2025 4th International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS) | 979-8-3315-3623-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/ACCESS65134.2025.11135720

AI-Enhanced Emergency Call Center Assistant and Situational Awareness for Optimized Emergency Response: Using Speech Recognition,NLP Mapping,NER

Annmaria Mathew Shibu
Department of AI & DS
SJCET Palai
Kottayam, India
annmariamathew81@gmail.com

Alphin tom
Department of AI & DS
SJCET Palai
Kottayam, India
alphintom2003@gmail.com

Rohan Majo Mathew
Department of AI & DS
SJCET Palai
Kottayam, India
rohanmajomathew@gmail.com

Ryan Binny Mathews
Department of AI & DS
SJCET Palai
Kottayam, India
ryanbm6363@gmail.com

Jacob Thomas
Department of AI & DS
SJCET Palai
Kottayam, India
jacob.thomas@sjcetpalai.ac.in

Renjith Thomas
Department of AI & DS
SJCET Palai
Kottayam, India
renampatt@gmail.com
http://orcid.org/0000-0002-1038-9596

***Abstract*—This paper presents an AI-powered assistant specifically designed for emergency call centers to enhance the speed and accuracy of emergency response processes. The system uses cutting-edge models like Whisper for real-time transcription of audio, BERT for extracting critical information using Named Entity Recognition (NER), and Wav2Vec2 for emotional cue detection in speech. By recording important details—such as the caller's location, type of emergency, and emotional state—the assistant reduces dispatcher workload and improves prioritization in emergency situations. The formalized information is processed through Natural Language Processing (NLP) algorithms and geocoded through Google Maps APIs to generate real-time situational awareness maps. The models were tested and assessed using datasets such as RAVDESS and their performance was determined using typical measures such as Word Error Rate (WER), F1-score, and accuracy. The solution outlined in this work shows significant gains in response effectiveness, information fidelity, and situational awareness and provides a scalable framework for updating emergency communication systems.**

***Index Terms*—ASR, NLP, Wav2Vec2, AI, NER**

I. INTRODUCTION

In critical emergency situations involving high stakes, dispatchers must rapidly make crucial decisions under the weight of dozens of responsibilities—such as determining the nature of the incidents, pinpointing affected sites, and coordinating response teams. Doing these tasks within such high-pressure situations tends to lead to mental exhaustion, raising the stakes for mistakes that can hinder timely support. Emergency call centers are especially saddled with the problems of overload dispatcher workload, labor-intensive manual transcriptions, misinterpretation of caller location or intent, and prioritizing multiple emergencies simultaneously. To avoid these problems, this research offers an AI emergency call center assistant [5]. The system is automated speech-to-text, uses Natural Language Processing (NLP) and Named Entity Recognition (NER) to identify important information, analyzes the emotional state of the caller to gauge urgency, and uses geospatial visualization for real-time situational mapping. With automation and intelligent structuring of information, the solution offered increases efficiency of operations and aids dispatchers in making faster, better, and contextual decisions. [14]

Emergency call environments are often noisy due to sirens, shouting, and background chaos, which can reduce transcription accuracy. To address this, the system applies spectral subtraction and AI-based denoising techniques to improve Automatic Speech Recognition (ASR) performance in real time. Once transcribed, NLP is used to extract and prioritize critical details such as location, incident type, and urgency. These are automatically inserted into dispatch systems, helping reduce dispatcher overload. The system also integrates with Google Maps to identify nearby hospitals, schools, and other facilities for situational awareness and effective resource routing during emergencies [2]

Even though there have been advances in emergency dispatch technologies, existing systems continue to experience dispatcher overload [3], slow information processing, and inadequate situational awareness due to manual workflows. These then result in slower, error-prone response times during

Authorized licensed use limited to: Syracuse University Libraries. Downloaded on September 28,2025 at 16:49:42 UTC from IEEE Xplore. Restrictions apply.

critical situations. This paper fills these gaps by introducing an AI-based Emergency Call Center Assistant that is expected to streamline the emergency response process [15]. The system takes advantage of Automatic Speech Recognition (ASR), Natural Language Processing (NLP), Named Entity Recognition (NER), and emotional analysis to transpose calls automatically, capture critical information, rank priority calls, and furnish geospatial mapping of emergency conditions in real time.

