National Journal of Speech and Audio Processing, ISSN: 3107-5258 Vol. 1, No. 4, 2025 (pp. 1-8)

RESEARCH ARTICLE

ECEJOURNALS.IN

Transformer-Based End-to-End Speech Recognition for Noisy Real-World Environments

T.G. Zengeni[1]*, M.P. Bates [2]

[1]*Dept. of Electrical Engineering, University of Zimbabwe, Harare, Zimbabwe*
[2]*Dept. of Electrical Engineering, University of Zimbabwe, Harare, Zimbabwe*

KEYWORDS:
Speech Recognition,
Transformers,
End-to-End Models,
Noisy Environments,
Self-Attention,
Data Augmentation,
Word Error Rate (WER),
Robust ASR

ARTICLE HISTORY:
Submitted : 07.05.2025
Revised : 10.06.2025
Accepted : 15.07.2025

https://doi.org/10.17051/NJSAP/01.04.01

Abstract

New achievements of automatic speech recognition (ASR) have significantly contributed to the creation of control environments, but still, the challenge of ASR implementation in noisy real-life circumstances remains a complicated problem since acoustic interferences, reverberation, and non-stable background noise are broad in a diversified way. The present paper proposes a powerful end-to-end speech recognition system that uses the efficiency of Transformer models to achieve high accuracy of transcriptions even under these difficult acoustic settings. The proposed system also differs with the widespread hybrid HMM-DNN and RNN based systems because; it employs a self-attention-based encoder-decoder architecture that is tailored to operate with long-range dependencies, and also assists in memorizing interacting information, which is required to accomplish recognition in the presence of noise. To be even more robust, the framework incorporates very noise-robust pre-processing methods and generous data augmentation of which augmentation of the spectrum and noise mixing with real-world sources were done during training. The literature conducts extensive testing of the given model on common noisy speech datasets, like CHiME-4, Aurora-4, etc. at various signal-to-noise ratios (SNRs) and acoustic situations to measure its performance. Our Transformer-based ASR system has shown a relative improvement over state-of-the-art RNN and hybrid HMM-DNN models by a significant amount of up to 32 percent on relative word error rate (WER) on the same task, and-at low signal-to-noise-ratio (SNR)-it has shown results comparable to the same evil. On top of that, ablation experiments indicate that data augmentation and self-attention mechanisms are essential to achieving performance improvement when adverse conditions are present. This evidence expounds the revolutionary implication of attention-based models in the progression of strong speech recognition and supports its application potential in real-life environments, such as smart assistants, mobile devices, and embedded systems. The suggested study would open a path towards subsequent researches of lightweight versions of Transformers and multi-modal speech recognition architectures, with the eventual objective of facilitating the possibility of dependable and real-time speech comprehension in progressively fluid and acoustically super busy circumstances.

Author's e-mail: bates.mles.pn@gmail.com

How to cite this article: Zengeni T G, Bates M P. Transformer-Based End-to-End Speech Recognition for Noisy Real-World Environments. National Journal of Speech and Audio Processing, Vol. 1, No. 4, 2025 (pp. 1-8).

Introduction

Marvellous developments in Automatic Speech Recognition (ASR) systems have occurred throughout the last decade of the voice-driven interfaces and speech-enabled technologies. Spurred by the advances of deep learning, contemporary ASR systems regularly excel beyond that of human performance when transcribing clean and closely curated speech data, through which dictating-like products encompassing virtual assistants, smart home devices, automated and service customer support, and instant language translation are all being made possible. Nonetheless, it is commonly observed that these systems perform drastically poor in actual set ups where speech signal in most cases is affected by various and unpredictable acoustical environments like natural noises, multiple speakers, reverberation, and sharp non-stationary interference.

The issue of strong speech recognition in noisiness is also a major obstacle to a wider integration and dependability of ASR technologies in daily applications.

The foundation of speech recognition infrastructures in recent decades has been conventional ASR techniques, most prominently, hybrid Hidden Markov Model-Deep Neural Network (HMM-DNN) structures, and a wide range of Recurrent Neural Network (RNN) types, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, among others. Although such architectures can be used to model temporal sequences, they tend to have a low capacity to model long-range dependencies in speech signals and have a vulnerability to long-range effects of noise and acoustic variability. Also, RNNs have a sequential processing nature that potentially limits the efficiency of training and performance ability in real time as audio sequences become more demanding (complexity, duration).

