

Smart chatbot for surveys by convolutional networks speech recognition

Robinson Jimenez-Moreno¹, Javier Eduardo Martínez Baquero², Luis Alfredo Rodríguez Umaña²

¹Department of Mechatronic Engineering, Universidad Militar Nueva Granada, Bogotá, Colombia

²Engineering School, Faculty of Basic Sciences and Engineering, Universidad de los Llanos, Villavicencio, Colombia

Article Info

Article history:

Received Aug 8, 2024

Revised Jan 3, 2025

Accepted Mar 3, 2025

Keywords:

Chatbot

Convolutional network

Database

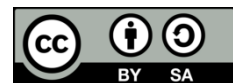
Deep learning

Voice selection

ABSTRACT

This paper details the development of an innovative voice chatbot interface specifically designed for evaluating user options using a Likert scale by color. The core of this interface is designing a convolutional neural network architecture, which has been trained with MEL spectrogram inputs from seven possible words for each answer. These spectrograms are crucial in capturing the audio features necessary for effective voice recognition and establishing the interactions that occur between the chatbot and the user, allowing the convolutional network to learn and distinguish between different types of user responses accurately. During the training phase, the convolutional neural network achieved an accuracy rate of 91.4%, indicating its robust performance in processing and interpreting voice commands. The interface was tested in a controlled environment, with a group of ten users and a survey of 5 questions, where it achieved a perfect detection accuracy of 100%. The results demonstrate the system's capacity for natural user interaction by voice and employing a free text to speech (TTS) algorithm for the chatbot voice.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Robinson Jiménez Moreno

Mechatronics Engineering Program, Faculty of Engineering, Universidad Militar Nueva Granada

Carrera 11 # 101-80, Bogotá D.C., Colombia

Email: robinson.jimenez@unimilitar.edu.co

1. INTRODUCTION

User perception systems often rely on surveys using the Likert scale [1] to gauge responses. Recent advancements in artificial intelligence have enhanced the application of this scale by incorporating intelligent algorithms [2]. These developments enable the integration of sophisticated tools that facilitate the creation of intelligent chatbots, thereby revolutionizing the way user feedback is collected and analyzed. Among the primary beneficiaries of these innovations is the elderly population. Intelligent chatbots [3] can significantly improve the survey experience for older adults by providing a more intuitive and accessible interface, which is particularly important given the challenges they may face with traditional survey methods. This integration not only streamlines the data collection process but also ensures that the feedback is more accurately and efficiently gathered, ultimately leading to better insights and more informed decision-making. The fusion of the Likert scale with artificial intelligence (AI)-driven technologies thus represents a significant leap forward in enhancing user perception systems, offering a more tailored and user-friendly approach for diverse demographics, especially the elderly [4].

The development of chatbots capable of natural voice interaction is achievable using advanced techniques like convolutional networks [5], [6]. These networks excel in processing and interpreting voice data, enabling the creation of chatbots that can understand and respond to spoken language with high accuracy. Text classification is an important tool in the development of chatbots, which can be carried out with AI techniques

[7], [8], Text classification is an important tool in the development of chatbots, which can be carried out with AI techniques. Where its integration with voice recognition tools would enhance its application in natural interfaces.

The applications of voice recognition [9] through convolutional networks extend far beyond basic interaction. For instance, this technique can be employed for emotion recognition, allowing systems to detect and respond to the emotional tone of a user's voice, thereby enhancing the user experience. Additionally, voice recognition can be utilized for controlling various technological devices, providing a hands-free and intuitive way to interact with smart home systems, mobile applications, and other digital interfaces.

Another significant advantage of convolutional networks in voice recognition is their potential for integration into portable systems. This capability makes it possible to embed sophisticated voice recognition functionalities into compact, mobile devices, broadening the scope of their application and accessibility. Whether for personal use, healthcare, or smart environments, the portability of these systems ensures that advanced voice interaction features are readily available in a wide range of contexts [10]. Thus, the use of convolutional networks in developing natural voice interaction chatbots represents a versatile and powerful tool in modern technology [11].

