Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 10, No. 3, September 2022, pp. 584~591
ISSN: 2089-3272, DOI: 10.52549/ijeei.v10i3.3888

Development of Javanese Speech Emotion Database (Java-SED)

Fatchul Arifin[1], Ardy Seto Priambodo[2], Aris Nasuha[3], Anggun Winursito[4], Teddy Surya Gunawan[5]
[1,2,3,4]Department of Electronics and Informatics Engineering of Education, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia
[5]Department of Electrical and Computer Engineering, International Islamic University Malaysia, Malaysia

Article Info

Article history:

Received May 18, 2022
Revised Jul 14, 2022
Accepted Aug 8, 2022

Keyword:

Speech Emotion
Javanese Database
Speech Emotion Recognition
Machine Learning

ABSTRACT

Javanese is one of Indonesia's most widely spoken regional languages, alongside other regional languages. Emotions can be recognized in various ways, including facial expression, behavior, and speech. Recognizing emotions through speech is a straightforward process, but the outcomes are significant. Currently, there is no database for identifying emotions in Javanese speech. This paper aims to describe the creation of a Javanese emotional speech database. Actors from the Kamasetra UNY community who are accustomed to performing in dramatic roles participated in the recording. The location where recordings are made is free of interference and noise. The actors of Kamasetra have simulated six types of emotions: happy, sad, fear, angry, neutral, and surprised. The cast comprises ten people between 20 and 30 years old, including five men and five women. Both humans (30 Javanese-speaking verifiers ranging in age from 17 to 50) and a machine learning system (30 Javanese-speaking verifiers with ages between 17 and 50) verify the database that has been created. The verification results indicate that the database can be used for Javanese emotion recognition. The developed database is open-source and freely available to the research community at this link https://beais-uny.id/dataset/.

Corresponding Author:

Fatchul Arifin,
Department of Electronic and Informatics Engineering,
Universitas Negeri Yogyakarta,
Yogyakarta, Indonesia.
Email: fatchul@uny.ac.id

1. INTRODUCTION

Java is the thirteenth most significant island in the world and is located in the Greater Sunda archipelago in Indonesia. The population on the island of Java is approximately 150 million. Java Island is home to sixty percent of Indonesia's population [1]. Natives of the island of Java use Javanese as their everyday language. Javanese is the most widely spoken language in Indonesia. The Javanese, whose territory includes Central Java, Yogyakarta, and East Java, speak this language. There are three levels of usage for the Javanese language: Krama Inggil, Madya, and Ngoko. The Javanese term "ngoko" is used to address someone of the same age or who is already well-known. While Javanese "Krama" is a subtle form of Javanese, it is typically used when speaking to parents or older individuals. In comparison, Java's "Madya" language is between the two [2]. Javanese usually use the "Ngoko" language to communicate their emotions. This paper will create a database to recognize the "Ngoko" Javanese emotion words.

Emotions are a collection of subjective, bodily, and behaviorally influenced feelings towards something. If emotions are appropriately controlled, they will positively impact life, but if they are not, they will have a negative effect on life. Therefore, early recognition of emotions is necessary to make their management more effortless. Recognizing speech or the Speech Emotion Recognition (SER) process is one way to identify emotions [3-5]. 1973 marked the beginning of the proliferation of SER applications [6, 7]. The existence of a database is a crucial aspect of SER development. There are three types of SER-related databases:

actor (simulated), in which emotions are played by professional actors, elicited (induced), which is constructed/obtained without the speaker's knowledge; and natural, which is obtained without simulation or influencing other parties/speakers, but recorded naturally through conversation [8, 9].

There are multi-languages speech emotion databases available [10]. SAVEE [11] and LDC [12] for English, EmodDB [13] for German, SES [14] for Spanish, MES [15] for Mandarin, CASIA [16] for Chinese, and DES [17] for Danish, and IITKGP-SEHSC [18] for Hindi are some examples of SER-related databases with multiple languages. Although many speech emotion databases have been developed to date, a Javanese language speech emotion database is still required. To the best of our knowledge, the developed database is the only speech emotion database available in the Javanese language. Moreover, the created database is open source and is freely available to the research community.

