ShEMO). The database includes 3000 semi-natural utterances, equivalent to 3 hours and 25 minutes of speech data extracted from online radio plays. The ShEMO covers speech samples of 87 native-Persian speakers for five basic emotions including *anger*, *fear*, *happiness*, *sadness* and *surprise*, as well as neutral state. Twelve annotators label the underlying emotional state of utterances and majority voting is used to decide on the final labels. According to the kappa measure, the inter-annotator agreement is 64% which is interpreted as "substantial agreement". We also present benchmark results based on common classification methods in speech emotion detection task. According to the experiments, support vector machine achieves the best results for both gender-independent (58.2%) and gender-dependent models (female=59.4%,

male=57.6%). The ShEMO is available for academic purposes free of charge to provide a baseline for further research on Persian emotional speech.

Download Database

- [Link to female utterances](#)
- [Link to male utterances](#)
- [Link to transcripts](#)

Overview of Sharif Emotional Speech (ShEMO) Database

Feature	Status
Access	Open source
Language	Persian
Modality	Speech
Duration	3 hours and 25 minutes
#Utterances	3000
#Speakers	87 (31 females, 56 males)
#Emotions	5 basic emotions (anger, fear, happiness, sadness and surprise) and neutral state
Orthographic transcripts	Available
Phonetic transcripts	Available

Read our paper on [Springer](#) or [arxiv](#)

Description of Labels

The characters used in the label of the utterances and their corresponding meaning:

- **A**: anger emotion
- **F**: female speaker (if used at the beginning of the label e.g. `F14A09`) or fear (if used in the middle of the label e.g. `M02F01`)

- **H** : happiness
- **M** : male speaker
- **N** : neutral
- **S** : sadness
- **W** : surprise

e.g. `F03S02` **F** means the speaker is **female**, **03** denotes **the speaker code**, **S** refers to the underlying emotion of the utterance which is **sadness**, **02** means this is the **second utterance for this speaker in sad emotion**.

(دادگان گفتار احساسی شریف (شمو

برای دریافت مقاله اینجا کلیک کنید

Citation

If you use this dataset, please cite the following paper:

```
@Article{MohamadNezami2019,
author  = {Mohamad Nezami, Omid and Jamshid Lou, Paria and Karami, Mansoureh},
title = {ShEMO: a large-scale validated database for Persian speech emotion
detection},
journal = {Language Resources and Evaluation},
year  = {2019},
volume  = {53},
number  = {1},
pages = {1--16},
issn  = {1574-0218},
doi = {10.1007/s10579-018-9427-x},
url = {https://doi.org/10.1007/s10579-018-9427-x}
}
```

Contact

Paria Jamshid Lou paria.jamshid-lou@hdr.mq.edu.au