The use of chatbots nowadays supports a wide variety of applications in fields such as intelligent agriculture [12], support for students on career guidance [13], customer service [14] and the tourism sector [15]. Where in these cases a chatbot is employed as an initial interaction in the perception of a social robot [16], which improve their interactivity when integrated with deep learning techniques [17], and with inclusions that can improve the quality of life in smart homes [18]. However, there is a clear need for natural interaction tools both inbound and outbound, i.e., voice interaction in both directions.

This article presents the development of a voice chatbot interface oriented to evaluate the perception of a user through a survey based on a minimum Likert scale [19]. To do this, a convolutional neural network architecture [20] is proposed which is trained to recognize specific words to answer the survey by color choices. The paper proposes a complement to the developments exposed in the state of the art, developing a basic interface which uses a text-to-speech (TTS) algorithm to play a specific script of a five-question survey to a group of stakeholders. Where the user will also respond by voice within the response range based on red (disagree), yellow (partially agree) or green (strongly agree) options. Thus facilitating technological interaction with ChaBots and applications in industrial [21] and home [22], [23], with human-robot work environments.

The structure of the paper is organized to provide a comprehensive overview of the research and its findings. The present introductory section outlines the state of the art in the field and highlights the specific contribution of this research. Following this, section 2 details the methodology and development environment used in the study, offering insights into the tools, techniques, and processes employed. In section 3, the results obtained from the research are thoroughly analyzed, providing a critical examination of the data and findings. Finally, section 4 presents the conclusion, summarizing the key outcomes of the study and discussing the implications. This section also outlines the associated future work, suggesting areas for further research and potential improvements to the current study. This structured approach ensures that the paper covers all essential aspects of the research, from the initial context and methodology to the final analysis and future directions.

2. METHOD

To achieve the goal of implementing a voice chatbot for evaluating survey-type services, a 16-layer convolutional network [23] is trained. The architecture of this network is detailed in Figure 1. Training parameters for the network are outlined in Table 1 [24]. This setup aims to ensure the chatbot effectively processes and evaluates voice inputs to provide accurate survey responses.

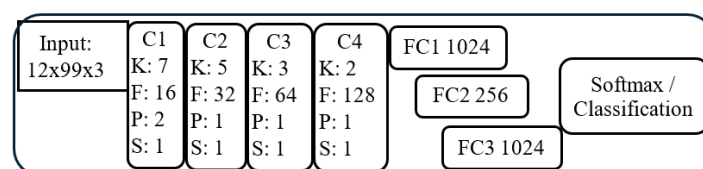


Figure 1. CNN architecture

The network is trained using a comprehensive database of voice recordings that align with the expected responses during interactions with the chatbot. Specifically, the system is designed to interpret binary responses, such as "yes" or "no". In addition, for user evaluations, the system employs a traffic light system, in

which responses are classified as green, yellow or red. These colors correspond to the user's positive, neutral or negative perception of the question asked or can be interpreted as red for disagreeing, yellow for partially agreeing or green for strongly agreeing with the specific evaluation.

Table 1. CNN training characteristics

Hype parameters	Value
Minibatch size	5
Optimizer	SGDM
Epochs	150
Learn rate	1e-6
Frequency	940
Iteration	14,700

The dataset utilized for training consists of 700 samples per MEL spectrogram [24], ensuring a robust representation of voice data, with 5 test users pronouncing the 7 different words and 100 samples per word. To optimize the training process, 70% of the samples for each class are allocated for training the network, while the remaining 30% are reserved for validation purposes. Figure 2 illustrates an example of the network input, with dimensions 12×9 and three channels corresponding to the MFCC of the input word, the derivative of the MFCC and its second derivative.

