International Journal of Electrical and Computer Engineering (IJECE)
Vol. 12, No. 5, October 2022, pp. 4926~4934
ISSN: 2088-8708, DOI: 10.11591/ijece.v12i5.pp4926-4934

A computationally efficient learning model to classify audio signal attributes

Maha Veera Vara Prasad Kantipudi[1], Satish Kumar[2]
[1]Department of Electronics and Telecommunication Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India
[2]Department of Robotics and Automation, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India

Article Info

Article history:

Received Jul 22, 2021
Revised May 19, 2022
Accepted Jun 12, 2022

Keywords:

Deep neural networks
Machine learning
Pulse audio signal
Signal processing

ABSTRACT

The era of machine learning has opened up groundbreaking realities and opportunities in the field of medical diagnosis. However, it is also observed that faster and proper diagnosis of any diseases/medical conditions require proper analysis and classification of digital signal data. It indicates the proper identification of tumors in the brain. Brain magnetic resonance imaging (MRI) data has to be appropriately classified, and similarly, pulse signal analysis is required to evaluate the human heart operating condition. Several studies have used machine learning (ML) modeling to classify speech signals, but very few studies have explored the classification of audio signal attributes in the context of intelligent healthcare monitoring. The study thereby aims to introduce novel mathematical modeling to analyze and classify synthetic pulse audio signal attributes with cost-effective computation. The numerical modeling is composed of several functional blocks where deep neural network-based learning (DNNL) plays a crucial role during the training phase, and also it is further combined with a recurrent structure of long-short term memory (R-LSTM) feedback connections (FCs). The design approaches further experiment in a numerical computing environment in terms of accuracy and computational aspects. The classification outcome of the proposed approach shows that it attains approximately 85% accuracy, which is comparable to the baseline approaches and execution time.

Corresponding Author:

Maha Veera Vara Prasad Kantipudi
Department of Electronics and Telecommunication Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University)
Pune 412115, India
Email: mvvprasad.kantipudi@gmail.com

1. INTRODUCTION

Since the last decade of research in audio technology has evolved up with various open directions. Moreover, there is a wide range of audio and speech signal processing applications, such as sensor-based speech processing, acoustic fingerprinting, and sound recognition. Apart from deriving 4-tuple aspects such as: i) storing audio data, ii) transmission of an audio data object, iii) capturing audio data, and iv) reconstruction of audio data signals, the conventional approaches in this technological advancement have found an immense scope to analyze the audio-related information and their meta-data very profoundly to have more potential insights [1]. The principle of audio signal classification in this regard has gained much more practical and theoretical values in the context of both pattern recognition and machine learning (ML) [2]. However, a clear view of the conventional research attempts reveals that applying and extending a supervised machine learning algorithm on speech signal processing algorithms poses a set of computational challenges during classification. The prime reason for this is that estimating signal labels from raw captured audio signal data is computationally challenging. However, training models based on neural networks (NN) play a crucial role in learning from in-depth audio embedded features [3]. The prime computational procedure to classify any audio signal attributes involves a stage of feature extraction where the extracted feature attributes (fA) are further explored to validate which class this fA belongs to. A gap exists in the research evolution of audio signal classification with ML approaches shows that relevant significant features from speech-based signals are well studied and less likely explored when other types of audio-based signals are concerned. It has to be considered that different types of audio signals pose distinct characteristic features. Thereby there is a notion of class-dependent feature analysis and study. Thus, it is essential to extract structured features with semantics, leading to proper deep processing of audio information required to construct an appropriate training model [4], [5]. The study introduces a novel analytical model that considers pulse audio data attributes and applies NN based learning model for computationally efficient and faster classification of data. The study, in this case, introduces a mathematical approach to construct the design of the neural network-based learning model and further apply it to the signal processing application to classify the discriminate features from the pulse audio signal. The training model is also validated in a numerical computing platform, considering different audio datasets corresponding to the pulse signals.

The overall theme of the formulated research manuscript is organized and presented for various sections. Section 2 represents the existing ML approaches deployed for audio signal classification; section 3 highlights the design methodology of the formulated system and the core backbone of workflows. Finally, section 4 talks about the numerical outcome, and section 5 illustrates the conclusion of the proposed research study.

