Topic: Depression Detection Using Hearing-Aid Integrated Sensors

Depression is a prevalent clinical condition, affecting over 280 million individuals globally, including 5% of the adult population [1]. For those affected, the condition represents a considerable burden, entailing a loss of quality of life and a reduction in life expectancy [2]. Early detection is crucial, as timely intervention can significantly enhance outcomes [3]. The duration of untreated depression plays a critical role in recovery, with earlier treatment being associated with higher remission rates [4], ultimately improving mental and social functioning [5].

Depression is usually diagnosed by a trained professional (e.g. physicians, psychiatrists, psychologists) using structured interviews and standardized questionnaires such as the Hamilton Rating Scale for Depression (HAM-D) [6], the Beck Depression Inventory-II (BDI-II) [7], and the Patient Health Questionnaire (PHQ) [8]. However, these assessments can only be performed when patients seek medical care from these professionals.

There is an emerging field of research investigating alternative approaches for depression detection by assessing the potential of objective digital biomarkers, such as voice, activity, gait, sleep time and many more [9]. A frequently employed biomarker for the detection of depression is voice [10, 11], given that individuals with depression exhibit notable alterations in their speech patterns. These include, among others, a reduction in intensity, a decrease in fundamental frequency deviation (FFD) or pitch range, and a slowing of speech rate [12]. The application of machine learning has enabled the development of numerous automated speech analysis techniques that use these paralinguistic features [13]. Another objective digital biomarker that can be indicative of depression is gait, including a slumped posture, impaired dynamic balance, and reduced gait speed [14], and other generic and expert motion features, which have already been shown to be promising for assessing the influence of stress on body posture and movement [15].

Data from these digital biomarkers can be obtained in a number of ways. In addition to external or ambient sensors, which are fixed in the environment, wearable sensing devices are often used, for instance, inertial sensors, smartphones, or smartwatches [16].

While it has been shown that voice analysis, as well as motion and gait patterns, can be used to detect depression, it has not been analyzed whether the combination of these modalities can be used to enhance depression detection. Furthermore, the majority of existing approaches are not suitable for everyday use, as they require the user to wear one or more sensors. Therefore, a hearing aid, a wearable device on the head, has the advantage of recording voice, gait, and other movements directly, as it inherently integrates a motion sensor as well as a microphone sensor. While earables in general are gaining popularity [18, 21], the combination of sensors on the ear (or head) to detect depression represents a novel approach that has not yet been documented in the literature.

The objective of this thesis is to develop an algorithm that can predict depression based on the combined analysis of voice, gait, and motion from hearing aid sensors. This entails data acquisition, feature extraction, and the development and evaluation of various machine and deep learning approaches. The data will be acquired through the utilization of hearing aid-integrated sensors, specifically the inertial sensor and the microphone. A study will be planned and conducted to collect data from 20 participants (10 healthy and 10 depressed individuals). The PHQ-8 will be used to assess the presence and severity of depression. To obtain voice data, a semi-structured interview will be conducted, which is recorded by the hearing aid microphones worn by the participants. To gather data related to the person's gait, a 6-minute walking test will be conducted. Additionally, head movements will be recorded in different situations to be able to gain insight into the participants' activity and movement. All tests will be conducted in a laboratory setting.

To predict the depression score, different algorithmic approaches will be implemented and evaluated. Following the extraction of features from the gait and motion data, different classifiers are evaluated with respect to their classification performance of distinguishing healthy from depressed individuals. In the first approach, voice features will be extracted by using established frameworks such as openSMILE [17] as well as gait and motion features, e.g. by using EarGait [18] and tsfresh [19]. Subsequently, a classification is performed. The second method differs primarily in the processing of voice. In this case, wav2vec 2.0, a

present deep learning model for emotion recognition [20], is used for depression detection. Prior to training, the model is pre-trained with speech data from other datasets. Following, fine-tuning is done by using speech recorded by hearing aid microphones. Different information fusion techniques will be applied to find the best-performing depression classification method.

The proposed work consists of the following parts:

- Literature research of relevant work in the field of voice analysis and motion analysis for depression detection resulting in a comprehensive understanding of the phenomenon.

- Research of existing algorithms (with a focus on machine and deep learning models) for automatic detection and classification of depression based on digital biomarkers.

- Planning and realization of a study with (if possible) 10 healthy and 10 depressed participants with the objective of collecting voice, gait, and motion data.

- Extraction of meaningful voice, gait and motion features, amongst others using open source libraries such as [16, 17, 18, 19].

- Statistical analysis of extracted features to understand their influence on the depressive state of individuals.

- Implementation and evaluation of feature-based machine learning classifiers to predict depression.

- Detailed evaluation of the different models in terms of classification performance.

Optional:

- Design, implementation, and evaluation of a deep-learning model for depression detection using data from both sensor modalities.

The thesis must contain a detailed description of all developed and used algorithms as well as a profound result evaluation and discussion. The implemented code has to be documented and provided. An extended research on literature, existing patents and related work in the corresponding areas has to be performed.