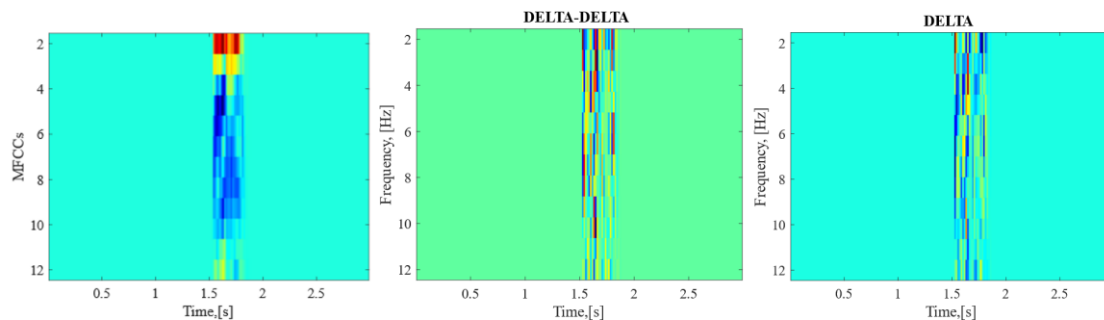


Figure 2. MEL coefficients

This approach, as shown in Figure 3, helps in fine-tuning the network's accuracy and reliability. By dividing the data in this manner, the system is better able to generalize from the training examples and perform effectively in real-world scenarios. This rigorous training protocol ensures that the network can accurately interpret and respond to a wide range of user inputs.

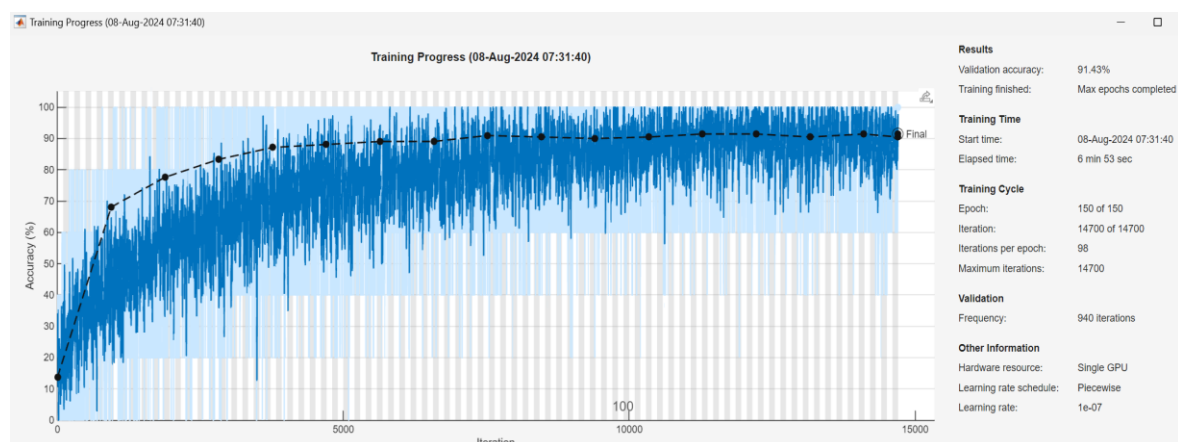


Figure 3. Network learning graph

Figure 4 illustrates the confusion matrix obtained [25]–[28], where a high level of accuracy of 91.4% is observed. The class that presents the highest error is 'yellow' being confused with the words finish and no. The network finally learns 554 thousand parameters and takes a classification time of 0.78 seconds.

		Confusion Matrix						
Output Class	red	29 13.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	yellow	0 0.0%	25 11.9%	2 1.0%	0 0.0%	0 0.0%	0 0.0%	92.6% 7.4%
	green	0 0.0%	0 0.0%	26 12.4%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	finish	0 0.0%	4 1.9%	2 1.0%	29 13.8%	2 1.0%	5 2.4%	69.0% 31.0%
	repeat	0 0.0%	0 0.0%	0 0.0%	0 0.0%	28 13.3%	0 0.0%	100% 0.0%
	no	0 0.0%	1 0.5%	0 0.0%	1 0.5%	0 0.0%	25 11.9%	92.6% 7.4%
	yes	1 0.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	96.8% 3.2%
		96.7% 3.3%	83.3% 16.7%	86.7% 13.3%	96.7% 3.3%	93.3% 6.7%	83.3% 16.7%	100% 0.0%
		Target Class						
		red	yellow	green	finish	repeat	no	yes

Figure 4. Speech recognition confusion matrix

3. RESULTS AND DISCUSSION

The interface is designed in MATLAB's App Designer, as shown in Figure 5, where it is graphically shown how to answer the questionnaire to be applied, by means of a color. When pressing the start key, an audio is played, explaining that each question will be answered with a color depending on whether one agrees (green), disagrees (red) or a neutral position with yellow, based on a Likert scale. The interface has a box to indicate the number of the question that is being asked in sequence and another one to visualize it, however, the audio of the current question is also played.