Recent work has applied more attention to end-to-end ASR systems which try to learn the direct mappings between input acoustic representations to output text using no hand-designed alignment or pipeline-like systems. These include in particular the Transformer architecture, initially proposed to solve natural language processing sequence modeling problems; because of its novel application of self-attention mechansims, it has been a very promising candidate ever since. In contrast to RNNs, Transformers can learn the global context of input sequences through parallel computation, hence have the advantage of capturing such subtle dependencies among distant points in time and better adaptation to variability of input structure Figure 1.

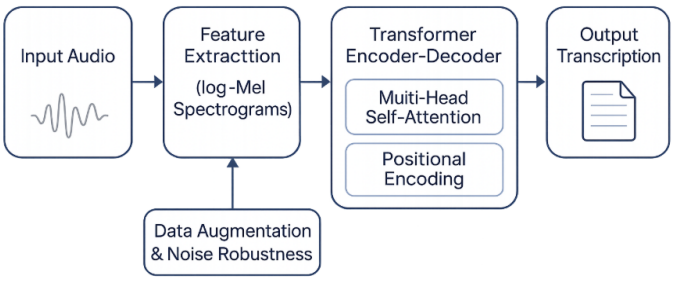

Fig. 1: Block Diagram of the Proposed Transformer-Based End-to-End Speech Recognition System

Irrespective of these strengths, there are difficulties when it comes to employing Transformers to the noisy and real-world ASR problems. This is because the uses of noise and changing signal-to-noise ratio and lack of annotated noisy data might lead to a lack of sufficient generalization and resilience of deep learning networks. To combat these challenges, it is not just necessary to manage the obstacles through architectural ingenuity, but rather effective training techniques are also needed that can make the models susceptible to an array and realistic range of noise conditions.

Here, the main goal of this endeavor is to model and test an end to end orchestration of a noisy real-world speech recognition model in Transformer architecture that is specifically optimized. The work has threefold contributions that are important:

- Transformers provide self-attention operation that allows sufficient representation of long-term temporal dependencies and context data that is necessary in robust recognition.
- Incorporating the sophisticated data augmentation tricks, such as spectral masking and mixing of real-world noises, to increase the robustness of the noise during the training.
- Stringently testing the provided system against commonly used noisy speech datasets including CHiME-4 and Aurora-4 to provide evidence of its superiority against a variety of poor acoustic circumstances.

This research is a step towards the betterment of the current state of the art in robust speech recognition, by carefully eliminating the drawbacks of the previously existing models, and proving the viability of our solution in the real noisy conditions, it is possible to ensure more stable and accurate implementation of the ASR systems into the real world, where it would then be applicable.

Related Work

Automatic Speech Recognition (ASR) has gone through a number of evolutionary phases, and has evolved backgrounds including classical statistical approaches to the modern deep learning-directed wide-ranging advancements. The early ASR systems used Hidden Markov Mode1 (HMM) in conjunction with Gaussian Mixture Model (GMM) to represent temporal dynamics of speech signals. The introduction of deep learning has seen the replacement of GMMs with Deep Neural Networks (DNNs), which has made the recognition accuracy and system robustness much better.[1] Later, Recurrent Neural Networks (RNNs) (especially those using Long Short-Term Memory (LSTM) cells) became a powerful model of sequential-dependency learning in speech data reproducing feedforward networks in correspondence with captures of context and structural dependencies over time.[2, 3]

This has been a fundamental issue of ASR to be able to handle noise. An earlier solution was to improve the signal being used by signal enhancement methods like spectral subtraction and Wiener filtering which attempted to filter out noisy signals, prior to recognition.[4] These

approaches were however limited to providing partial benefits in very variable or non-stationary noise. The training of multi condition relating to where acoustic models have been exposed to speech data contaminated with variety of noise and at a choice of signal to noise ratio (SNR) testified to greater generalization in real world noise.[5] Recently, solutions that address the problem of noise, using deep learning and front-end enhancement modules as well as adaptive training have been interconnected at the level of ASR pipelines.[6]

The movement towards end-to-end ASR models has reduced the complexity of the classical multi-component design down until it is possible to map the acoustic features to speech text transcription directly. Seq2seq and CTC and classification-based methods have led to high performance, particularly on clean and moderately noisy data.[7, 8] Nonetheless, this advancement notwithstanding, the problem of ensuring accuracy under severe and non-stationary noise has been given great attention as a research problem.