2. DESIGN OF JAVANESE SPEECH EMOTION DATABASE

The Javanese speech emotion database was recorded under specific and standardized conditions to be free of noise and echo. Some identical sentences/utterances are systematically recorded with different emotions. The most important aspect of a sound dataset is the recording quality. Before beginning the sound recording process, the location of recording equipment and other supporting equipment is planned. Figure 1 depicts a sketch of the placement of the recording device. Note that we recorded audio-visual emotion, but in this paper, we will focus on speech emotion. Figure 2 depicts the actual recording process. During the recording process, a video camera is also used to record the recording process to verify the sound data. A high-quality microphone is required to make a good recording. The microphone is positioned approximately 6 to 10 inches in front of the speaker at a height parallel to the chin.

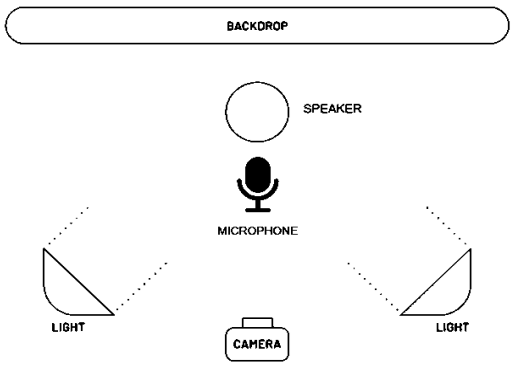

Figure 1. Audio-visual Equipment Setup for Javanese Emotion Database Recording

Figure 2. Actual Recording Process

2.1. Speakers Selection

As stated in the introduction, there are three ways to construct a database for this SER: actor (simulated), elicited (induced), and natural. The model employing the actor (simulation) has the highest probability of success among the three approaches [13, 18]. Therefore, the speakers utilized in this study are seasoned actors from one of the extracurricular arts and theater organizations at UNY (Kamasetra UNY). Each actor will deliver identical lines expressing six standard emotions: happiness, sadness, fear, disgust, surprise, and anger [19]. There are ten speakers, consisting of five men and five women between the ages of 20 and 30 years old.

2.2. Speech Material

The speech material used in this study consists of everyday sentences. The artist will later deliver this speech with various emotions. The sentences that will be used have been chosen by Javanese language experts in theater, radio drama, Javanese song films, and others because they are widely used in daily life. The research team initially composed twenty sentences. These phrases are shown in Table 1.

Table 1. Selected Speech Utterances for Javanese Speech Emotion Recording

Selected Speech Utterances	Source
Atiku arep nangis koyo diiris-iris	(Geguritan: "Ngapuro"- J. Sumarto)
Yen wis etuk jeneng, bakal etuk jenang	(Geguritan Ala Abbash CH)
Durung usaha nanging wis pasrah	(Film Anak Lanang 2017)
Kangmasmu kuwi wis mati	(Film Sultan Agung 2018)
Aku ra ngerti mas karo pikirmu	(Film Sultan Agung 2018)
Kang mas, sakjane iki gek ana apa ta	(Film Sang pencerah 2010)
Aku pengine diakui merga karyaku	(Film Yowes ben)
Cukup pisan iki aja dibaleni	(Film Yowes ben 2)
Apa ora isa kowe pisah karo dheweke	(Roman Nona Secretary-Suparto Brata)
Kabeh kuwi wis diatur sing gawe urip	(Djaka Lodang Magazine Edition No.37)
Wis ngomong wae rasah sungkan	(Sultan Agung 2018)
Diajak kenalan kok nglamun to dhik	(Short story: "Dhik" -Eko Wahyudi)
Apike ora waton nek ngendikan	(Tilik Film 2018)
Tenangna pikirmu, kowe kudu sabar	(Film Tilik 2018)
Pancene sapa sing tau kapusan	(Film Tilik 2018)
Ora ana salahe kowe njaluk ngapura	(Geguritan "Nalika Kali-Kali Nagih Janji"-Eko Nuryono)
Ana mobil banter sing nabrak pitku	(Djaka Lodang Magazine No.37)
Ayo tangi mengko mundhak telat sekolahe	(Novel Sumingkir - Faris Febri Utama)
kowe aja ngajari aku goroh	(Panjebar Spirit Magazine 2012)
Awake dhewe kudu nyukuri apa sing ana	(Panjebar Spirit Magazine 2012)
Aku tresna karo kowe	(Javanese song)