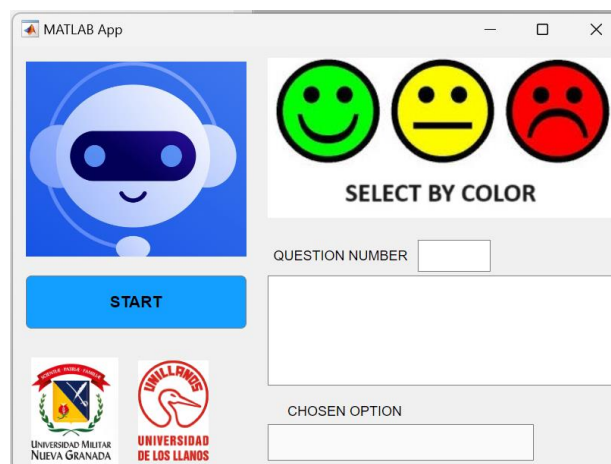


Figure 5. Developed interface

The bot's chosen and interpreted response is displayed in a text box located at the end of the app. Additionally, the system includes a feature that allows the survey to be terminated if the user says "end." At the start of the survey, the app provides an initial explanation of how the questions operate. If the user indicates

that the explanation is unclear, they have the option to request a repetition of the instructions. This ensures that users fully understand how to engage with the survey and can easily conclude the session if needed.

For validation purposes, a group of five generic questions was established as follows:

- How do you consider the attention received?
- Do you consider that the waiting time was adequate?
- Would you request this service again?
- The facilities are to your liking?
- Do you find this survey easy to answer?

Figure 6 shows the interface in operation. In this case it is not of interest to evaluate the Likert scale against the answers, but to validate that the interface was operational, for which no different problems are evidenced since this development phase operates in a controlled environment, giving 100% effectiveness in the test environment with a group of ten users. When evaluating a semi-noisy environment (people chatting in the background) there is evidence of class confusion, so the level of recognition reduces the effectiveness of the application. In low ambient noise conditions users reported a user-friendly facial interface, however they indicated that the semi-robotic voice employed by the TTS could be improved.

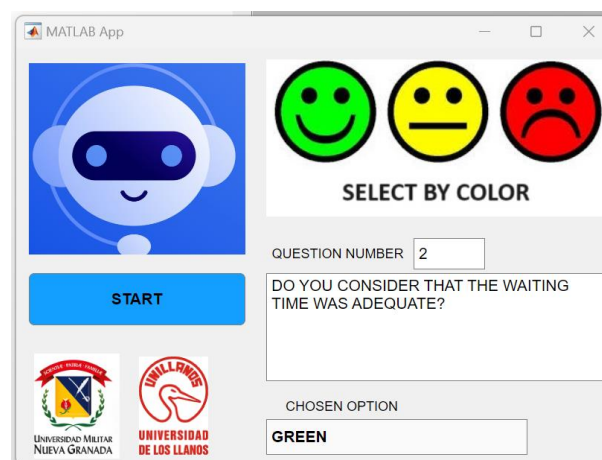


Figure 6. Interface in operation

For future work, it is proposed to undertake a comprehensive evaluation of various types of convolutional neural networks to identify the most suitable option for embedding in a portable system. This involves experimenting with different network architectures, hyperparameters, and training techniques to determine which configuration offers the best balance between performance, efficiency, and resource utilization. Additionally, an important aspect of this future work will be to extend the current selection scale, which is presently limited, to encompass five possible responses. This expansion aims to provide a more nuanced and detailed framework for user feedback, enhancing the system's capability to capture a broader range of user sentiments and preferences. By integrating a more sophisticated network and extending the response scale, the goal is to significantly improve the system's versatility, accuracy, and user satisfaction in practical, portable applications. This dual approach will help ensure that the final product is both highly effective and user-friendly. More complex, multi-word responses can be handled by other types of associative networks, e.g., long short-term memory (LSTM) [29]–[33].