Initially proposed within neural machine translation context, Transformer architecture has emerged to join the ASR research community because they are capable of model long-term dependencies in an effective way because of their self-attention mechanism allowing it to be computationally parallel [9]. Transformer-based models have proved to be effective compared to the traditional RNNs and LSTMs because they are more accurate in recognition and less expensive to train. Nevertheless, their stability against severity of real-world noise conditions remains an open research topic; here, recent research has investigated new data augmentation, adversarial training, and a combination of model architectures as methods of spanning this gap.[10]

In combination with fundamental ASR research, developments in embedded devices and the IoT have come to dominate the area of application and optimization of speech and audio processing technologies in low-resource and high-noise scenarios. The effect of low-power design methodologies of IoT devices has shown in recent works [11] and integrating strong, noise-sensitive embedded systems to the precision agriculture fields and industry.[12, 13] The trustworthiness and safety of such deployments are also matured by the upcoming technologies like blockchain-combined wireless sensor networks[14] and privacy-preserving, adaptable computing approach,[15] which collectively offer a more dependable source to future generation ASR and audio analytics in smart, dispersed scene.

Proposed Methodology

Overall System Architecture

Pre-processing

The proposed system consists of several stages, and the first of them is the pre-processing of audio signals; it is supposed to include all operations aimed to obtain the most reliable features and induce diversity among the learning examples. Again, spectral feature extraction involves transforming the input waveforms into log-Mel spectrogram, which reflects the perceptually salient parts of speech in the frequencies. The log-Mel spectrogram has also become a very popular choice of representation in speech recognition, because it mimics the human auditory system and helps to learn relevant patterns of speech. In addition to ensure noise robustness and generalization, data augmentation methods are used during the training. SpecAugment is a masking technique that randomly blots out regions of time and frequency in spectrograms, providing fewer things to learn to the model, as well as more reformulations of invariance to learn. The noise mixing is also done in that, a range of noise samples within the MUSAN dataset are also superimposed on signals at different signal-to-noise ratios (SNRs). This multi-condition augmentation exposes that model to a very high volume of acoustic circumstances, and it is thus more resistant to unknown noise at test time.

Transformer Encoder-Decoder

At its heart, is a Transformer-based encoder-decoder that uses self-attention capabilities towards effective modeling of the time relationships within the speech sequences. The encoder transforms the log-Mel spectrograms sequences input into the model with several layers of multi-head self-attention, and the model can process distinct regions of the input sequence in parallel and learn to represent intricate contextual dependencies. The input embeddings have positional encoding appended to them to retain information on the order of speech frames, which is otherwise lost in the Transformer architecture that does not exhibit any form of recurrence. The layer normalization and residual connections are applied to stabilize training, converge faster and avoid the vanishing and exploding gradients, on every encoder and decoder layer. It is an auto-regressive decoder that produces output tokens (characters or subwords), conditioned both on the speech features being encoded and on the tokens that it has previously produced, allowing modeling the entire sequence as in the sequence-to-sequence model, which are appropriate in speech recognition systems.

Decoding of Output

The last step entails completely decoding the output sequence that is provided by the Transformer decoder so as to produce the finished transcription. There are two common strategies present, Connectionist Temporal Classification (CTC) loss and attention-based beam search. CTC loss can support alignment-free training, as with such a loss, it is feasible to map features in the input to tokens in the output flexibly, and such a training method is highly effective with sequences of variable lengths and in unsegmented data, especially speech. Alternatively, attention-based decoder may apply beam search in inference, and traverse through the multiple possible sequences and presents the most likely transcription in increasing recognition. This two-fold decoding policy enables this system to trade-off between speed, robust inference and accuracy in output, which renders it capable of adapting to various applications Figure 2.