After consulting with experts on the Javanese language, it was decided to select four of the most neutral Javanese sentences so that they could be used to express a variety of emotions. These are the four sentences selected for this recording:

- *Yen wis entuk jeneng, bakal entuk jenang* (If you have earned trust, you will benefit)
- *Kabeh kuwi wis diatur sing gawe urip* (Everything is arranged by God)
- *Wis ngomong wae rasah sungkan* (Just say it, no need to be ashamed)
- *Aku pengine diakui merga karyaku* (I want to be recognized for my achievements.)

2.3. Selected Emotions

To analyze emotions in Javanese, six primary emotions experienced by the Javanese population are utilized. The six emotions are happiness, neutrality, sadness, anger, fear, and shock. All speakers must deliver four sentences for each of six fundamental emotions during a single session. Seven times per sentence, one emotion is repeated. For a total of 1,680 utterances in the database (4 sentences 7 repetitions 6 emotions 10 speakers). Each emotion has 280 expressions. The word count of each statement ranges between 5 and 7.

3. JAVANESE SPEECH EMOTION DATABASE RECORDING

In this section, the recording environment and recording process will be further elaborated on.

3.1. Recording Environment Setup

To obtain high-quality audio recordings of emotional Javanese speech, the recording must be in a standard location, free of noise, echoes, and other disturbances. This is required as the audio recording will serve as a benchmark for the SER. The recording for this study was conducted in the instrumentation system laboratory of FT UNY's Electronic Engineering Study Program. Audacity is used to record with a sampling

rate of 16 kHz and mono-channel recording. The microphone utilized is a unidirectional condenser microphone. In this research, Krisp AI, a noise-canceling software, is also used to remove noise originating from background noise or the surrounding environment.

3.2. Recording Process

The potential speakers/artists are briefed before recording. This briefing is intended to provide the speaker with a technical explanation. Each speaker will convey six distinct emotions. Each emotion has four sentence types. Each sentence within a single emotion is repeated seven times. So, 168 recorded data will be obtained from each talent (speaker). In the study, ten speakers ranging in age from 20 to 30 years, with five female speakers and five male speakers. Thus, a total of 1680 recorded data will be obtained and stored in the database. Ten speakers with 1680 are considered sufficient because they already represent various kinds of emotions and utterances. Several researchers related to the previous SER database also used personnel and routes that were not much different. For example, Vryzas et al. [20] recorded 500 utterances from 5 actors; James et al. [21] recorded 2400 utterances from four actors (two males and two females); and Neto et al.[22] recorded 12 speakers (6 female + 6 male) for a total of 1167 recordings.

The recording administrator will signal the beginning when the speaker is prepared to begin the speech (one type of emotion, one sentence). This sign is needed to synchronize the performer with the recording medium. When the recording is complete, the artist provides a specific code word that will be used to decode the emotions and sentence structures he expressed earlier. This process is done to reduce labeling errors resulting from the recording process.

The speakers are experienced actors who are part of the Kamasetra drama community at Yogyakarta State University, Indonesia. They practice expressing emotions in Javanese with the provided sentences under the supervision of a senior supervisor before the scheduled recording date.

4. EXPERIMENTAL RESULTS AND DISCUSSION

After collecting 1680 recorded data, evaluation and verification of the data are conducted. There are two evaluation and verification methods. The first is an expert evaluation conducted by a Javanese language expert. Second, machine learning-based verification is conducted by applying machine learning algorithms to the compiled database to identify Javanese emotions.

4.1. Expert Evaluation

After the recording session, experts in the Javanese language conduct a verification. This Java language expert is a Javanese language lecturer and a Javanese cultural practitioner who resides near Yogyakarta and Surakarta. The expert must assess the emotional content of the statements. Before choosing an emotional category, the expert is permitted to hear the speech twice or thrice. According to the results of this validation, 84 % of speech emotions were correctly identified.