Advisors:	Ann-Kristin Seifer
	Robert Richer
	Prof. Dr. Björn Eskofier
Student:	Maximilian Renz
Start – End:	09.12.2024 – 08.06.2025

References

[1] World Health Organization: *Depressive disorder (depression)*. Available: `https://www.who.int/news-room/fact-sheets/detail/depression`, [Accessed: Sept. 4, 2024]

[2] Jia, Haomiao and Zack, Matthew M. and Thompson, William W. et al.: *Impact of depression on quality-adjusted life expectancy (QALE) directly as well as indirectly through suicide*. Social Psychiatry and Psychiatric Epidemiology, no. 6 (2015): 939–949.

[3] Hetrick, S.E. and Parker, A.G. and Hickie, I.B. et al.: *Early Identification and Intervention in Depressive Disorders: Towards a Clinical Staging Model*. Psychotherapy and Psychosomatics, no. 5 (2008): 263–270.

[4] Bukh, J. D. and Bock, C. and Vinberg, M. and Kessing, L. V.: *The effect of prolonged duration of untreated depression on antidepressant treatment outcome*. Journal of Affective Disorders, no. 1 (2013): 42–48.

[5] Coulehan, J.L. and Schulberg, H.C. and Block, M.R. et al.: *Treating depressed prima-ry care patients improves their physical, mental, and social functioning.* Arch Intern Med, no. 10 (1997): 1113–1120.

[6] Hamilton, Max.: *A rating scale for depression.* Journal of Neurology, Neurosurgery and Psychiatry, no. 23 (1960): 56–62.

[7] Beck, A.T. and Ward, C.H. and Mendelson, M. and Mock, J. and Erbaughx J.: *An Inventory for Measuring Depression.* Archives of General Psychiatry, no. 4 (1961): 53–63.

[8] Richardson, L.P. and McCauley, E. and Grossman, D.C. et al.: *Evaluation of the Patient Health Questionnaire-9 Item for detecting major depression among adolescents.* Pediatrics, no. 126 (2010): 1117–1123.

[9] Mohr, David and Zhang, Mi and Schueller, Stephen: *Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning.* Annual Review of Clinical Psychology, no. 1 (2017): 23–47.

[10] Wu, P. and Wang, R. and Lin, H. et al.: *Automatic depression recognition by intelligent speech signal processing: A systematic survey.* CAAI Transactions on Intelligence Technology, no. 3 (2023): 701–711.

[11] Low, D. M. and Bentley, K. H. and Ghosh, S. S.: *Automated assessment of psychiatric disorders using speech: A systematic review.* Laryngoscope Investigative Otolaryngology, no. 1, pp. (2020): 96–116.

[12] Hollien, Harry.: *Vocal Indicators of Psychological Stress.* Annals of the New York Academy of Sciences, no. 1 (1980): 47–72.

[13] Wu, Pingping and Wang, Ruihao and Lin, Han et al.: *Automatic depression recognition by intelligent speech signal processing: A systematic survey.* CAAI Transactions on Intelligence Technology, no. 3 (2023): 701–711.

[14] Feldman, Ron and Schreiber, Shaul and Pick, Chaim and Been, Ella.: *Gait, Balance and Posture in Major Mental Illnesses: Depression, Anxiety and Schizophrenia.* (2020).

[15] Richer, Robert and Koch, Veronika and Abel, Luca et al.: *Machine learning-based detection of acute psychosocial stress from body posture and movements.* Sci Rep, no. 1 (2024): 8251.

[16] Garcia-Ceja, E. and Riegler, E. and Nordgreen, T. and et al.: *Mental health monitoring with multimodal sensing and machine learning: A survey.* Pervasive and Mobile Computing, no. 1 (2018): 1–26.

[17] Eyben, F. and Wöllmer, M. and Schuller, B.: *Opensmile: the munich versatile and fast open-source audio feature extractor.* Proceedings of the 18th ACM international conference on Multimedia (2010): 1459–1462.

[18] Seifer, A.-K. and Dorschky, E. and Küderle, A. et al.: *EarGait: Estimation of Temporal Gait Parameters from Hearing Aid Integrated Inertial Sensors.* Sensors, no. 14 (2023): 6565.

[19] M. Christ, M. and Braun, N. and Neuffer, J. and Kempa-Liehr, A. W.: *Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh - A Python package).* Neurocomputing (2018): 72–77.

[20] Wagner, Johannes et al.: *Dawn of the Transformer Era in Speech Emotion Recognition: Closing the Valence Gap.* IEEE Transactions on Pattern Analysis and Machine Intelligence, no. 9 (2023): pp. 10745–10759.

[21] Roeddiger, Tobias and Clarke, Christopher and Breitling, Paula and Schneegans, Tim and Zhao, Haibin and Gellersen, Hans and Beigl, Michael. *Sensing with Earables: A Systematic Literature Review and Taxonomy of Phenomena* Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 6.3 (2022): pp.1-57.