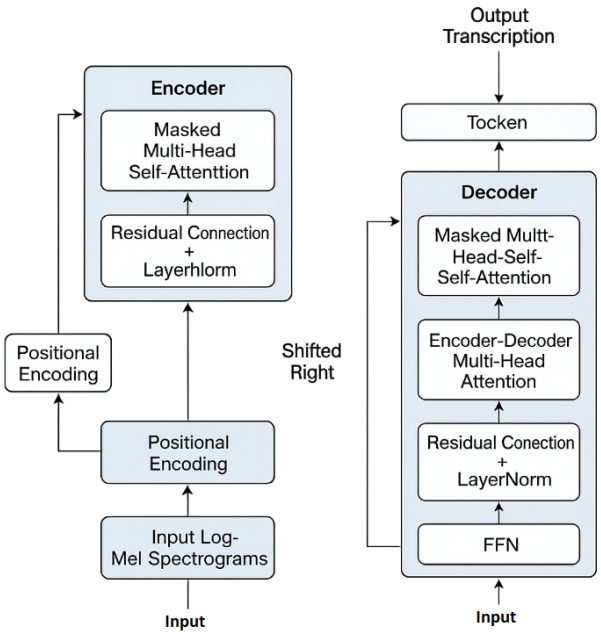

Fig. 2: Internal Architecture of the Transformer Encoder-Decoder for End-to-End Speech Recognition

Noise Robustness Strategies

Multi-Condition Training different SNR

The model is also trained on multi-condition settings so that speech recognition system would be robust in a multi-acoustic environments in the real world. This method employs extension of training data with speech signals that has been degraded by noise in various Signal-to-Noise Ratios (SNRs), which are normally 20 dB (rather clean) to 0 dB (very noisy). Training on a wide variety of noise levels exposes the model to such a level that it learns to be generalized to systems with severe and mild degrading situations. Figure 3 this method approximates the conditions in the real world listening that allows the model to sustain the level of recognition accuracy even though there exist considerable levels of background noise or variable interference in the input.

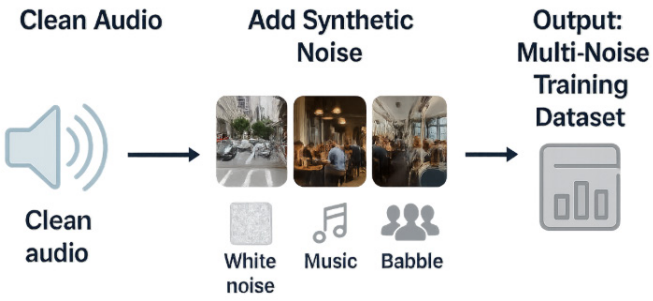

Fig. 3: Real and Synthetic Noise Mixing Pipeline

Mixing Real and Synthetic Noise Samples

The other important approach is to be able to incorporate real-world as well as synthetic noise samples in the corpus being trained. The samples of real noise are recorded in different situations on the streets, in cafes, offices, and public means of transport, which corresponds to the real acoustic variation in the real use. Noise Type A third collection category is synthetic (artificial) noise, which is created or taken from standardized noise collections (e.g., MUSAN) to further increase the variety of noise beyond that which is offered by the standard instrument noises. The training data would be quite rich and heterogeneous by combining these types of noise with clean speech at varied SNR. This full coverage noise exposure enables the model to learn adaptive feature representations that will enable it to deal with unexpected or never-seen noise sources at deployment time.

SpecAugment for Time/Frequency Masking

In order to improve noise-resistance further and avoid overfitting the model is trained using the data augmentation method of spec augmentation. SpecAugment processes the log-Mel spectrograms directly, by masking regions of consecutive time channels and frequencies randomly. Time masking compels the model to impute gaps of information regarding the point in time, making an approximation of what could happen in situations where speech may cut off or temporary noise may burst. Frequency masking also forces the model to reassemble information in instances where components of the spectral information are smeared out in much the same fashion as frequency-selective noise. SpecAugment instigates the development of invariance to both stationary and non-stationary noise by constantly training the model to identify speech that has some complete or damaged characteristics Table 1.