4.2. Machine Learning Evaluation

In addition to expert evaluation, machine learning was used to test the database's ability to recognize emotional patterns. The recorded data are used as training and test data to introduce Javanese emotions. Due to the use of stereo audio during recording, the stereo data format must be converted to the mono data format. In addition, features were extracted from the existing data using the Mel-frequency cepstral coefficient (MFCC) method, one of the most popular and noise-sensitive speech recognition techniques.

K-Nearest Neighbors, Linear SVC, Random Forest Classifier, Neural Network, and Quadratic Discriminant Analysis (QDA) Classifier are the five machine learning algorithms utilized for this evaluation. Using the Python programming language and the Scikit-learn module, evaluations are conducted. Table 2 depicts the apparatus used during the evaluation phase.

Table 2. Hardware and software specifications used for evaluation

Parameter	Description
Processor	Intel Core i5 12400F
RAM	DDR4 3200Mhz 2x8GB
Storage	SSD NVMe 1TB
VGA	RTX3060 12GB GDDR6
Python	3.9
Scikit-Learn	1.1.1
Jupyter Notebook	5.0

The classification accuracy of each algorithm following the hyperparameter tuning process is displayed in Table 3.

Table 3. Machine learning performance and its tuned parameter

Method	Parameter	Accuracy
K-Nearest Neighbors	algorithm='kd_tree', n_neighbors=3	83%
Poly SVC	kernel='poly', C=3.685, degree=3.0	88%
Random Forest	max_depth=10, n_estimators=458, max_leaf_nodes=391	85%
Neural Network	random_state=1, hidden_layer_sizes=(488,443), max_iter=325, learning_rate_init=0.00233, n_iter_no_change=27, alpha=0. 00185, activation='tanh'	90%
Quadratic Discriminant Analysis	reg_param=0.0	86%

In the evaluation/testing described previously, 75% of the dataset was used for training and 25% was used for testing. The test results indicate that the KNN method has an accuracy of 83 %, the Linear SVC method has an accuracy of 88 %, the Random Forest method has an accuracy of 85 %, the Neural Network method has an accuracy of 90 %, and the QDA method has an accuracy of 86 %. Based on Table 3, it is evident that the Neural Network algorithm yields the optimal solution. Furthermore, the accuracy for each emotion is shown in Table 4, while the confusion matrix is shown in Figure 3.

Table 4. The accuracy of the neural network classifier

	precision	recall	f1-score	support
angry	0.96	0.90	0.93	84
fearful	0.90	0.89	0.89	61
happy	0.88	0.91	0.90	66
neutral	0.95	0.97	0.96	72
sad	0.87	0.81	0.84	64
surprised	0.82	0.89	0.86	73