4. CONCLUSION

The implementation of a voice survey system significantly enhances user interaction, making it more natural and intuitive. Using artificial intelligence techniques grounded in deep learning, this system achieves efficient functional development. Such advancements in technology, as demonstrated in this case, improve user-friendly interactions, particularly benefiting elderly individuals. By leveraging deep learning, the voice survey system can better understand and respond to user inputs, ensuring a smoother and more accessible experience for all users, especially those who may face challenges with traditional interfaces. This approach not only streamlines the interaction process but also promotes inclusivity and ease of use, showcasing the potential of AI-driven solutions in improving technological accessibility and user engagement.

The operation of the system in both controlled environments and noisy conditions highlights the crucial need for implementing noise suppression filters in the final interface. During testing, it became evident that the presence of ambient noise adversely affects the system's performance and accuracy. This observation underscores the importance of incorporating robust noise suppression mechanisms to ensure the interface functions effectively in diverse real-world settings. Addressing this aspect is essential for enhancing the overall reliability and user experience of the system. Consequently, this issue will be a primary focus in future work, where we will evaluate and integrate advanced noise suppression techniques. By doing so, we aim to refine the interface, making it more resilient to background noise and thereby improving its practical applicability and performance in everyday use scenarios.

ACKNOWLEDGMENTS

The authors would like to thank the Universidad Militar Nueva Granada where they are associate professors. The authors would like to thank the Universidad de los Llanos where they are associate professors.

FUNDING INFORMATION

Product derived from the research project titled “Design of a human-robot interaction model using deep learning algorithms” INV-ING-3971 financed by the vice-rector for research of the Universidad Militar Nueva Granada, year 2024.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration. The following table describes the contribution by author.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Robinson Jimenez-Moreno	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓
Javier Eduardo Martínez Baquero		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Luis Alfredo Rodríguez Umaña		✓		✓	✓	✓	✓			✓	✓	✓	✓	

C : Conceptualization	I : Investigation	Vi : Visualization
M : Methodology	R : Resources	Su : Supervision
So : Software	D : Data Curation	P : Project administration
Va : Validation	O : Writing - Original Draft	Fu : Funding acquisition
Fo : Formal analysis	E : Writing - Review & Editing	

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

Not applicable as it requires the involvement of personnel from outside the work team, no sensitive information was handled.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, R.J.M., upon reasonable request.

REFERENCESs




[1] T. A. Malapane and N. K. Ndlovu, “Assessing the reliability of likert scale statements in an E-commerce quantitative study: A cronbach alpha analysis using SPSS statistics,” in *2024 Systems and Information Engineering Design Symposium (SIEDS)*, May 2024, pp. 90–95, doi: 10.1109/SIEDS61124.2024.10534753.