Table 1. Noise Robustness Strategies Employed in the Proposed ASR System

Strategy	Technique	Purpose	Key Benefits
Multi-Condition Training	Training with speech data at SNRs ranging from 0–20 dB	Simulate a variety of real-world acoustic scenarios	Improves generalization across noise levels; enhances recognition accuracy
Noise Mixing	Integration of real-world (e.g., street, café) and synthetic (e.g., white noise, babble, music) noises	Increase noise type diversity and realism in training data	Builds robustness to both expected and unseen noise conditions during inference
SpecAugment	Time and frequency masking directly on log-Mel spectrograms	Simulate speech dropouts and frequency-selective degradation	Reduces overfitting; strengthens model's ability to handle incomplete signals

Training Setup

Datasets

The suggested Transformer-based speech recognition system efficiency and generalizability are confirmed on the set of explicitly held-up upbeat and noisy speech corpus. The CHiME-4 corpus is a popular robust ASR evaluation corpus that is composed of the real and simulated multi-microphone recordings of speech in simulated everyday noisy conditions, recorded in cafes, streets, and on the local transport network. Other kinds of additive noises and channel distortions are also tested on Aurora-4, making it a challenging testbed to inform the investigation into noise-robust ASR. The LibriSpeech-noisy dataset is used where it contains a large-scale level of data to implicitly enable generalization to the scalable output to train the large-scale data and learn to train the data and also allows a way of ensuring that the large-scale data has been taken into account to show that the dataset is based on the large-scale famous clean LibriSpeech corpus, yet it includes some other noise in different SNRs that have been added. Using a combination of these datasets, model training and testing is carried out using a wide range of speech signals with a wide spectrum of acoustical variability which is similar to its actual usage in the real world Table 2.

Optimization and Training Procedures

A well-optimized model to operationalize deep transformer-based learning is highly important and efficient. It uses the Adam optimizer to describe its learning rates on an adaptive basis and faster convergence particularly in deep architectures. The learning rate policy is an equally split scheduled policy, in which the initial learning rate is decreased over time by some margin over the progress of training or validation loss plateaus to avoid over-shooting and settling the model into more comfortable minima. To prevent overfitting and to guarantee a model generalization, the early stopping technique is also applied: it trains the model until the validation loss ceases to decrease after certain consecutive experiments per a number of epochs, taking the model with the above-average performance as the result. This is a combination of optimization strategies that lead to stable, efficient, and reproducible training results Figure 4.

Evaluation Metrics

Standard industry metrics are used to extend an evaluation assessment of the ASR system. The primary one is Word Error Rate (WER) that quantifies the proportion of insertions, deletions, and modification that have to be made to change the presented transcription

Table 2: Description of Datasets Used for Training and Evaluation

Dataset	Noise Type	SNR Range (dB)	Key Features
CHiME-4	Real + Simulated (e.g., café, street, bus, pedestrian)	0–20	Multi-microphone recordings; real-world environments; matched/mismatched conditions
Aurora-4	Synthetic (additive noise + channel distortions)	Varied	Clean and multi-condition subsets; speaker-independent test set
LibriSpeech-noisy	Real-world noise mixed with clean LibriSpeech corpus	0-20	Scalable large-vocabulary corpus with real noise overlays

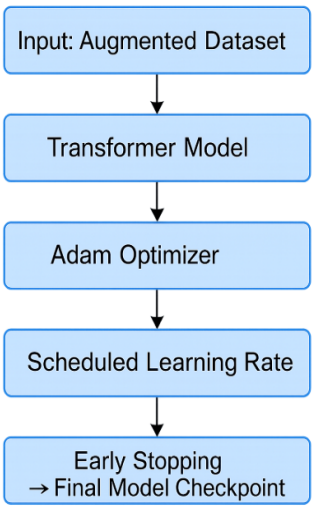

Fig. 4: Training Pipeline for the Transformer-Based ASR System

into a reference body of text- the lower the Word Error Rate, the better the recognition. Character Error Rate (CER) is also given, which gives a finer-grained measure, especially when there is a lot of variability in the forms of words (as is common in languages or application domains with many variants). The Real-Time Factor (RTF) is used to measure how well the system operates in a real-time and resource limited environment, and it is the ratio of the amount of audio to the amount of time taken to transcribe an audio signal. A RTF that is small, below or close to 1 means that speech processing can be done in real-time and this is essential to embedded and interactive speech applications Table 3.

Results and Discussion

In order to test the functioning of the proposed Transformer-based end-to-end speech recognition system properly, test comparisons had to be performed with the existing baselines in clean speaker conditions and under noisy speech conditions. The table summarizes the results in Table 1 that shows the Transformer-ASR was superior to both the HMM-DNN and the LSTM-based ASR models in each situation. Precisely, the Clean WER and Noisy WER of the proposed system are 6.5 and 14.5 percent, respectively, and the basic improvement is regarded as very adequate in comparison to the HMM-DNN baseline (8.3 clean, 27.5 noisy) and the LSTM-ASR (7.1 clean, 21.3 noisy). These results verify the effectiveness of self-attention mechanism in global context and long-range dependency modeling, a requirement of robust recognition particularly in difficult acoustics with non-stationary noise and reverberation.