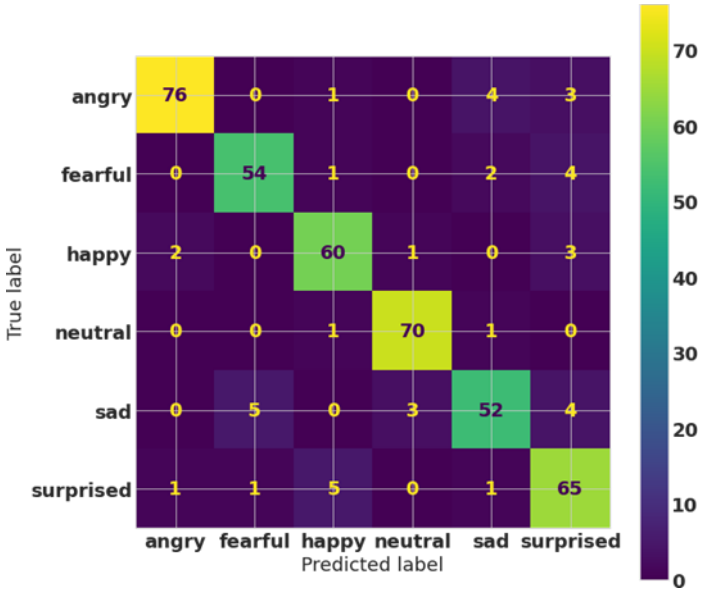

Figure 3. Confusion Matrix of Neural Network Classifier

4.3. Benchmarking Evaluation

The Speech Emotion of Java Language database development has been completed. In the study, there were ten speakers between the ages of 20 and 30, with five female and five male speakers. Each speaker will express six unique emotions (angry, fearful, happy, neutral, sad, and surprised). Each emotion corresponds to four sentence types. Each sentence is repeated seven times within a single emotion. Thus, the database will be populated with a total of 1680 recorded data. Ten speakers with a score of 1680 are deemed sufficient because it already represents a variety of emotions and speech types. For researchers who have previously created a database, the number of actors and complete recordings is comparable. This is shown in Table 5. The table demonstrates that the number of speakers and recordings produced is still acceptable and usable. Vryzas et al.[20] recorded 500 utterances from 5 actors, James et al. [21] recorded 2400 utterances from 4 actors (2 males and 2 females), and Neto et al. [22] recorded 12 speakers (6 female + 6 male) for a total of 1167 recordings.

Table 5. Benchmarking Evaluation

No	Database Name	Language	Participant	Emotion	Quantity of Database (voice recording)	Quality of Database
1	Mexican Emotional Speech Database (MESD)	Mexican	3 female, 2 male, and 6 children	six different prosodies: anger, disgust, fear, happiness, neutral, and sadness	864	89.4%, 93.9%, and 83.3% accuracy on female, male, and child voices
2	Emotional Speech Database (ESD)	Chinese, English	10 native English speakers and 10 native Chinese speakers	5 emotions: angry, happy, neutral, sad, and surprise	3500	baseline achieves better performance for neutral-to-happy conversion due to the poor performance of SER on happy (29.95%) compared with 84.32% on sad, and 70.47% on angry
3	The Variably Intense Vocalizations of Affect and Emotion Corpus (VIVAE)	English	12 speakers	non-speech 6 emotions: achievement, anger, fear, pain, pleasure, and surprise with 3 emotional intensities (low, moderate, strong, peak)	1085	
4	Acted Emotional Speech Dynamic Database (AESDD)	Greek	5 actors	5 emotions: anger, disgust, fear, happiness, and sadness	500	The accuracy of human listeners was estimated at around 74%
6	The Emotional Voices Database (EmoV-DB)	English	4 speakers- 2 males and 2 females	The emotional styles are neutral, sleepiness, anger, disgust, and amused	7000	

Table 5 depicts the quality/accuracy of the created database. The average precision of the AESDD database is 74%. The accuracy of the ESD database for the emotion cheerful is 29.5%, for sad it is 84.3%, and for furious it is 70.4%. The Speech Emotion for Javanese Language database created in this study, for instance, has a maximum accuracy of 90% when tested using the Neural Network algorithm (see Table 3). Our proposed algorithms achieved accuracies of 96%, 90%, 88%, 95%, 87%, and 82% for angry, fearful, happy, neutral, sad, and surprised, respectively. This demonstrates that the developed database can be used to detect emotion in the Javanese language. Nevertheless, other researchers can access the Javanese Speech Emotion Database at http://beais-uny.id/dataset/.

5. CONCLUSIONS

The purpose of this paper is to create a Java emotional database comprised of sound recordings with authentic emotional content. The design, recording, and validation of databases adhere to various standards. In one session, 10 speakers between the ages of 20 and 45 were selected and four sentences were recorded for each of the six basic emotions of happiness, neutrality, sadness, anger, fear, and surprise. Seven times within a single sentence, the same emotion is repeated. So that the database contains a total of 1680 utterances, each emotion has 280 expressions. The range of each statement's word count is between 5 and 7. The emotional Javanese speech recording was evaluated by Javanese language experts (84 % accuracy). The evaluation was also conducted utilizing a machine learning system with multiple algorithms, with each KNN method achieving 83 percent accuracy, the Linear SVC method achieving 88 percent accuracy, the Random Forest method achieving 85 percent accuracy, the Neural Network method achieving 90 percent accuracy, and the QDA method achieving 86 percent accuracy. Neural Network is the best algorithm in this situation, with a 90 percent accuracy rate. This demonstrates that the developed database has sufficient quality to be used in future research on speech emotion recognition.

ACKNOWLEDGMENTS

The researcher expresses profound gratitude and appreciation to Yogyakarta State University for granting permission and providing research funds so that the Javanese speech emotion database can be developed. The researcher would like to thank the Javanese experts as well who have validated our sentence selection and the Kamasetra UNY artist who has performed the actual recording. We would like also to thank IIUM for providing research facilities.