This section introduces the conventional approaches that have used machine learning tools to correctly classify the audio signal (pulse-signal (pS)) discriminant features considering a spectrum analysis. The study [6] introduced an analytical approach based on decomposition and synthetic analysis, which further applied to the non-stationary audio signal for classification of its intrinsic features. The following are the steps summarized to depict the workflow of the presented approach, such as: i) the design analysis of the formulated approach comprises a set of functional modules where initially a pre-processing block is adopted to deal with non-stationary attributes of an audio signal, ii) it is also used to classify the features of the original signal in terms of energy and intrinsic based function, and iii) the process also further evaluates the sinusoidal parameters, which are further applied in audio synthesis.

The experimental outcome shows that the presented approach is practical for audio signal synthesis [7]. The study of [7] introduced an ML-based predictive approach to efficiently determine the perceived level of reverberation from the audio signal [7]. The architectural design of the proposed solution evaluates a class-level schema to validate the presented model under different types of audio sources. The outcome obtained shows that the ML-based trained model accurately predicts the perceptual score value [6].

Similar approaches also derived in the study of [8]–[12], where different ML approaches are used to classify the audio spectrum data. It is also observed that out of different approaches, NN-based learning approaches have been widely studied in audio signal attributes to deal with various synthesis and processing parameters. The cutting-edge conceptual modelings have provided a wide range of solutions in audio-data classification for different use-cases. It also presented NN based learning approach to speed up the process of audio synthesis by introducing a notion of interconnected, networked computational cells [13].

Similarly a new spectral estimation modeling is introduced considering radial basis function enabled NN methodology [14]. The study's prime aim was to classify the audio signal to recover the higher frequency (HF) component features. The Table 1 highlights a few relevant studies on audio signal processing, where NN approaches are widely used.

Table 1. Summary of relevant studies on audio signal classification using NN

Authors	Problem Labelled	Design Approach
Xu *et al.* [15]	Audio attribute tagging and classification	Recurrent convolutional NN learning approach for logMet audio spectrum classification
Kelz and Widmer [16]	Labeled noise estimation in the audio spectrum	Classification approach based on NN based learning and labeling
Başbuğ and Sert [17]	Scene classification in the audio spectrum	Long-short term memory (LSTM) architectural design
Garcia [18]	Detection of spectral peaks	Learning approach of frequency estimation

Other approaches have considered various NN based coding mode of selection approach to classifying the audio signal spectrum, such as the study of [19]–[21]. A few approaches have found their applicability in the speech audio spectrum classifications with in-depth features using recurrent convolutional NN approaches [22]–[26]. The studies of [27], [28] have a higher scope in audio signal classification and synthesis.

As highlighted in the prior section, a thorough background study of the research problem clearly shows that a wide range of research attempts are taken towards classifying different types of audio spectrum attributes using ML approaches. Still, most of the studies are limited to only speech signal processing applications. It is also found that despite various analytical solutions towards audio signal classification being designed using deep learning statistical modeling schema, a gap still exists due to the complexity and classification accuracy problems. Another problem in this broad area of application also shows that significantly less focus is laid towards the pulse-signal classification problem in the healthcare domain, which is crucial to making a proper patient diagnosis from a clinical viewpoint. Therefore, the problem statement of the study is derived: "*It is computationally challenging to design a conceptual model of learning approach based on LSTM architecture to classify the audio spectrum attribute with higher accuracy and by meeting the constraints of computational complexity aspects*." The subsequent sections will discuss the design approach of formulated conceptual design modeling of the pulse-audio classification model.

2. PROPOSED PROCEDURE

The prime aim of the formulated system is to classify the pulse audio signal attributes with the aid of both cost-effective computation and accuracy aspects. The system design and modeling corresponding to the formulated approach comprise a set of core functional blocks visually and combinedly represented in Figure 1. The core modeling of the system is constructed considering the functional module such as: pre-processing module *Pp(X)*, the feature extraction module *fe(X)*, and classification module *Cm(X)*. The connectivity among these three prime modules can be established with a notion of fundamental workflow:

$$Pp(X) \rightarrow fe(X) \rightarrow Cm(X).$$

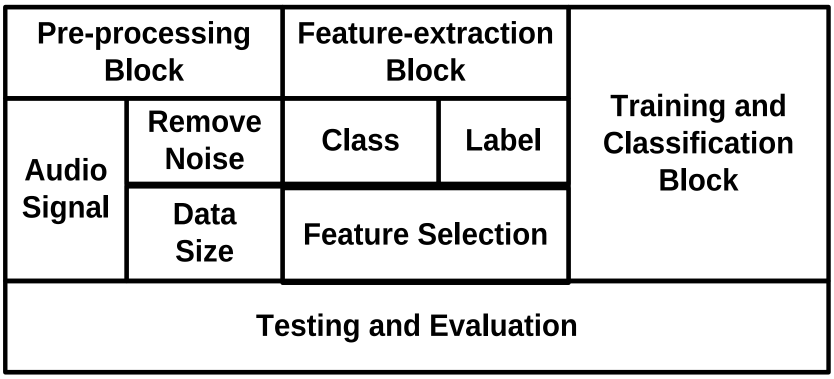


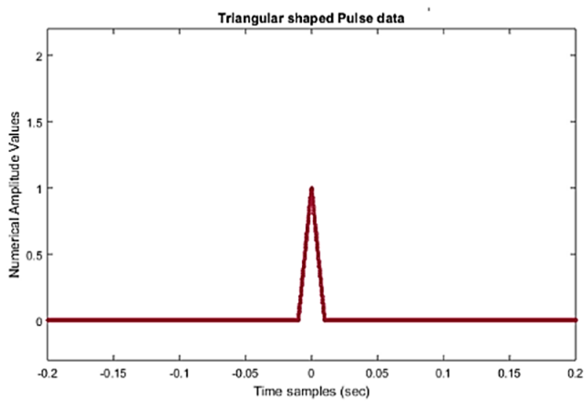
Figure 1. Functional block-based representation

The experimental pulse data set (*pData[]*) is generated using a numerical computing environment consisting of a set of pulse signals, as highlighted in Figure 2. The experimental approach can also be extended for another dataset [6] of pulse audio (heart beat-oriented) signal labeled feature attributes for the classification purpose. The system also considers novel data structuring operations on the *pData[]* computed frames from the files, and here each file is considered a specific period of seconds with sampling rate (*Sr*). The sampling rate here refers to the frame structuring values *(fs)* ∈ sfile of 1 sec. Here sfile refers to sound file object. The total frames in *pData[]* corresponding audio files can be computed with (1).

$$\mathrm{nfTot} = \mathrm{Sr} \times \mathrm{t} \tag{1}$$

The data structuring and framing operations here basically normalize the Sr for each data in *pData[]* also reduces the dimensionality factor in the sound signal wave, resulting in better execution time of the classifier and other involved procedures.

Numerically generated synthetic pulse-data signal type-1
TimeSamples (T_s): (-0.2 to 0.2)



Numerically generated synthetic pulse-data signal type-2
TimeSamples (T_s): (-0.2 to 0.4)

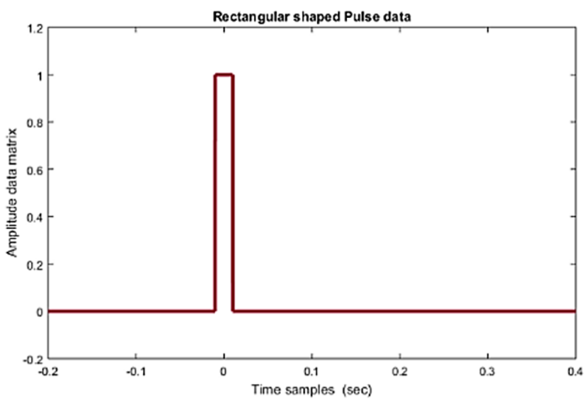


Figure 2. Synthetic pulse signal

3. RESEARCH METHOD

Initially, *pData[]* is divided into two sets of attributes, such as training attributes (*tA*) and testing and validation attribute (*teA*). The workflow further exhibits the segment-wise sequential execution model of the overall design architecture of the formulated conceptual model. The numerical simulation and formulation of the conceptual model initially consider two different types of pulse-audio data signal before performing classification, as highlighted:

– Design 1: $Pp(X) \rightarrow$ pre − processing functional block: This functional block enables pre-processing of tA and teA data where *tA*→ $[Class\ Label]$ this means in this supervised learning model, the audio signal tA is labeled for various classes for ease of extraction of features (fA). The tA and teA pulse signal attributes are initially undergone through a band-pass filter modeling to minimize noisy attributes. Also, further, it reduces the complexity of data by re-shaping the pulse-signal data considering the rate of frame (rF) instances by applying a lower-sampling approach. The Figure 3 shows the activity of execution of the formulated: $Pp(X)$ block.