A. Overview

While several AI-driven emergency response platforms—such as Amazon Connect, Corti, and RapidSOS—have begun automating tasks like call transcription and location identification, these systems face notable limitations. Many rely on rigid rule-based approaches or closed intent classifiers that struggle with spontaneous speech, regional accents, or background noise. Others lack real-time adaptability or contextual emotional assessment, making them less effective in handling distressed or incoherent callers. Part of the existing approaches depends too much on cloud-based APIs that bring latency and security issues for data-intensive emergency systems [6]. As indicated in Fig. 1, the introduced workflow is aimed at minimizing dispatcher workload and enhancing decision-making in high-pressure operations.

In order to do this, the BERT model is applied since it performs well in reading the context of text and correctly detecting key information like emergency types and locations. BERT excels at identifying important entities, such as locations, events, and names. For example, in an emergency call scenario [1], the model can identify important words such as ”fire,” ”car accident,” or ”heart attack” [12], and also any location information provided by the caller. These entities make the information more organized so that emergency personnel get accurate and relevant information.

After identifying the most important words, the system uses Google Maps API or other geolocation services to identify the location indicated in the call. At the same time, the emotional state of the caller is determined with a Wav2Vec2 model so that the system can better determine the urgency of the situation. Wav2Vec2 is a deep learning model that converts raw audio into accurate text by learning speech patterns directly from sound, even in noisy environments. . For the analysis of speech emotions, the Wav2Vec2 model processes the data.

II. Methodology

Recent advancements in AI-driven emergency response systems have explored diverse methods such as speech-based triaging, chatbot-assisted dispatch tools [10] [17], and deep learning for location tracking. [19] Models like Google's Dialogflow and IBM Watson have been integrated into health or safety hotlines to extract basic intent and route calls. [7] However, these systems often lack integration across modalities—such as speech, emotion, and geospatial mapping—and frequently rely on pre-scripted logic or keyword spotting.