REFERENCES

[1] J. J. Stockdale, *Island of Java*. Tuttle Publishing, 2011.

[2] D. E. Subroto, M. Dwirahardjo, and B. Setiawan, "Endangered krama and krama Inggil varieties of the Javanese language," *Linguistik Indonesia*, vol. 26, no. 1, pp. 89-96, 2008.

[3] I. Luengo, E. Navas, and I. Hernáez, "Feature analysis and evaluation for automatic emotion identification in speech," *IEEE Transactions on Multimedia*, vol. 12, no. 6, pp. 490-501, 2010.

[4] S. Kuchibhotla, H. D. Vankayalapati, R. Vaddi, and K. R. Anne, "A comparative analysis of classifiers in emotion recognition through acoustic features," *International Journal of Speech Technology*, vol. 17, no. 4, pp. 401-408, 2014.

[5] T. M. Wani, T. S. Gunawan, S. A. A. Qadri, M. Kartiwi, and E. Ambikairajah, "A comprehensive review of speech emotion recognition systems," *IEEE Access*, vol. 9, pp. 47795-47814, 2021.

[6] T. S. Gunawan, M. F. Alghifari, M. A. Morshidi, and M. Kartiwi, "A review on emotion recognition algorithms using speech analysis," *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, vol. 6, no. 1, pp. 12-20, 2018.

[7] M. F. Alghifari, T. S. Gunawan, S. A. A. Qadri, M. Kartiwi, and Z. Janin, "On the use of voice activity detection in speech emotion recognition," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 4, pp. 1324–1332, 2019.

[8] S. G. Koolagudi and K. S. Rao, "Emotion recognition from speech: a review," *International journal of speech technology*, vol. 15, no. 2, pp. 99-117, 2012.

[9] Z. Zhang, E. Coutinho, J. Deng, and B. Schuller, "Cooperative learning and its application to emotion recognition from speech," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 1, pp. 115-126, 2014.

[10] S. A. A. Qadri, T. S. Gunawan, M. F. Alghifari, H. Mansor, M. Kartiwi, and Z. Janin, "A critical insight into multi-languages speech emotion databases," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 4, pp. 1312-1323, 2019.

[11] P. Jackson and S. Haq, "Surrey audio-visual expressed emotion (SAVEE) database," *University of Surrey: Guildford, UK*, 2014.

[12] M. Liberman. "Emotional prosody speech and transcripts." http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002S28 (accessed May 2022.

[13] F. Burkhardt, A. Paeschke, M. Rolfes, W. F. Sendlmeier, and B. Weiss, "A database of German emotional speech," in *Interspeech*, 2005, vol. 5, pp. 1517-1520.

[14] V. Hozjan, Z. Kacic, A. Moreno, A. Bonafonte, and A. Nogueiras, "Interface Databases: Design and Collection of a Multilingual Emotional Speech Database," in *LREC*, 2002.

[15] L. Chen, X. Mao, Y. Xue, and L. L. Cheng, "Speech emotion recognition: Features and classification models," *Digital signal processing*, vol. 22, no. 6, pp. 1154-1160, 2012.

[16] W. Bao *et al.*, "Building a chinese natural emotional audio-visual database," in *2014 12th International Conference on Signal Processing (ICSP)*, 2014: IEEE, pp. 583-587.

[17] I. S. Engberg, A. V. Hansen, O. Andersen, and P. Dalsgaard, "Design, recording and verification of a Danish emotional speech database," in *Fifth European conference on speech communication and technology*, 1997.

[18] S. G. Koolagudi, R. Reddy, J. Yadav, and K. S. Rao, "IITKGP-SEHSC: Hindi speech corpus for emotion analysis," in *2011 International conference on devices and communications (ICDeCom)*, 2011: IEEE, pp. 1-5.

[19] W. Putra and F. Arifin, "Real-time emotion recognition system to monitor student's mood in a classroom," in *Journal of Physics: Conference Series*, 2019, vol. 1413, no. 1: IOP Publishing, p. 012021.

[20] N. Vryzas, M. Matsiola, R. Kotsakis, C. Dimoulas, and G. Kalliris, "Subjective evaluation of a speech emotion recognition interaction framework," in *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion*, 2018, pp. 1-7.

