Lecture Notes in Electrical Engineering • Volume 835, Pages 711 - 724 • 2022 • 8th International Conference on Computational Science and Technology, ICCST 2021 • Virtual, Online • 28 August 2021 through 29 August 2021 • Code 275849

Document type
Conference Paper

Source type
Book Series

ISSN
18761100

ISBN
978-981168514-9

DOI
10.1007/978-981-16-8515-6_54

Publisher
Springer Science and Business Media Deutschland GmbH

Original language
English

Volume Editors
Alfred R., Lim Y.

View less

Microphone-Independent Speech Features for Automatic Depression Detection Using Recurrent Neural Network

Ezzi, Mugahed Al-Ezzi Ahmed[a] ; Hashim, Nik Nur Wahidah Nik[a] ; Basri, Nadzirah Ahmad[b]

Save all to author list

[a] Department of Mechatronics Engineering, Faculty of Engineering, International Islamic University Malaysia, Gombak, Malaysia

[b] Department of Psychiatry, Faculty of Medicine, International Islamic University Malaysia, Jalan Hospital, Pahang, Kuantan, 25000, Malaysia

Full text options

Abstract

Author keywords

Indexed keywords

SciVal Topics

Metrics

Funding details

Abstract

Depression is a common mental disorder that has a negative impact on individuals, society, and the economy. Traditional clinical diagnosis methods are subjective and necessitate extensive expert participation. Because it is fast, convenient, and non-invasive, automatic depression detection using speech signals is a promising depression objective biomarker. Acoustic feature extraction is one of the most challenging techniques for speech analysis applications in mobile phones. The values of the extracted acoustic features are significantly influenced by adverse environmental noises, a wide range of microphone specifications, and various types of recording software. This study identified microphone-independent acoustic features and utilized them in developing an end-to-end recurrent neural network model to classify depression from Bahasa Malaysia speech . The dataset includes 110 female participants. Patient Health Questionnaire 9, Malay Beck Depression Inventory-II, and subjects' declaration of Major Depressive Disorder diagnosis by a trained clinician were used to determine depression status. Multiple combinations of speech types were compared and discussed. Robust acoustic features derived from female spontaneous speech achieved an accuracy of 85%. © 2022, The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd.

Author keywords

Acoustic features ; Deep learning; Depression detection ; Speech analysis

Indexed keywords

SciVal Topics

Metrics

Funding details

References (38)

View in search results format >

All

Export Print E-mail Save to PDF Create bibliography

1 (2017) *Depression and Other Common Mental Disorders: Global Health Estimates*. Cited 3386 times.
World Health Organization

2 Mukhtar, F., Oei, O.T.P.S.
A review on the prevalence of depression in Malaysia

(2011) *Current Psychiatry Reviews*, 7 (3), pp. 234-238. Cited 33 times.
doi: 10.2174/157340011797183201

View at Publisher

3 Institute for Public Health (2015) National health and morbidity survey 2015 (NHMS
(2015) *Ministry of Health*
Malaysia Kuala Lumpur

4 World Health Organization (2019) GHO|Human resources—data by country. In: World Health Organization. https://apps.who.int/gho/data/view.main.MHHRv. Accessed 25 Jan 2021

5 Guan, N.C., Lee, T.C., Francis, B., Yen, T.S.
Psychiatrists in Malaysia: The ratio and distribution
(2018) *Malays J Psychiatry*, 27, pp. 4-12. Cited 8 times.

6 Jiang, H., Hu, B., Liu, Z., Yan, L., Wang, T., Liu, F., Kang, H., (...), Li, X.
Investigation of different speech types and emotions for detecting depression using different classifiers

(2017) *Speech Communication*, 90, pp. 39-46. Cited 52 times.
doi: 10.1016/j.specom.2017.04.001

View at Publisher

7 Kraepelin, E.
Manic Depressive Insanity and Paranoia
(1921) *J Nerv Ment Dis*, 53, p. 350. Cited 66 times.

8 Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., Quatieri, T.F.
A review of depression and suicide risk assessment using speech analysis

(2015) *Speech Communication*, 71, pp. 10-49. Cited 364 times.
doi: 10.1016/j.specom.2015.03.004

View at Publisher

9 Stasak, B., Epps, J., Goecke, R.
Elicitation design for acoustic depression classification: An investigation of articulation effort, linguistic complexity, and word affect

(2017) *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2017-August, pp. 834-838. Cited 17 times.
http://www.isca-speech.org
doi: 10.21437/Interspeech.2017-1223

View at Publisher

10 Afshan, A., Guo, J., Park, S.J., Ravi, V., Flint, J., Alwan, A.
Effectiveness of voice quality features in detecting depression

(2018) *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2018-September, pp. 1676-1680. Cited 22 times.
https://www.isca-speech.org/iscaweb/index.php/online-archive
doi: 10.21437/Interspeech.2018-1399