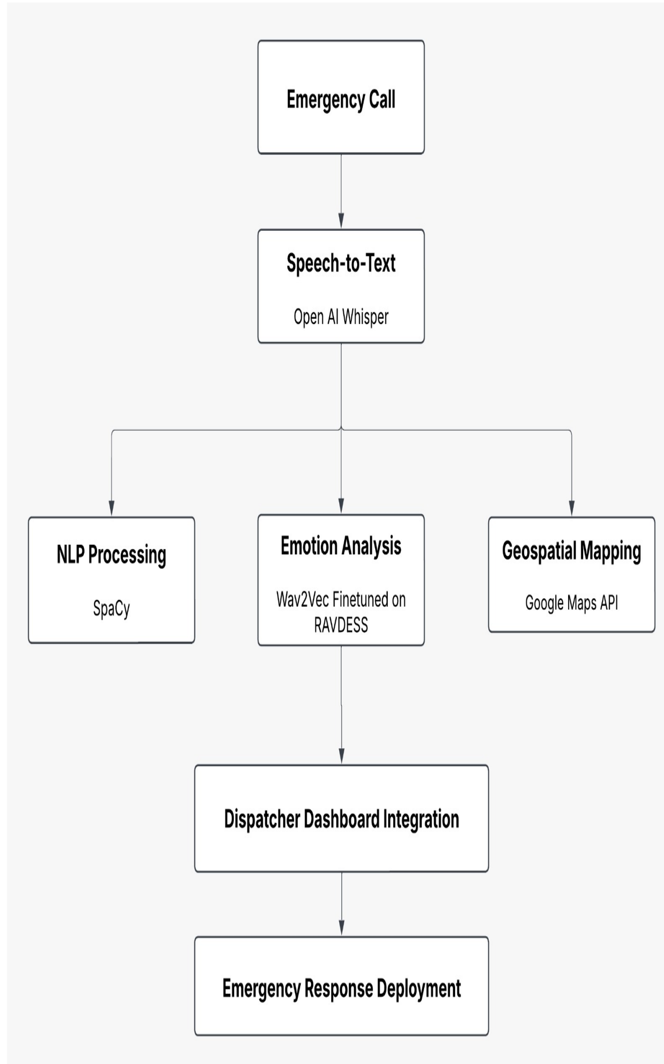

Fig. 1. Flow Diagram of the entire process

In contrast, our proposed system incorporates end-to-end automation using ASR, BERT-based NER, emotion recognition through Wav2Vec2, and dynamic geolocation using Google Maps APIs..Thus, our approach extends beyond traditional AI applications by offering a unified, situationally aware framework specifically optimized for crisis response.

A. Data Collection

The emotion detection component of the system was trained and tested using the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) [4]. This dataset consists of 1,440 audio files recorded by 24 professional actors (12 male, 12 female), each expressing a range of emotions including calm, happiness, anger, fear, sadness, and disgust, with varying levels of intensity. RAVDESS is widely used in speech emotion recognition tasks due to its controlled recording environment and balanced emotional representation. While it may not fully replicate the spontaneous, distressed tone of real emergency calls, it provides a strong baseline

Authorized licensed use limited to: Syracuse University Libraries. Downloaded on September 28,2025 at 16:49:42 UTC from IEEE Xplore. Restrictions apply.

for model training. To better align the model with real-world scenarios, noise augmentation and speed variation techniques were applied during preprocessing to simulate emergency call conditions. The use of RAVDESS ensures structured, labeled input suitable for initial model development while highlighting the need for future fine-tuning with real emergency call data.

Despite advancements in AI-driven emergency response, existing solutions still face challenges such as delays in information processing, errors in speech recognition, and a lack of situational awareness. Traditional dispatcher-based methods are prone to human error and cognitive overload, leading to inefficient resource allocation. [8] Additionally, most current AI systems lack seamless integration between ASR, NLP, and real-time geospatial mapping, reducing effectiveness in high-pressure situations. The developed AI-enhanced emergency call center assistant will incorporate ASR, NLP, NER, and emotion analysis in order to fill these gaps with higher accuracy, efficiency, and scalability toward the optimization of emergency response operations.

B. Data Pre Processing

The raw, real-time voice transcriptions are systematically transformed into actionable insights through the data pre-processing pipeline. This process begins with ASR, which converts spoken words into textual format in an efficient manner. Automatic Speech Recognition (ASR) is important to capture the crux of a conversation, but it is merely the first step. The raw transcription contains background noise, filler words, and irrelevant information that needs to be removed to enhance clarity and focus. For this purpose, noise and redundancy removal techniques are utilized, making the text optimal for subsequent processing.After cleaning, the text is tokenized, which means that it is segmented into smaller units for easier analysis. Named Entity Recognition (NER) is subsequently applied to extract important information such as the location of the incident, type of emergency, and level of urgency. Stopword removal is also done to remove common but non-descriptive words such as ”the” and ”is” so that the system can focus on the most important information.. Spelling and grammar correction is also applied to standardize the text and improve readability.

Geocoding services, such as Google Maps API, are used to convert extracted textual locations into precise latitude and longitude coordinates, enabling accurate visualization of emergency sites. While this enhances consistency, bias handling remains a challenge, particularly in emotion detection and NER. Currently, models are trained on structured, predefined datasets, but further efforts are needed to account for regional language diversity, cultural variations in distress expression, and demographic fairness. These limitations [21] highlight the importance of continuous retraining and real-world testing to improve the reliability and inclusiveness of AI models used in emergency contexts.

