Scopus

Search   Sources   Lists   SciVal ↗

SS

Document details

CSV export ⌄   ⬇ Download   🖶 Print   ✉ E-mail   🗎 Save to PDF   ☆ Save to list   More... ›

Full Text   View at Publisher

Advances in Intelligent Systems and Computing
Volume 1350 AISC, 2021, Pages 21-30
2nd International Conference on Innovative Technology, Engineering and Sciences, iCITES 2020;
Pekan; Malaysia; 22 December 2020 through 22 December 2020; Code 256319

Speech Emotion Recognition Using Deep Neural Networks on Multilingual Databases (Conference Paper)

Qadri, S.A.A.[a], Gunawan, T.S.[a,b] ✉, Wani, T.M.[a], Ambikairajah, E.[b], Kartiwi, M.[c], Ihsanto, E.[d]   👤

🗐 Save all to author list

[a] Electrical and Computer Engineering Department, International Islamic University Malaysia, Gombak, Malaysia
[b] School of Electrical Engineering and Telecommunications, UNSW, Sydney, Australia
[c] Information Systems Department, International Islamic University Malaysia, Gombak, Malaysia

View additional affiliations ⌄

Abstract   ⌄ View references (14)

The research community's ever-increasing interest in studying human-computer interactions (HCI), systems deducing, and identifying a speech signal's emotional aspects has emerged as a hot research topic. Speech Emotion Recognition (SER) has brought the development of automated and intelligent analysis of human utterances to reality. Typically, an SER system focuses on extracting the features from speech signals such as pitch frequency, formant features, energy-related and spectral features, tailing it with a classification quest to understand the underlying emotion. The key issues pivotal for a successful SER system are driven by the proper selection of proper emotional feature extraction techniques. In this paper, Mel-frequency Cepstral Coefficient (MFCC) and Teager Energy Operator (TEO) along with a new proposed Feature Fusion of MFCC and TEO referred to as Teager-MFCC (TMFCC) is examined over a multilingual database consisting of English, German and Hindi languages. Deep Neural Networks have been used to classify the different emotions considered, happy, sad, angry, and neutral. Evaluation results show that the proposed fusion TMFCC with a recognition rate of 92.7% outperforms TEO and MFCC. With TEO and MFCC configurations, the recognition rate has been found as 88.5% and 90.0%, respectively. © 2021, The Author(s), under exclusive license to Springer Nature Switzerland AG.

SciVal Topic Prominence ⓘ

Topic: Emotion Recognition | Emotional Speech | Paralanguage

Prominence percentile: 98.667   ⓘ

Author keywords

(Deep Neural Networks (DNN)) (Mel-frequency Cepstral Coefficient (MFCC)) (Speech Emotion Recognition)
(Teager Energy Operator (TEO))

Indexed keywords

Engineering
controlled terms: (Continuous speech recognition) (Database systems) (Deep neural networks) (Feature extraction)
                       (Human computer interaction) (Speech communication)

Metrics ⓘ   View all metrics ›

✳

PlumX Metrics   ⌄
Usage, Captures, Mentions,
Social Media and Citations
beyond Scopus.

Cited by 0 documents

Inform me when this document
is cited in Scopus:

Set citation alert ›

Related documents

Comparison of similarity among
sub-categories of angry speech
emotion

Palo, H.K. , Mohanty, J. ,
Mohanty, M.N.
(2017) Proceedings of IEEE
International Conference on
Emerging Technological Trends
in Computing, Communications
and Electrical Engineering, ICETT
2016

Human Emotion Recognition
using Acoustic Features with
Optimized Feature Selection and
Fusion Techniques

Lingampeta, D. , Yalamanchili, B.
(2020) Proceedings of the 5th
International Conference on
Inventive Computation
Technologies, ICICT 2020

Design and development of a
multi-lingual speech corpora
(TaMaR-EmoDB) for emotion
analysis

Rajan, R. , Haritha, U.G. , Sujitha,
A.C.
(2019) Proceedings of the Annual
Conference of the International
Speech Communication
Association, INTERSPEECH

View all related documents based
on references

Find more related documents in
Scopus based on:

Engineering uncontrolled terms	Feature extraction techniques Human computer interaction (HCI) Intelligent analysis Mel-frequency cepstral coefficients Multilingual database Research communities Speech emotion recognition Teager energy operators

Engineering main heading: Neural networks

Funding details

Funding sponsor	Funding number	Acronym
Ministry of Education - Singapore	FRGS/1/2018/ICT02/UIAM/02/4,FRGS19-076-0684	MOE
University of New South Wales See opportunities by UNSW↗		UNSW
Ministry of Higher Education, Malaysia		MOHE
Universitas Mercu Buana		UMB

References (14)

View in search results format ⟩

☐ All   CSV export ⌄   🖨 Print   ✉ E-mail   📄 Save to PDF   Create bibliography

☐ 1   Lu, X., Li, S., Fujimoto, M.