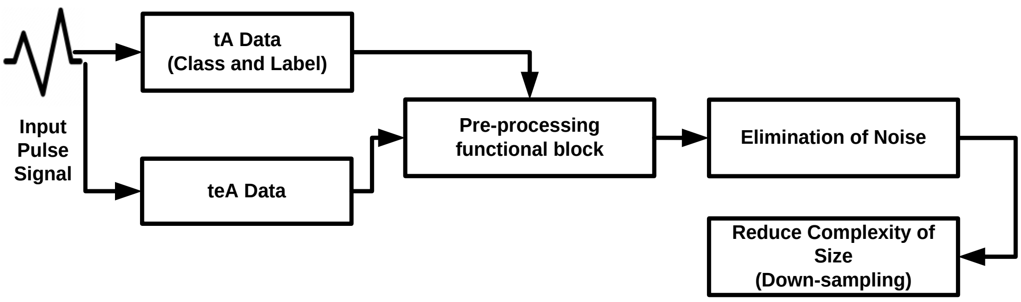


Figure 3. Functional backbone of pre-processing block

The input pulse signal p(t) cleans the undergoes through a transformation process to minimize the noise and eliminates the data redundancy by performing extraction of specific frequency labeled data. This phase also performs feature selection and extraction from the p(t) and performs dimensionality reduction concerning filtering. The transformation process can be mathematically realized.

$$p'(t) \leftarrow T(p(t)) \tag{2}$$

The process also applies lower-sampling approach modeling to set the exact frame rate adjustment. The process computational process applies a lower-sampling approach procedure for dimensionality reduction with an efficient feature selection process. The down-sampling procedure here helps deal with massive features in the audio signal data, which makes the computing process more efficient and robust. It applies a low-pass filter attribute on the data and covert approximately 30,000 fs and 765 fs which can also be expressed as normalized pulse signal attributes. The study adopted the methodical philosophy adopted in [29] and [30], which enables the functional module $fe(X)$. The lower-sampling approach can be mathematically expressed:

$$p'(t) = \sum \frac{p(t)}{\max(p(t)} \tag{3}$$

Here $p'(t)$ denotes the normalized pulse signal.

– Design 2: $Cm(X)$ →training and classification module: This functional module is designed for two prime functional blocks such as i) training block and ii) testing block. The Figure 4 shows the core components of the formulated system where LSTM based recurrent neural network-enabled learning is utilized for deep pulse audio feature classification.

Figure 4 shows how the learning model of the formulated concept is designed considering the LSTM reference recurrent NN architectural design [31]–[33]. The training data set is pre-processed to minimize the complexity and noise associated with pulse-audio data attributes. Further lower-sampling approach techniques also perform filtering of specific frequency attributes for feature selection and extraction process. The extracted labeled features of different classes are further used to train the LSTM NN model to classify the audio signal intrinsic in-depth features better. The LSTM reference NN architecture consists of different prime gateways such as iG, oG, and fG. These prime attributes are used for reading, writing, and reset computational operations.

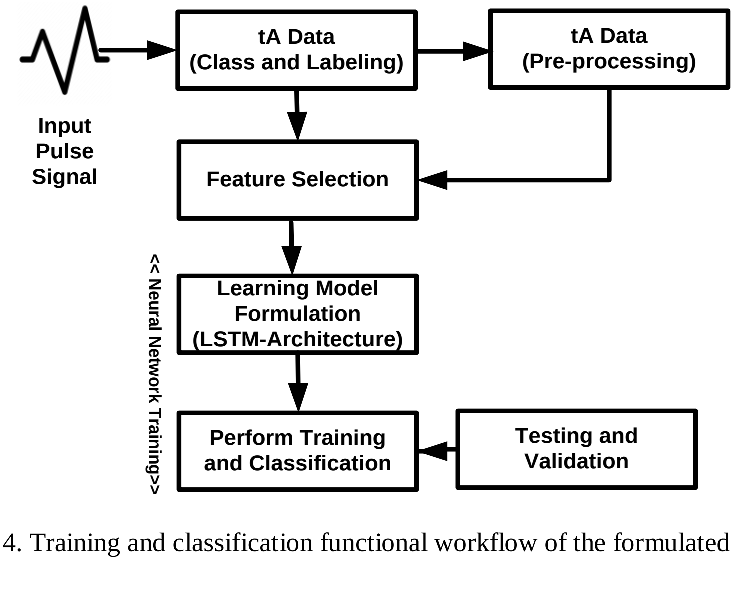


Figure 4. Training and classification functional workflow of the formulated concept