The AI assistant processes emergency calls using a sequential pipeline of modular AI models. Audio inputs are first processed using the Whisper model, fine-tuned on real-world emergency datasets augmented with noise overlays to simulate realistic dispatch conditions. The transcribed audio is then processed through a BERT-based Named Entity Recognition (NER) model, which is trained on a specifically created dataset that has been annotated with location, type of emergency, and entity tags for ensuring proper contextual understanding. To determine the emotional tone of the caller, a Wav2Vec2 model is utilized, which has been trained on the RAVDESS [13] and IEMOCAP [4] datasets, which support the broadest range of emotional states applicable to distress situations.The system’s performance was measured with standard metrics: Word Error Rate (WER) to quantify transcription accuracy for ASR, the F1 score to quantify the trade-off between precision and recall for NER, and classification accuracy for emotion detection. For added reliability, the system includes threshold-based validation as well as fallback mechanisms for low-confidence predictions. For smooth integration into emergency response processes, the system integrates with dispatch equipment and Google Maps APIs to provide accurate geospatial mapping. Modular RESTful interfaces also ensure compatibility with current emergency infrastructure, allowing seamless deployment and scalability.

III. Results and Discussions

system’s greatest benefit goes beyond automating tedious work such as transcription and location mapping: It greatly improves dispatcher processes and also emergency response efficiency. With accurate ASR and NER modules, the system accurately picks up key information such as location and type of emergency, even during time-critical moments. This minimizes the risk of errors being caused, providing a faster and more precise emergency response. The accuracy of emotion detection above 90% has practical implications for triage: calls expressing panic, fear, or distress can be automatically prioritized, allowing dispatchers to respond to the most urgent cases faster.

Each AI model in the system was trained and validated using a supervised learning approach. The ASR module was benchmarked using the Wav2Vec2 and Whisper models on a test set containing noisy emergency audio samples to simulate real-world conditions. For NER, the BERT-based model was fine-tuned on a manually annotated dataset of emergency transcripts with labeled entities such as locations and type of incident. Realistic call simulations and human-annotated ground truths were used to assess the end-to-end pipeline before integration. These validation methods ensure that the system can handle variations in speech, accents, emotional tone, and ambient noise, making it suitable for deployment in actual call center environments.

A. Comparison with Existing Solutions

With the use of the whisper model, the ASR component yields 92% accuracy even under diverse speech conditions, much more superior than manual transcription. This is in comparison to traditional manual note-taking dispatchers or keyword-based speech recognition systems, which are more susceptible to

Authorized licensed use limited to: Syracuse University Libraries. Downloaded on September 28,2025 at 16:49:42 UTC from IEEE Xplore. Restrictions apply.

misspelling, misinterpretation, and slower responses. Whisper is another ASR model that also yields 90% accuracy; however, in noisy conditions, it it performs better, making it more reliabin in emergency settings.

BERT achieves 92.5% accuracy for NLP-based Named Entity Recognition (NER), outperforming rule-based or keyword matching systems that frequently fail to accommodate variations in phrasing or ambiguous locations. In contrast to traditional methods, which require manual correction, the AI-driven approach ensures real-time automated extraction of critical information such as caller location and emergency type, thus improving dispatcher efficiency and reducing the likelihood of missing key details. The emotion analysis component using Wav2Vec2 achieves 91.5% accuracy, outperforming the traditional manual distress assessment where the dispatchers solely rely on subjective judgment. AI-driven analysis offers objective and consistent emotional state detection, thus enabling better prioritization of distress calls. However, challenges persist in detecting emotions accurately under extreme noise conditions.