- [2] K. Memmedova and B. Ertuna, "Development of a fuzzy Likert scales to measure variables in social sciences," *Information Sciences*, vol. 654, Jan. 2024, doi: 10.1016/j.ins.2023.119792.
- [3] K. Balaji, S. Karim, and P. S. Rao, "Unleashing the power of smart chatbots: Transforming banking with artificial intelligence," in *2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, May 2024, pp. 1–7, doi: 10.1109/ACCAI61061.2024.10602456.
- [4] L. Küntzer, S. Schwab, H. Spaderna, and G. Rock, "Measuring user experience of older adults during virtual reality exergaming," in *2024 16th International Conference on Quality of Multimedia Experience (QoMEX)*, Jun. 2024, pp. 153–159, doi: 10.1109/QoMEX61742.2024.10598263.
- [5] A. S. Shinde, A. S. Jamdar, K. D. Joshi, and S. T. Sarode, "A CNN based speech recognition approach for voice controlled elevator," in *2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT)*, Dec. 2021, pp. 728–733, doi: 10.1109/ICEECCOT52851.2021.9707928.
- [6] M. K. Muchamad, Z. Fuadi, and N. Nasaruddin, "Prototype design of deep learning-based voice control model for smart home," in *2022 IEEE International Conference on Internet of Things and Intelligence Systems (IoT&IS)*, Nov. 2022, pp. 25–30, doi: 10.1109/IoT&IS56727.2022.9975901.
- [7] V. Dutt, S. M. Sasubilli, and A. E. Yerrapati, "Dynamic information retrieval with chatbots: A review of artificial intelligence methodology," in *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Nov. 2020, pp. 1299–1303, doi: 10.1109/ICECA49313.2020.9297533.
- [8] Y. Su, Y. Huang, and C.-C. J. Kuo, "Efficient text classification using tree-structured multi-linear principal component analysis," in *24th International Conference on Pattern Recognition (ICPR)*, Aug. 2018, pp. 585–590, doi: 10.1109/ICPR.2018.8545832.
- [9] H. Avula, R. Ranjith, and S. P. Anju, "CNN based recognition of emotion and speech from gestures and facial expressions," in *2022 6th International Conference on Electronics, Communication and Aerospace Technology*, Dec. 2022, pp. 1360–1365, doi: 10.1109/ICECA55336.2022.10009316.
- [10] K. Dharavath, G. M. Kumar, K. R. Reddy, and M. H. Reddy, "Gesture and voice controlled virtual mouse for elderly people," in *2024 2nd International Conference on Networking and Communications (ICNWC)*, Apr. 2024, pp. 1–5, doi: 10.1109/ICNWC60771.2024.10537425.
- [11] M. I. Ramadani P, I. B. H. H. N. P. Wisudawan, S. Murnani, and H. Setiawan, "On-device MFCC-CNN voice recognition system with ESP-32 and web-based application," in *2023 15th International Conference on Information Technology and Electrical Engineering (ICITEE)*, Oct. 2023, pp. 161–166, doi: 10.1109/ICITEE59582.2023.10317720.
- [12] P. Suebsombut, P. Sureephong, A. Sekhari, S. Chermumroong, and A. Bouras, "Chatbot application to support smart agriculture in Thailand," in *2022 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*, Jan. 2022, pp. 364–367, doi: 10.1109/ECTIDAMTNCN53731.2022.9720318.
- [13] S. K. Assayed, M. Alkhatib, and K. Shaalan, "Advising chatbot for high school in smart cities," in *2023 8th International Conference on Smart and Sustainable Technologies (SpliTech)*, Jun. 2023, pp. 1–6, doi: 10.23919/SpliTech58164.2023.10193065.
- [14] A. Fadhluohman, A. Sulasikin, Y. Nugraha, N. L. R. Husna, M. E. Aminanto, and J. I. Kanggrawan, "Development of Indonesian language intelligent chatbot for public services in JAKI application," in *2023 IEEE International Smart Cities Conference (ISC2)*, Sep. 2023, pp. 1–7, doi: 10.1109/ISC257844.2023.10293514.
- [15] S. Anand, A. M. Abhishek Sai, and M. Karthikeya, "Chatbot enabled smart tourism service for Indian cities: An AI approach," in *2023 11th International Conference on Internet of Everything, Microwave Engineering, Communication and Networks (IEMECON)*, Feb. 2023, pp. 1–7, doi: 10.1109/IEMECON56962.2023.10092286.
- [16] V. Somashekarappa, C. Howes, and A. Sayeed, "Good looking: How gaze patterns affect users' perceptions of an interactive social robot," in *2024 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO)*, May 2024, pp. 128–133, doi: 10.1109/ARSO60199.2024.10557795.
- [17] A. M. A. Sai, O. Balamurali, M. Karthikeya, and S. Anand, "A web-based chatbot for Indian cities: A comparison of CNN, ANN, and LSTM models," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Jul. 2023, pp. 1–6, doi: 10.1109/ICCCNT56998.2023.10307912.
- [18] R. Aadithya, M. Harish, S. M. Faathih, P. S. Manoharan, M. Muthuramalingam, and G. Kannayeram, "Smart home automation using discord chat bots," in *2023 International Conference on Energy, Materials and Communication Engineering (ICEMCE)*, Dec. 2023, pp. 1–4, doi: 10.1109/ICEMCE57940.2023.10433982.
- [19] T. Yamashita, "Analyzing likert scale surveys with rasch models," *Research Methods in Applied Linguistics*, vol. 1, no. 3, Dec. 2022, doi: 10.1016/j.rmal.2022.100022.
- [20] D. Fu, X. Zhang, D. Chen, and W. Hu, "Pathological voice detection based on phase reconstitution and convolutional neural network," *Journal of Voice*, Oct. 2022, doi: 10.1016/j.jvoice.2022.08.028.
- [21] O. Aydogmus, M. C. Bingol, G. Boztas, and T. Tuncer, "An automated voice command classification model based on an attention-deep convolutional neural network for industrial automation system," *Engineering Applications of Artificial Intelligence*, vol. 126, Nov. 2023, doi: 10.1016/j.engappai.2023.107120.
- [22] R. Jiménez-Moreno and R. A. Castillo, "Deep learning speech recognition for residential assistant robot," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 2, pp. 585–592, Jun. 2023, doi: 10.11591/ijai.v12.i2.pp585-592.
- [23] M. S. I. Sharifuddin, S. Nordin, and A. Mohd Ali, "Comparison of CNNs and SVM for voice control wheelchair," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 3, pp. 387–393, Sep. 2020, doi: 10.11591/ijai.v9.i3.pp387-393.
- [24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
- [25] I. Rafegas, M. Vanrell, L. A. Alexandre, and G. Arias, "Understanding trained CNNs by indexing neuron selectivity," *Pattern Recognition Letters*, vol. 136, pp. 318–325, Aug. 2020, doi: 10.1016/j.patrec.2019.10.013.
- [26] H. Yun, "Prediction model of algal blooms using logistic regression and confusion matrix," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 3, pp. 2407–2413, Jun. 2021, doi: 10.11591/ijece.v11i3.pp2407-2413.
- [27] L.-E. Pomme, R. Bourqui, R. Giot, and D. Auber, "Relative confusion matrix: Efficient comparison of decision models," in *2022 26th International Conference Information Visualisation (IV)*, Jul. 2022, pp. 98–103, doi: 10.1109/IV56949.2022.00025.
- [28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [29] K. B. Pranav and J. Manikandan, "Design and evaluation of a real-time face recognition system using convolutional neural networks," *Procedia Computer Science*, vol. 171, pp. 1651–1659, 2020, doi: 10.1016/j.procs.2020.04.177.
- [30] Y. Su and C. C. J. Kuo, "On extended long short-term memory and dependent bidirectional recurrent neural network," *Neurocomputing*, vol. 356, pp. 151–161, Sep. 2019, doi: 10.1016/j.neucom.2019.04.044.