$$fG \leftarrow Sig(w1 \times c1 + h(t-1)_{fG} + bV_{fG}) \tag{4}$$

$$iG \leftarrow Sig(w2 \times c2 + h(t-1)_{iG} + bV_{iG}) \tag{5}$$

$$oG \leftarrow Sig(w3 \times c3 + h(t-1)_{oG} + bV_{oG}) \tag{6}$$

$$compute \rightarrow c(s) = fG \times c(s-1) + iG \times hyper(W \times c(t) + h(t-1) + bV(c) \quad (7)$$

$$h(t) \rightarrow oG \times hyper(C(s)) \quad (8)$$

The equations (3) to (4) shows how LSTM neural network modeling is utilized here where a function sigmoid sig is used for different operational attributes such as weight (w), coefficient C, hidden layer state h(t), and a bias vector b. The computation of cell state vector c(s) also utilized hyperbolic hyper (X). Along with the Input layer, the reference architecture of LSTM also used a dense layer and softmax layer during the classification and training. The reference model of LSTM contains output height of 1 along with output width 782 and output depth 64. The Figure 5 shows the testing module of LSTM based audio signal classification. The accuracy performance is evaluated during the classification prediction stage, and also the outcome of both computation and accuracy is further validated for comparative performance analysis, as shown in the next section.

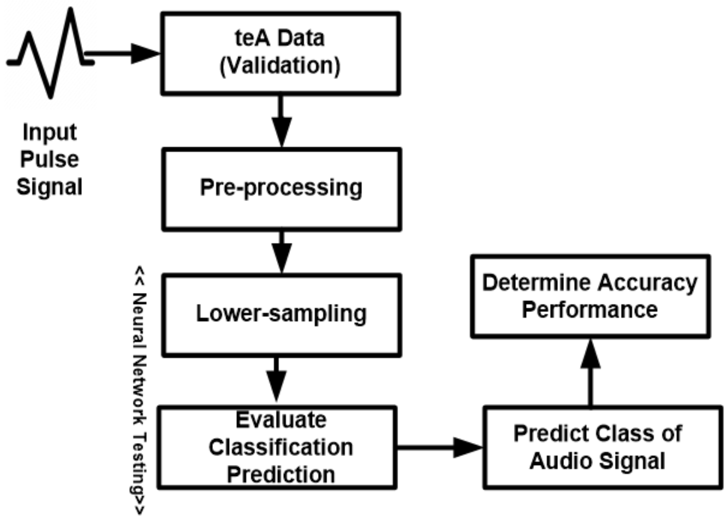


Figure 5. Testing module of LSTM based audio signal classification

4. RESULT AND DISCUSSION

This section talks about the outcome obtained after simulating the numerical modeling of the learning approach for audio classification. This phase of the research manuscript discusses the validation outcome of the classification prediction accuracy of the formulated conceptualized modeling. The design model is simulated under MATLAB numerical computing environment supported with system type 64-bit operating system, x64-based processor, 4 GB RAM, and 2.00, 1.99 GHz processing speed.

The dataset corresponds to the pulse signal [6] consists of 30,000 frames and a time of 12.34 secs. From this dataset, the training data and data for validation are programmatically generated in synthetic form. The analytical system design is simulated with respect to a set of operational constraints, and the operating frequency of input synthetic audio signal is considered to be in a range of 55-800 Hz. The validation of the prediction accuracy is performed by comparing the classification accuracy score with three other types of frequently adopted machine learning models, such as SVM, decision tree (DT), and random forest (RF). During the training and validation phase, the hyperparameters consider dropout rates ranging between (0.05-0.25). It results in an accuracy of 77% and 82.1%, with a loss of 48.2 and 47.65. The Figure 6 shows that the formulated conceptualized modeling attain better validation performance in classification accuracy, which is ~85% and superior to other learning models.

The prime reason for obtaining this outcome is that LSTN based NN models apply better learning from the labeled features, considering deep feature extraction from the synthetic audio signal data. There are various performance metrics to evaluate the classification model's performance, such as accuracy, precision, recall, and sensitivity. However, the proposed solution computes the accuracy performance (Ap) for true positive (tP), true negative (tN), false positive (fP), and false negative (fN).

$$Ap \leftarrow (tP + tN)/(tP + tN + fP + fN) \quad (9)$$

The formulated approach applies the dimensionality reduction process of data and a filtering approach to make the data more suitable for the classification model. Thereby the computational time complexity and memory constraints are also significantly reduced. The validation outcome also shows that for ten epochs, the formulated approach attains a processing time of 0.0879 sec and 0.2124 sec. of execution time, comparable to the existing baselines. In random forest approach the processing time is found 0.1234 sec where as in the case of support vector machine (SVM) and DT the execution time is approximately 0.78 secs and 0.034 secs. The study also refers to the method introduced in [32], [34] to overcome overfitting issue in LSTM and NN based solutions.