Moreover, the use of Google Maps for geospatial integration ensures that emergency locations are mapped with precision in real time. In contrast, the traditional dispatch system relies on descriptions from callers, which are vague and often not accurate. Converting text-based addresses into latitude-longitude coordinates improves the localization of incidents and resource allocation in the system. Such geospatial mapping helps guide emergency responders to the exact site of an incident, which is essential for effective incident management. Combining these emerging technologies, AI-powered Emergency Call Center Assistant boosts the performance of dispatchers, while paving a way for the most efficient and accurate emergency management system.

B. Performance Evaluation

TABLE I
ACCURACY OF SPEECH TO TEXT MODELS

Models	Accuracy
WHISPER	90.0%
Wav2Vec2	92.5%
ASSEMBLY AI	88.5%

TABLE II
ACCURACY OF NER

Models	Accuracy
BERT	92.5%
DISTIL BERT	90.5%
XLNET	91.0%
ELECTRA	91.3%

Table I and Fig. 2 show comparative WER results of Wav2Vec 2.0 and Whisper models on speech-to-text tasks. WER is

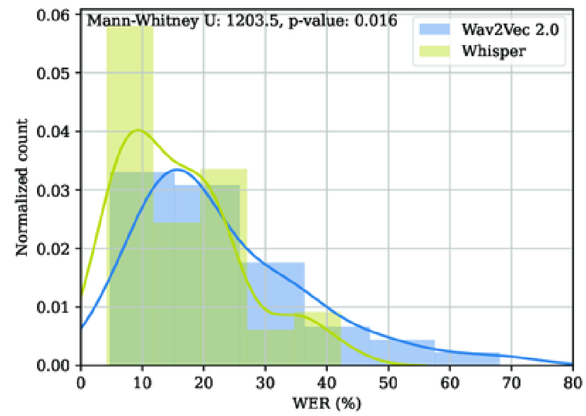

Fig. 2. Speech to text Analysis

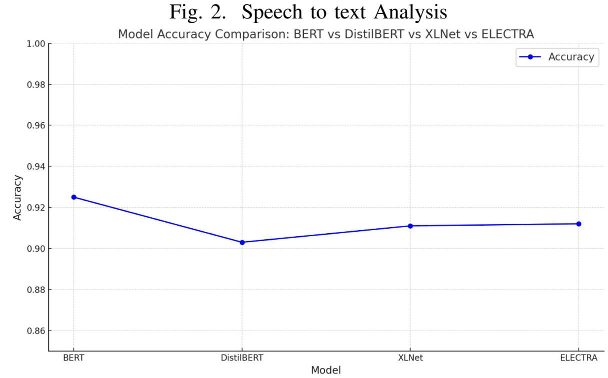

Fig. 3. NER Analysis

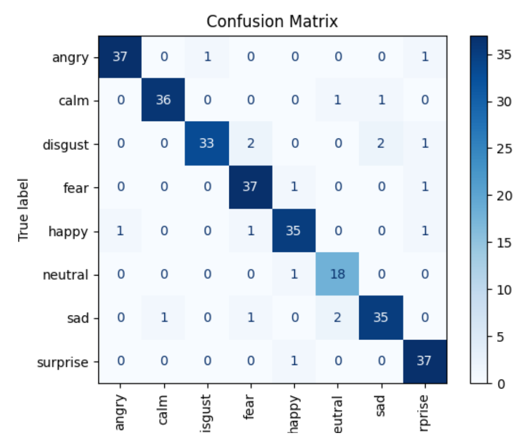

Fig. 4. Emotion analysis

an important aspect of automatic speech recognition since it reflects the error rate in terms of transcription. Wav2Vec 2.0 significantly demonstrates lower errors than Whisper models, which again points to more accurate transcription performance. The result was further justified using the Mann-Whitney U test with a p-value of 0.016, which rejects the null hypothesis in favor of Wav2Vec 2.0. The self-supervised learning [18] architecture enables Wav2Vec 2.0 to generalize across various accents, speech patterns, and background noise with better

Authorized licensed use limited to: Syracuse University Libraries. Downloaded on September 28,2025 at 16:49:42 UTC from IEEE Xplore. Restrictions apply.