- [31] Y. Su, K. Fan, N. Bach, C.-C. J. Kuo, and F. Huang, "Unsupervised multi-modal neural machine translation," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 10474–10483, doi: 10.1109/CVPR.2019.01073.
- [32] Y. Su, Yuzhong Huang, and C.-C. Jay Kuo, "Dependent bidirectional RNN with extended-long short-term memory," *ICLR 2018 Conference Blind Submission*, no. 1990, pp. 1–14, 2018.
- [33] Y. Su and C.-C. J. Kuo, "Recurrent neural networks and their memory behavior: A survey," *APSIPA Transactions on Signal and Information Processing*, vol. 11, no. 1, 2022, doi: 10.1561/116.00000123.

BIOGRAPHIES OF AUTHORS






Robinson Jiménez-Moreno    is an electronic engineer graduated from Universidad Distrital Francisco José de Caldas in 2002. He received a M.Sc. in Engineering from Universidad Nacional de Colombia in 2012 and Ph.D in Engineering at Universidad Distrital Francisco José de Caldas in 2018. His current working as Assistant Professor of Universidad Militar Nueva Granada and research focuses on the use of convolutional neural networks for object recognition and image processing for robotic applications such as human-machine interaction. Email: robinson.jimenez@unimilitar.edu.co.



Javier Eduardo Martínez Baquero    is an electronic engineer graduated from Universidad de los Llanos in 2002. Posgraduated in Electronic Instrumentation from Universidad Santo Tomas in 2004, posgraduated in Instrumentation and Industrial Control at Universidad de los Llanos in 2020 and Msc in Educative Technology and Innovative Media for Education at Universidad Autonoma de Bucaramanga in 2013. His current working as Associated Professor of Universidad de los Llanos and research focuses on Instrumentation, Automation, Control and Renewable Energies. Email: jmartinez@unillanos.edu.co.



Luis Alfredo Rodríguez Umaña    is an electronic engineer graduated from Universidad de los Llanos in 2002. Posgraduated in Automatic and Industrial Computing from Universidad Autonoma de Colombia in 2006 and Msc. in Information Technology Management at Universidad Cooperativa de Colombia in 2022. His current work as Assistant Professor of Universidad de los Llanos and research focuses on Automation and Control. Email: lrodriguez@unillanos.edu.co.