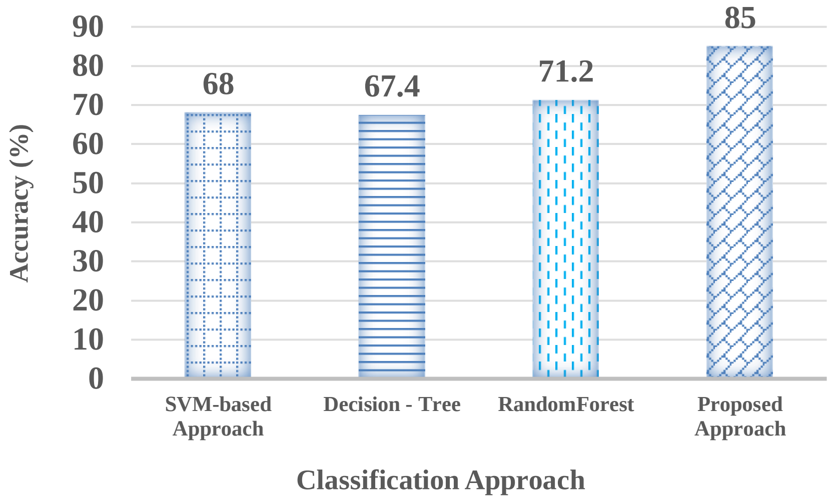


Figure 6. Analysis of classification accuracy

5. CONCLUSION

The study presented a novel learning model that adopts the reference architecture of LSTM to classify pulse-audio synthetic data. The methodology constructed also considers hypothetical factors by justifying their practicability into modern healthcare diagnosis. The computational analysis poses robustness by differing the training ratio and shows that the numerical computation's computational time complexity is significantly reduced. The comparative performance analysis and the quantified outcome show that the proposed approach attains better classification accuracy than the existing solutions. The system does not effectively work with the spectrogram technique on computing more distinctive features from pulse signal attributes. The limitation of the study is that it has not assessed the false positive and negative scores for the proposed LSTM based learning model. However, it anticipates its scope in future innovative healthcare applications in the context of pulse-data monitoring systems.

ACKNOWLEDGEMENTS

This research was funded by Symbiosis International University (SIU) under the Research Support Fund.

REFERENCES

[1] M. Karjalainen, "Immersion and content-a framework for audio research," in *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. WASPAA'99 (Cat. No.99TH8452)*, 1999, pp. 71–74, doi: 10.1109/ASPAA.1999.810852.

[2] F. Rong, "Audio classification method based on machine learning," in *2016 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, Dec. 2016, pp. 81–84, doi: 10.1109/ICITBS.2016.98.

[3] J. Cramer, H.-H. Wu, J. Salamon, and J. P. Bello, "Look, listen, and learn more: design choices for deep audio embeddings," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 3852–3856, doi: 10.1109/ICASSP.2019.8682475.

[4] Z. Nina, S. Wee, Y. Zhuliang, Y. Jufeng, and C. Huawei, "Enhanced class-dependent classification of audio signals," in *2009 WRI World Congress on Computer Science and Information Engineering*, 2009, pp. 100–104, doi: 10.1109/CSIE.2009.664.

[5] D. Wu, "An audio classification approach based on machine learning," in *2019 International Conference on Intelligent Transportation, Big Data and Smart City (ICITBS)*, Jan. 2019, pp. 626–629, doi: 10.1109/ICITBS.2019.00156.

[6] X. ming Li, C. chun Bao, and M.-shen Jia, "A sinusoidal audio and speech analysis/synthesis model based on improved EMD by adding pure tone," in *2011 IEEE International Workshop on Machine Learning for Signal Processing*, Sep. 2011, pp. 1–5, doi: 10.1109/MLSP.2011.6064614.

[7] S. Safavi, A. Pearce, W. Wang, and M. Plumbley, "Predicting the perceived level of reverberation using machine learning," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, Oct. 2018, pp. 27–30, doi: 10.1109/ACSSC.2018.8645201.