View at Publisher

11 Baranyi, P., Csapo, A., Sallai, G.
Cognitive infocommunications (CogInfoCom)

(2015) *Cognitive Infocommunications (CogInfoCom)*, pp. 1-219. Cited 357 times.
http://dx.doi.org/10.1007/978-3-319-19608-4
ISBN: 978-331919608-4; 978-331919607-7
doi: 10.1007/978-3-319-19608-4

View at Publisher

12 Alpert, M., Pouget, E.R., Silva, R.R.
Reflections of depression in acoustic measures of the patient's speech

(2001) *Journal of Affective Disorders*, 66 (1), pp. 59-69. Cited 118 times.
doi: 10.1016/S0165-0327(00)00335-9

View at Publisher

13 Cannizzaro, M., Harel, B., Reilly, N., Chappell, P., Snyder, P.J.
Voice acoustical measurement of the severity of major depression

(2004) *Brain and Cognition*, 56 (1), pp. 30-35. Cited 110 times.
doi: 10.1016/j.bandc.2004.05.003

View at Publisher

14 Hönig, F., Batliner, A., Nöth, E., Schnieder, S., Krajewski, J.
Automatic modelling of depressed speech: Relevant features and relevance of gender

(2014) *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pp. 1248-1252. Cited 34 times.
http://www.isca-speech.org

View at Publisher

15 Mundt, J.C., Vogel, A.P., Feltner, D.E., Lenderking, W.R.
Vocal acoustic biomarkers of depression severity and treatment response (Open Access)

(2012) *Biological Psychiatry*, 72 (7), pp. 580-587. Cited 141 times.
doi: 10.1016/j.biopsych.2012.03.015

View at Publisher

16 Stassen, H.H., Kuny, S., Hell, D.
The speech analysis approach to determining onset of improvement under antidepressants

(1998) *European Neuropsychopharmacology*, 8 (4), pp. 303-310. Cited 58 times.
doi: 10.1016/S0924-977X(97)00090-4

View at Publisher

17 Liu, Z., Kang, H., Feng, L., Zhang, L.

Speech pause time: A potential biomarker for depression detection

(2017) *Proceedings - 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2017*, 2017-January, pp. 2020-2025. Cited 6 times.
ISBN: 978-150903049-1
doi: 10.1109/BIBM.2017.8217971

View at Publisher

18 Low, L.-S.A., Maddage, N.C., Lech, M., Sheeber, L.B., Allen, N.B.

Detection of clinical depression in adolescents' speech during family interactions (Open Access)

(2011) *IEEE Transactions on Biomedical Engineering*, 58 (3 PART 1), pp. 574-586. Cited 129 times.
doi: 10.1109/TBME.2010.2091640

View at Publisher

19 Cummins, N., Epps, J., Breakspear, M., Goecke, R.

An investigation of depressed speech detection: Features and normalization

(2011) *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pp. 2997-3000. Cited 95 times.

View at Publisher

20 Scherer, S., Stratou, G., Mahmoud, M., Boberg, J., Gratch, J., Rizzo, A., Morency, L.-P.

Automatic behavior descriptors for psychological disorder analysis

(2013) *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2013*, art. no. 6553789. Cited 83 times.
ISBN: 978-146735545-2
doi: 10.1109/FG.2013.6553789

View at Publisher

21 Alghowinem, S., Goecke, R., Wagner, M., Epps, J., Gedeon, T., Breakspear, M., Parker, G.

A comparative study of different classifiers for detecting depression from spontaneous speech

(2013) *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, art. no. 6639227, pp. 8022-8026. Cited 67 times.
ISBN: 978-147990356-6
doi: 10.1109/ICASSP.2013.6639227

View at Publisher

22 Kiss, G., Tulics, M.G., Sztahó, D., Esposito, A., Vicsi, K.

Language independent detection possibilities of depression by speech

(2016) *Smart Innovation, Systems and Technologies*, 48, pp. 103-114. Cited 17 times.
http://www.springer.com/series/8767
doi: 10.1007/978-3-319-28109-4_11

View at Publisher

23 Kiss, G., Vicsi, K.

Physiological and cognitive status monitoring on the base of acoustic-phonetic speech parameters

(2014) *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8791, pp. 120-131. Cited 12 times.
http://springerlink.com/content/0302-9743/copyright/2005/
doi: 10.1007/978-3-319-11397-5_9

View at Publisher

24 Kiss, G., Vicsi, K.
Comparison of read and spontaneous speech in case of automatic detection of depression. In: 2017 8th IEEE international conference on cognitive infocommunications (CogInfoCom). IEEE, Debrecen
(2017) *Pp 000213–000218*