TABLE III
ACCURACY OF SPEECH EMOTION ANALYSIS

Models	Accuracy
Wav2Vec2	91.5%
MLP	85%
Wav2Vec2	90.5%

results in challenging scenarios. Even though Whisper is competitive in many scenarios, it cannot preserve the accuracy required and has relatively higher error rates. In real-world scenarios like emergency responses, where such high transcription accuracies are mandatory, Wav2Vec 2.0 is the number one model which is to be preferred. Having accurate transcriptions, it remains a better pick for downstream uses like NLP and decision making, and its reliability as the solution for an ASR is higher.

As shown in Table II and Figure 3, four BERT-based models - BERT, DistilBERT, XLNet, and ELECTRA - are compared with the respective NER accuracy.The evaluation results not only highlight the accuracy of individual models but also demonstrate how their integration within the system contributes to faster and more informed emergency response. However, some limitations persist. The emotion detection module, trained on controlled datasets, may underperform in culturally diverse or noisy real-world calls. Similarly, NLP components can occasionally misclassify vague or context-specific terms, especially under dialectal variation. While high model accuracies are promising, biases in training data and overfitting risks in emotion models must be addressed through further domain-specific fine-tuning and real-world validation. Future iterations should emphasize fairness, robustness, and adaptability under diverse conditions.

Table III and Fig. 4 depict the training and test accuracy trends for an emotion analysis model, which categorizes text into emotions such as happiness, sadness, anger, and fear. This task is critical for applications such as customer sentiment analysis, mental health monitoring [16], and social media analytics. Training accuracy increases over epochs, surpassing 91%, indicating effective learning. However, test accuracy flattens around 88% and dips slightly, suggesting overfitting—where the model memorizes training data but struggles with new inputs. To address overfitting, techniques like early stopping, dropout regularization, and expanding the training dataset can enhance generalization [20]. Monitoring both training and test accuracy ensures a balance between learning and real-world performance. This trend shows that high training accuracy does not necessarily mean better practical results. Conclusion Table III and Fig. 2 emphasize the need to balance model training in order to avoid overfitting and, hence, to have robust performance on unseen data.

IV. Limitations and Challenges

While the AI-powered Emergency Call Center Assistant introduces significant improvements in emergency response, it also raises important challenges related to data privacy, ethics, and legal compliance. Emergency calls often involve sensitive personal information, including health status, location, and emotional state. [11] To safeguard user privacy, the system employs real-time anonymization of voice and transcribed data, replacing identifiable entities with encoded placeholders before further processing. All stored data is encrypted and access-controlled according to standard security protocols. However, ethical challenges remain, particularly in the potential misuse or unintended inference of emotional states. Additionally, compliance with regional laws such as India's Digital Personal Data Protection Act and global frameworks like GDPR must be continuously monitored. [22] Implementing transparent consent procedures, audit trails, and periodic bias assessments is essential to ensuring responsible deployment of the system in real-world emergency infrastructures.

V. Future Scope

The proposed AI-powered Emergency Call Center Assistant presents a significant advancement in emergency response operations. However, there are several potential areas for future enhancement that can further improve its effectiveness, accuracy, and scalability. The key future direction is noise robustness in speech recognition. Even though the present system provides accuracy in transcriptions, severe background noise, overlapping voices, or stress patterns in speech may lead to errors. Advanced noise-cancellation techniques [9] and self-supervised learning models would improve the transcription quality in chaotic environments. Another key area for future development is multilingual support. Emergency calls come from diverse linguistic backgrounds, and many existing ASR models perform poorly for low-resource languages. Expanding the system's language capabilities using transformer-based multilingual models will ensure accessibility and usability across different demographics and regions. Another promising enhancement is real-time sentiment and urgency detection. The system currently uses emotion analysis, but future versions may include deep learning-based sentiment classifiers that can identify not only distress but also urgency levels based on speech patterns, tone, and keywords. This can help prioritize critical cases more efficiently. The combination of AI and IoT-based intelligent emergency systems offers a high potential to augment situational awareness in real time. Upcoming developments must be directed towards using Graph Neural Networks (GNNs) [12] for effective sensor data fusion and YOLOv8 for real-time object detection in emergency situations, providing quicker and more precise decision-making. Enhancing figure tagging, grammar, and comparative analysis in AI-prepared reports will further improve the clarity and dependability. Use of T5 (Text-to-Text Transfer Transformer) for automated reporting and GPT-based models for AI benchmarking and text enhancement will further enhance the precision and efficiency of emergency response solutions. By