[8] J.-S. Liang and K. Wang, "Vibration feature extraction using audio spectrum analyzer based machine learning," in *2017 International Conference on Information, Communication and Engineering (ICICE)*, Nov. 2017, pp. 381–384, doi: 10.1109/ICICE.2017.8479273.

[9] H. Phan, L. Hertel, M. Maass, R. Mazur, and A. Mertins, "Learning representations for nonspeech audio events through their similarities to speech patterns," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 4, pp. 807–822, Apr. 2016, doi: 10.1109/TASLP.2016.2530401.

[10] T. Li and G. Tzanetakis, "Factors in automatic musical genre classification of audio signals," in *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (IEEE Cat. No.03TH8684)*, 2003, pp. 143–146, doi: 10.1109/ASPAA.2003.1285840.

[11] K. Qian, Z. Xu, H. Xu, and B. P. Ng, "Automatic detection of inspiration related snoring signals from original audio recording," in *2014 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)*, Jul. 2014, pp. 95–99, doi: 10.1109/ChinaSIP.2014.6889209.

[12] S. Karthik, S. Kumar, K. M. V. V Prasad, K. Mysurareddy, J. M. C, and B. D. Seshu, "Automated home-based physiotherapy," in *2020 International Conference on Decision Aid Sciences and Application (DASA)*, 2020, pp. 854–859, doi: 10.1109/DASA51403.2020.9317247.

[13] Z. Baracskai, "DANN: digital audio neural network," in *2019 4th International Conference on Smart and Sustainable Technologies (SpliTech)*, Jun. 2019, pp. 1–4, doi: 10.23919/SpliTech.2019.8783027.

[14] J. P. Dominguez-Morales *et al.*, "Deep spiking neural network model for time-variant signals classification: a real-time speech recognition approach," in *2018 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2018, pp. 1–8, doi: 10.1109/IJCNN.2018.8489381.

[15] Y. Xu, Q. Kong, W. Wang, and M. D. Plumbley, "Large-scale weakly supervised audio classification using gated convolutional neural network," in *2018 IEEE International Conference on Acoustics, Speech And Signal Processing (ICASSP)*, 2018, pp. 121–125.

[16] R. Kelz and G. Widmer, "Investigating label noise sensitivity of convolutional neural networks for fine grained audio signal labelling," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2018, pp. 2996–3000, doi: 10.1109/ICASSP.2018.8461291.

[17] A. M. Basbug and M. Sert, "Analysis of deep neural network models for acoustic scene classification," in *2019 27th Signal Processing and Communications Applications Conference (SIU)*, Apr. 2019, pp. 1–4, doi: 10.1109/SIU.2019.8806301.

[18] G. Garcia, "Estimation of sinusoids in audio signals using an analysis-by-synthesis neural network," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)*, 2001, vol. 5, pp. 3369–3372, doi: 10.1109/ICASSP.2001.940381.

[19] M.-L. Wang and M.-T. Lee, "A neural network based coding mode selection scheme of hybrid audio coder," in *2010 IEEE International Conference on Wireless Communications, Networking and Information Security*, Jun. 2010, pp. 107–110, doi: 10.1109/WCINS.2010.5541899.

[20] S. K. H. N. M. Asif, "A unified approach using neural networks efficient algorithms in audio signal processing," in *8th International Multitopic Conference, 2004. Proceedings of INMIC 2004.*, 2004, pp. 26–31, doi: 10.1109/INMIC.2004.1492841.

[21] S. Soni, S. Dey, and M. S. Manikandan, "Automatic audio event recognition schemes for context-aware audio computing devices," in *2019 Seventh International Conference on Digital Information Processing and Communications (ICDIPC)*, May 2019, pp. 23–28, doi: 10.1109/ICDIPC.2019.8723713.

[22] S. Hershey *et al.*, "CNN architectures for large-scale audio classification," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2017, pp. 131–135, doi: 10.1109/ICASSP.2017.7952132.

[23] M. P. Kantipudi, C. John Moses, R. K. Aluvalu, and G. T. Goud, "Impact of Covid-19 on Indian higher education," in *Library Philosophy and Practice*, vol. 2021, 2021.

[24] Y. Li, X. Li, Y. Zhang, W. Wang, M. Liu, and X. Feng, "Acoustic scene classification using deep audio feature and BLSTM network," in *2018 International Conference on Audio, Language and Image Processing (ICALIP)*, Jul. 2018, pp. 371–374, doi: 10.1109/ICALIP.2018.8455765.