Automatic speech recognition

(2020) *SpringerBriefs in Computer Science*, pp. 21-38.
http://www.springer.com/series/10028
doi: 10.1007/978-981-15-0595-9_2

View at Publisher

☐ 2   Dzedzickis, A., Kaklauskas, A., Bucinskas, V.

Human emotion recognition: Review of sensors and methods (Open Access)

(2020) *Sensors (Switzerland)*, 20 (3), art. no. 592. Cited 34 times.
https://www.mdpi.com/1424-8220/20/3/592/pdf
doi: 10.3390/s20030592

View at Publisher

☐ 3   Furui, S., Kikuchi, T., Shinnaka, Y., Hori, C.

Speech-to-text and speech-to-speech summarization of spontaneous speech
(Open Access)

(2004) *IEEE Transactions on Speech and Audio Processing*, 12 (4), pp. 401-408. Cited 128 times.
doi: 10.1109/TSA.2004.828699

View at Publisher

☐ 4   Swain, M., Routray, A., Kabisatpathy, P.

Databases, features and classifiers for speech emotion recognition: a review

(2018) *International Journal of Speech Technology*, 21 (1), pp. 93-120. Cited 77 times.
www.kluweronline.com/issn/1381-2416/
doi: 10.1007/s10772-018-9491-z

View at Publisher

☐ 5   Cited 2 times.
El Ayadi, M., Kamel, M., Karray, F.: Survey on speech emotion recognition: features, classification schemes,
and databases. Pattern Recogn. 44(3), 572–587 (2011). https://doi.org/10. 1016/j.patcog.2010.09.020

☐ 6   Kuchibhotla, S., Vankayalapati, H.D., Anne, K.R.

An optimal two stage feature selection for speech emotion recognition using acoustic
features

(2016) *International Journal of Speech Technology*, 19 (4), pp. 657-667. Cited 20 times.
www.kluweronline.com/issn/1381-2416/
doi: 10.1007/s10772-016-9358-0

View at Publisher

☐ 7   Bandela, S.R., Kishore Kumar, T.

Emotion recognition of stressed speech using teager energy and linear prediction
features

(2018) *Proceedings - IEEE 18th International Conference on Advanced Learning Technologies, ICALT
2018*, art. no. 8433554, pp. 422-425. Cited 11 times.
http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=8432371
ISBN: 978-153866049-2
doi: 10.1109/ICALT.2018.00107

View at Publisher

☐ 8   Shrestha, A., Mahmood, A.

Review of deep learning algorithms and architectures   (Open Access)

(2019) *IEEE Access*, 7, art. no. 8694781, pp. 53040-53065. Cited 163 times.
http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=6287639
doi: 10.1109/ACCESS.2019.2912200

View at Publisher

☐ 9   Fayek, H.M., Lech, M., Cavedon, L.

Towards real-time Speech Emotion Recognition using deep neural networks

(2015) *2015, 9th International Conference on Signal Processing and Communication Systems, ICSPCS 2015
- Proceedings*, art. no. 7391796. Cited 43 times.
ISBN: 978-146738118-5
doi: 10.1109/ICSPCS.2015.7391796

View at Publisher

☐ 10   Abdelwahab, M., Busso, C.

Study of Dense Network Approaches for Speech Emotion Recognition

(2018) *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing -
Proceedings*, 2018-April, art. no. 8461866, pp. 5084-5088. Cited 16 times.
ISBN: 978-153864658-8
doi: 10.1109/ICASSP.2018.8461866

View at Publisher

☐ 11   Tillman, T.W., Carhart, R.
(1966) *An Expanded Test for Speech Discimination Utilizing CNC Monosyllabic Words*. Cited 252 times.

☐ 12   Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., Weiss, B.
(2005) *A Database of German Emotional Speech*. Cited 229 times.

☐ 13   Koolagudi, S.G., Reddy, R., Yadav, J., Rao, K.S.

IITKGP-SEHSC : Hindi speech corpus for emotion analysis

(2011) *2011 International Conference on Devices and Communications, ICDeCom 2011 - Proceedings*, art.
no. 5738540. Cited 25 times.
ISBN: 978-142449190-2
doi: 10.1109/ICDECOM.2011.5738540

 View at Publisher

☐ 14   Zhou, G., Hansen, J.H.L., Kaiser, J.F.: Nonlinear feature based classification of speech under stress. IEEE
Trans. Speech Audio Process. 9, 201–216 (2001). https://doi.org/10.1109/89. 905995

 Gunawan, T.S.; Electrical and Computer Engineering Department, International Islamic University Malaysia,
Gombak, Malaysia; email:tsgunawan@iium.edu.my

About Scopus

What is Scopus

Content coverage

Scopus blog

Scopus API

Privacy matters

Language

日本語に切り替える

切换到简体中文

切換到繁體中文

Русский язык

Customer Service

Help

Contact us