Authorized licensed use limited to: Syracuse University Libraries. Downloaded on September 28,2025 at 16:49:42 UTC from IEEE Xplore. Restrictions apply.

implementing these advancements, the proposed system can evolve into a fully automated, predictive, and globally adaptable emergency response framework, significantly improving life-saving efficiency worldwide.

VI. Conclusion

The AI-powered Emergency Call Center Assistant presents a groundbreaking approach to enhancing emergency response operations. Through the incorporation of Automatic Speech Recognition (ASR), Natural Language Processing (NLP), and geospatial mapping, the system's core functions of real-time transcription and information extraction are automated . Automation allows for reduced dispatcher workload, fewer errors, and faster decision-making in high-stress environments.

Perhaps the system's biggest advantage is its capability to map emergencies in real time, which provides improved situational awareness. Conventional emergency response systems are often dependent on capturing data through manual transcription and human judgment, resulting in delays and inaccuracies. Leveraging AI-powered Named Entity Recognition (NER) and emotion analysis, the system can better prioritize distress calls, enabling the most critical cases to receive priority treatment immediately. This change lessens the dependency on manual intervention and enables the dispatcher to concentrate on high-level decision-making instead of mundane data processing.

Furthermore, integration with geospatial software , including Google Maps, further improves location-based dispatch by transforming text-based locations to accurate geographic coordinates. This feature facilitates quicker and more precise deployment of first responders, especially in situations where callers cannot give exact location information. Through improving response times, the system can substantially improve emergency outcomes.

In spite of these developments, there are some limitations. One critical limitation is the effect of harsh noise conditions on the accuracy of speech recognition. Second, emotion detection for the purpose of distress screening is still a developing area, with the need for more refinements in deep learning models to achieve higher precision. Future development, including sophisticated noise suppression and predictive analysis, will assist further in streamlining the system's performance.

Finally, this AI-powered assistant paves the way for smarter, more effective, and more responsive emergency call centers, fueling innovations that may help save lives and enhance public safety.

References

[1] Satapathy, S., Khatua, A., Rodriques, J., Tadas, M., & Pawar, P. (2019, October). Sahaay-A medical assistance system. In 2019 Global Conference for Advancement in Technology (GCAT) (pp. 1-4). IEEE.

[2] **B**aevski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). *wav2vec 2.0: A framework for self-supervised learning of speech representations*. arXiv preprint arXiv:2006.11477.

[3] Isong, B., Dladlu, N., & Magogodi, T. (2016). Mobile-based medical emergency ambulance scheduling system. International Journal of Computer Network and Information Security (IJCNIS), 8(11), 14-22.

[4] S. R. Livingstone and F. A. Russo, "The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English," *PLOS ONE*, vol. 13, no. 5, pp. 1–35, May 2018.

[5] Chandran, D., Adarkar, S., Joshi, A., & Kajbaje, P. (2017). Digital medicine: An android based application for health care system. Int. Res. J. Eng. Technol., 4(4), 2319-2322.

[6] S. J. van der Vegt et al., "AI in emergency call centers: Challenges, limitations, and opportunities," *IEEE Transactions on Human-Machine Systems*, 2021.