25 Kiss, G., Vicsi, K.

Mono- and multi-lingual depression prediction based on speech processing

(2017) *International Journal of Speech Technology*, 20 (4), pp. 919-935. Cited 17 times.
www.kluweronline.com/issn/1381-2416/
doi: 10.1007/s10772-017-9455-8

View at Publisher

26 Long, H., Guo, Z., Wu, X., Hu, B., Liu, Z., Cai, H.

Detecting depression in speech: Comparison and combination between different speech types

(2017) *Proceedings - 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2017*, 2017-January, pp. 1052-1058. Cited 14 times.
ISBN: 978-150903049-1
doi: 10.1109/BIBM.2017.8217802

View at Publisher

27 Vlasenko, B., Sagha, H., Cummins, N., Schuller, B.

Implementing gender-dependent vowel-level analysis for boosting speech-based depression recognition (Open Access)

(2017) *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2017-August, pp. 3266-3270. Cited 12 times.
http://www.isca-speech.org
doi: 10.21437/Interspeech.2017-887

View at Publisher

28 Liu, Z., Li, C., Gao, X., Wang, G., Yang, J.
Ensemble-based depression detection in speech

(2017) *Proceedings - 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2017*, 2017-January, pp. 975-980. Cited 6 times.
ISBN: 978-150903049-1
doi: 10.1109/BIBM.2017.8217789

View at Publisher

29 Stasak, B., Epps, J., Lawson, A.
Analysis of phonetic markedness and gestural effort measures for acoustic speech-based depression classification

(2018) *2017 7th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos, ACIIW 2017*, 2018-January, pp. 165-170. Cited 2 times.
ISBN: 978-153860680-3
doi: 10.1109/ACIIW.2017.8272608

View at Publisher

30 Wang, J., Sui, X., Hu, B., Flint, J., Bai, S., Gao, Y., Zhou, Y., (...), Zhu, T.
Detecting postpartum depression in depressed people by speech features

(2018) *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10745 LNCS, pp. 433-442. Cited 4 times.
http://springerlink.com/content/0302-9743/copyright/2005/
ISBN: 978-331974520-6
doi: 10.1007/978-3-319-74521-3_46

View at Publisher

31 Su, Y., Zhang, K., Wang, J., Zhou, D., Madani, K.
Performance analysis of multiple aggregated acoustic features for environment sound classification (Open Access)

(2020) *Applied Acoustics*, 158, art. no. 107050. Cited 25 times.
http://www.journals.elsevier.com/applied-acoustics/
doi: 10.1016/j.apacoust.2019.107050

View at Publisher

32 Ghosal, D., Kolekar, M.H.
Music genre recognition using deep neural networks and transfer learning

(2018) *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2018-September, pp. 2087-2091. Cited 37 times.
https://www.isca-speech.org/iscaweb/index.php/online-archive
doi: 10.21437/Interspeech.2018-2045

View at Publisher

33 Ellis, D.
(2007) *Chroma Feature Analysis and Synthesis. Resources of Laboratory for the Recognition and Organization of Speech and Audio-Labrosa*. Cited 29 times.

34 Kattel, M., Nepal, A., Shah, A.K., Shrestha, D.
Chroma feature extraction
(2019) *Conference: Chroma Feature Extraction Using Fourier Transform*. Cited 13 times.

35 Cohn, R.
Introduction to neo-riemannian theory: A survey and a historical perspective
(1998) *J Music Theory*, 42, p. 167. Cited 117 times.
https://doi.org/10.2307/843871

36 Jiang, D.-N., Lu, L., Zhang, H.-J., Tao, J.-H., Cai, L.-H.
Music type classification by spectral contrast feature
(Open Access)

(2002) *Proceedings - 2002 IEEE International Conference on Multimedia and Expo, ICME 2002*, 1, art. no. 1035731, pp. 113-116. Cited 221 times.
ISBN: 0780373049
doi: 10.1109/ICME.2002.1035731

View at Publisher

37 Davis, S.B., Mermelstein, P.
Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences

(1980) *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28 (4), pp. 357-366. Cited 3404 times.
doi: 10.1109/TASSP.1980.1163420

View at Publisher

38 McFee, B., Raffel, C., Liang, D., Ellis, D.P., McVicar, M., Battenberg, E., Nieto, O.
) librosa: Audio and music signal analysis in python
(2015) *Proceedings of the 14Th Python in Science Conference*, pp. 18-25. Cited 893 times.
Citeseer, pp

Hashim, N.N.W.N.; Department of Mechatronics Engineering, Faculty of Engineering, International Islamic University Malaysia, Gombak, Malaysia; email:niknurwahidah@iium.edu.my

About Scopus

What is Scopus

Content coverage

Scopus blog

Scopus API

Privacy matters

Language

日本語版を表示する

查看简体中文版本

查看繁體中文版本

Просмотр версии на русском языке

Customer Service

Help

Tutorials

Contact us