[25] K. Xu *et al.*, "General audio tagging with ensembling convolutional neural networks and statistical features," *The Journal of the Acoustical Society of America*, vol. 145, no. 6, pp. EL521--EL527, Jun. 2019, doi: 10.1121/1.5111059.

[26] G. Keren and B. Schuller, "Convolutional RNN: An enhanced model for extracting features from sequential data," in *2016 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2016, pp. 3412–3419, doi: 10.1109/IJCNN.2016.7727636.

[27] D. Subbarao, M. V. V. P. Kantipudi, M. A. Kumar, and D. Chandra, "Robust, knowledge-based, robust models," *International Journal of Computer Science and Technology*, vol. 2, no. 1, 2011.

[28] K. M. V. V Prasad and H. N. Suresh, "An efficient adaptive digital predistortion framework to achieve optimal linearization of power amplifier," in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Mar. 2016, pp. 2095–2101, doi: 10.1109/ICEEOT.2016.7755058.

[29] J. Díaz-García, P. Brunet, I. Navazo, and P.-P. Vázquez, "Downsampling methods for medical datasets," in *International Conferences Computer Graphics, Visualization, Computer Vision and Image Processing 2017 and Big Data Analytics, Data Mining and Computational Intelligence 2017*, 2017, pp. 12–20.

[30] M. Genussov and I. Cohen, "Musical genre classification of audio signals using geometric methods," in *European Signal Processing Conference*, 2010, pp. 497–501.

[31] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Networks*, vol. 18, no. 5–6, pp. 602–610, Jul. 2005, doi: 10.1016/j.neunet.2005.06.042.

[32] U. V Maheswari, R. Aluvalu, and K. Keerthi Chennam, "Chapter 5 Application of machine learning algorithms for facial expression analysis," in *Machine Learning for Sustainable Development*, De Gruyter, 2021, pp. 77–96.

[33] M. V. V. P. Kantipudi, S. Vemuri, and V. K. Sanipini, "Deep learning-based image super-resolution algorithms-a survey," *International Journal of Computing and Digital Systems*, vol. 11, no. 1, pp. 413–421, Jan. 2022, doi: 10.12785/ijcds/110134.

[34] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

BIOGRAPHIES OF AUTHORS

Maha Veera Vara Prasad Kantipudi is working as an Associate Professor in the Department of E and TC, Symbiosis Institute of Technology, Pune. He received his B.Tech. (Electronics and Communications) (2009) and MTech. (Digital Electronics and Communication Systems) (2011) degrees from Jawaharlal Nehru Technological University, Kakinada. He received his Ph.D. (Signal Processing specialization) from BITS, VTU, Belagavi (2018). He, previously, worked as the Director of Advancements for Sreyas Institute of Engineering and Technology, Hyderabad, and also as an Associate Professor with R.K. University, Rajkot. He is having teaching experience of around 10.8 years. His current research interests are in Signal Processing with Machine Learning, Education and Research. He is recognized as a technical resource person for Telangana state by the IIT Bombay Spoken tutorial team. He conducted key Training Workshops on Open-Source Tools for education, Signal Processing and Machine Learning focused topics, and Educational Technology. He has authored and co-authored many papers in International Journals, International Conferences, National Conferences and published five Indian Patents. Prasad is a Senior Member of IEEE (Membership ID: #93513961) and an active member of Machine Intelligence Research Labs and USERN (Universal Scientific Education and Research Network) (April 2020-present). He is one of the active reviewers for wireless networks, Journal of Springer Nature. His name is listed at 19th position in Top 100 Private University's Authors Research Productivity Rankings given by the Confederation of Indian Industry (CII) based on the "Indian Citation Index" Database 2016. He can be contacted at email: prasadb2016@gmail.com.

Satish Kumar is Assistant Professor in the Department of Mechanical Engineering at Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune India. He did his Master degree (MTech.) and Doctoral degree (Ph.D.) in 2013 and 2020 from K Visvesvaraya Technological University, Belgaum, Karnataka, India. He has 8 years of experience in teaching, research and industries. His area of research interests includes Smart Manufacturing, Condition Monitoring, Composites, Cryogenic Treatment, Additive manufacturing and Hard materials machining. He has authored more than 24international/national journal and conferences publications. According to Google Scholar, he has 120+ citations, with an H-index of 7 and an i10-index of 4. He can be contacted at email: satish.kumar@sitpune.edu.in.