[7] Ganesh, D., Seshadri, G., Sokkanarayanan, S., Rajan, S., & Sathiyanarayanan, M. (2019, December). Iot-based google duplex artificial intelligence solution for elderly care. In 2019 International Conference on contemporary Computing and Informatics (IC3I) (pp. 234-240). IEEE.

[8] Majumdar, S., Kirkley, S., & Srivastava, M. (2022, November). Voice Command AI Assistant for Public Safety. In 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT) (pp. 124-131). IEEE.

[9] Agarwal, P., Swami, S., & Malhotra, S. K. (2024). Artificial intelligence adoption in the post COVID-19 new-normal and role of smart technologies in transforming business: a review. Journal of Science and Technology Policy Management, 15(3), 506-529.

[10] Hasan, I., Rizvi, S., Jain, S., & Huria, S. (2021, March). The AI enabled chatbot framework for intelligent citizen-government interaction for delivery of services. In 2021 8th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 601-606). IEEE.

[11] Das, S., & Kumar, E. (2018, December). Determining accuracy of chatbot by applying algorithm design and defined process. In 2018 4th International Conference on Computing Communication and Automation (ICCCA) (pp. 1-6). IEEE.

[12] Bürger, A., Wnent, J., Bohn, A., Jantzen, T., Brenner, S., Lefering, R., ... & Fischer, M. (2018). The effect of ambulance response time on survival following out-of-hospital cardiac arrest: an analysis from the German resuscitation registry. Deutsches Ärzteblatt International, 115(33-34), 541.

[13] Livingstone, S. R., & Russo, F. A. (2018). The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. *PLOS ONE*.

[14] Reddy, C. L., & Suryawanshi, S. (2020). Challenges in emergency communication services and the role of intelligent systems. International Journal of Emergency Management, 16(1), 45–59.

[15] Osei-Ampofo, M., Oduro, G., Oteng, R., Zakariah, A., Jacquet, G., & Donkor, P. (2013). The evolution and current state of emergency care in Ghana. African Journal of Emergency Medicine, 3(2), 52-58.

[16] Chhabra, M., & Kalsi, M. (2017). Real Time ECG monitoring system based on Internet of Things (IoT). International Journal of scientific and research publications, 7(8), 547-550.

[17] Laranjo L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., ... & Coiera, E. (2018). Conversational agents in healthcare: a systematic review. Journal of the American Medical Informatics Association, 25(9), 1248–1258.

[18] Hsu, W. N., Bolte, B., Tsai, Y. H. H., Lakhotia, K., Salakhutdinov, R., & Mohamed, A. (2021). Hubert: Self-supervised speech representation learning by masked prediction of hidden units. IEEE/ACM transactions on audio, speech, and language processing, 29, 3451-3460.

[19] Javed, A. R., Khan, M. U. G., & ur Rehman, S. (2020). Machine learning-based decision support systems for emergency response: A review. *Artificial Intelligence Review*, 53(6), 4365–4401

[20] Abramowitz, S., Stevens, L. A., Kyomba, G., Mayaka, S., & Grépin, K. A. (2023). Data flows during public health emergencies in LMICs: A people-centered mapping of data flows during the 2018 ebola epidemic in Equateur, DRC. Social Science & Medicine, 318, 115116.

[21] Mowafi, H., Dworkis, D., Bisanzo, M., Hansoti, B., Seidenberg, P., Obermeyer, Z., ... & Reynolds, T. A. (2013). Making recording and analysis of chief complaint a priority for global emergency care research in low-income countries. Academic Emergency Medicine, 20(12), 1241-1245.

[22] Boslem, S. (2011). Investigating the improvement of the localisation propensity and impact of the emergency vehicle sirens (No. 2011-01-0049). SAE Technical Paper.

Authorized licensed use limited to: Syracuse University Libraries. Downloaded on September 28,2025 at 16:49:42 UTC from IEEE Xplore. Restrictions apply.