[21] J. James, L. Tian, and C. I. Watson, "An open source emotional speech corpus for human robot interaction applications," in *Interspeech*, 2018, pp. 2768-2772.

[22] J. R. Torres Neto, L. Y. Mano, and J. Ueyama, "Verbo: voice emotion recognition database in Portuguese language," *Journal of Computer Science*, vol. 14, no. 11, pp. 1420-1430, 2018.

BIOGRAPHY OF AUTHORS

Fatchul Arifin received a B.Sc. in Electrical Engineering at Universitas Diponegoro in 1996. Afterward, He received a Master's degree from the Department of Electrical Engineering, Institut Teknologi Bandung (ITB) in 2003 and a Doctoral degree in Electric Engineering from Institut Teknologi Surabaya in 2014. Currently, he is the lecturer at both the undergraduate and postgraduate electronic and informatic programs of the Engineering Faculty at Universitas Negeri Yogyakarta. His research interests include but are not limited to Artificial Intelligent Systems, Machine Learning, Fuzzy Logic, and Biomedical Engineering.

Aris Nasuha received a B.Sc. in Physics at Universitas Gadjah Mada in 1993. Afterward, he received a Master and Doctoral degree from the Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember Surabaya (ITS) in 2001 and 2019. His research interests include but are not limited to Artificial Intelligent Systems, Machine Learning, and Digital Signal Processing. He is the lecturer at Department of Electronics and Informatic Engineering (1994-now), and head of Bachelor Applied Science in Electronic Engineering (2019-now), Faculty of Engineering, Universitas Negeri Yogyakarta.

Ardy Seto Priambodo obtained his B.Eng. degree in Electrical Engineering from Institut Teknologi Sepuluh Nopember (ITS), Indonesia, in 2012. He received his M.Eng. degree in 2018 from Electrical Engineering, Universitas Gadjah Mada (UGM), Indonesia. He obtained a professional engineer degree (Ir.) from Universitas Negeri Yogyakarta in 2020. His research interests are on control systems, instrumentation, and robotics. He is currently a lecturer at Universitas Negeri Yogyakarta (UNY), Indonesia (2019-now), Head of Instrumentation Laboratory (2020-now), Flying Robot UNY Team Supervisor (2019-now).

Anggun Winursito is a researcher with a focus on signal processing and instrumentation. He has received M.Eng. from Gadjah Mada University and B.Ed. from Yogyakarta State University in 2014 and 2018. His areas of expertise are speech recognition, image recognition, analog and digital signal processing, and instrumentation. He has participated in several local and national research projects including speech pattern recognition and radar telecommunications systems. He was awarded the best paper at an international conference held in Malaysia for his research on the compression of speech features to improve the accuracy of speech recognition systems. Currently, he is the lecturer at Department of electronic and informatic engineering education, Faculty of Engineering, Universitas Negeri Yogyakarta. He can be contacted at email: anggunwinursito@uny.ac.id.

Professor Teddy Surya Gunawan received his BEng degree in Electrical Engineering with cum laude award from Institut Teknologi Bandung (ITB), Indonesia, in 1998, and Ir in 2022. He obtained his M.Eng degree in 2001 from the School of Computer Engineering at Nanyang Technological University, Singapore, and a Ph.D. degree in 2007 from the School of Electrical Engineering and Telecommunications, The University of New South Wales, Australia. His research interests are speech and audio processing, biomedical signal processing and instrumentation, image and video processing, and parallel computing. He was awarded the Best Researcher Award in 2018 from IIUM. He is currently an IEEE Senior Member (since 2012), was chairman of IEEE Instrumentation and Measurement Society – Malaysia Section (2013 and 2014), Professor (since 2019), Head of Department (2015-2016) at Department of Electrical and Computer Engineering, and Head of Programme Accreditation and Quality Assurance for Faculty of Engineering (2017-2018), International Islamic University Malaysia. He has been a Chartered Engineer (IET, UK) since 2016 and Insinyur Profesional Utama (PII, Indonesia) since 2021, a registered ASEAN engineer since 2018, and ASEAN Chartered Professional Engineer since 2020.