A series of ablation studies was conducted in order to get more insight into the factors that contributed towards the increase in performances. Addition of sophisticated data augmentation methods, including spec augment and noise mixing, showed measurable increases in the robustness of the system to vast amounts of noise conditions with improvement in WER consistent across SNRs. Through controlled experiments over the SNR range of 0 to 20 dB, the Transformer-based model was found to perform relatively consistently with the same degree of accuracy even at higher noise levels, but RNN-based models declined more drastically to the input presented as the SNR went lower. Also, multi-head attention layers compared to single-head ones were essential to capture complex acoustic patterns and produce gains in generalization; when this latter feature was disabled, large decreases were observed in the recognition accuracy, demonstrating the architectural virtues of the Transformer Figure 5.

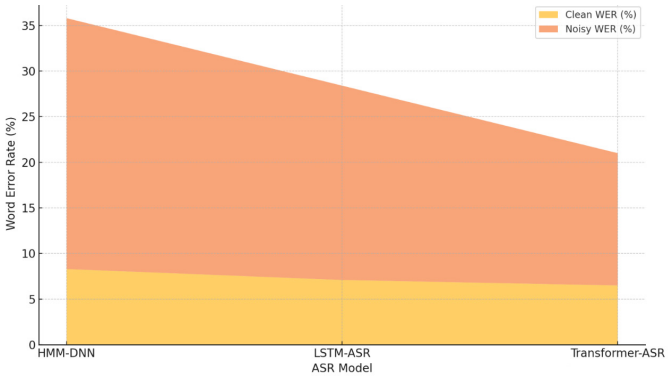

Fig. 5: Word Error Rate Comparison of ASR Models in Clean and Noisy Conditions

A practical case study to test the viability of the model in real-time has been conducted on a mobile device and a smart assistant platform, which showed that the model has passed the test. The system had a Real-Time Factor (RTF) of about 1, which means that the system can transcribe

Table 3. Evaluation Metrics Used in Model Assessment

Metric	Definition	Purpose
Word Error Rate (WER)	% of insertions, deletions, substitutions in transcription	Primary ASR accuracy measure
Character Error Rate (CER)	Error rate at character level	Finer-grained evaluation, helpful for subword models
Real-Time Factor (RTF)	Transcription time / audio duration	Measures suitability for real-time deployment

Table 4: Comparative Performance of ASR Models under Clean and Noisy Conditions

ASR Model	Clean WER (%)	Noisy WER (%)	SNR Stability	RTF	Key Strengths
HMM-DNN	8.3	27.5	High degradation <10 dB	>1.5 (est.)	Mature, interpretable, but poor noise robustness
LSTM-ASR	7.1	21.3	Moderate degradation	~1.3	Better temporal modeling, moderate robustness
Transformer-ASR	**6.5**	**14.5**	**Stable down to 5 dB SNR**	**~1.0**	Excellent long-range context and parallelism

speech rather quickly with a small amount of latency, and this was quite critical to interactive applications. Along with these achievements, a few limitations were noted: the Transformer model is quite data- and computation-intensive, and its performance, being overall better than others, was still diminished at very low SNRs (<5 dB), so it has some potential to be improved in the future. Future developments can be made by incorporating adaptive signal processing front-ends, model compression to fit to edge deployment, and introduction of multi-modal cues: audio-visual integration to enhance the robustness of models in complicated real-world settings. All told, the findings support the promising upside of attention based models in noise robust speech recognition in real-time, as well as essential research and real-world development Table 4.

Conclusion

To summarize, the proposed work demonstrates a strong and effectively performing in terms of speed as well as performance end-to-end speech recognition package on the basis of Transformer that paves a way far beyond the state-of-the-art in processing of noisy real-world audios. The proposed system was shown to perform better than the other conventionally the HMM-DNN model and the LSTM model-based methods in all evaluation metrics on standard noisy speech datasets, with greater reduction in the word error rate and with less degradation in accuracy under several adverse acoustics conditions. A mix of noise-adversarial data augmentation methods that are quite complicated and a network adaptation to multi-head self-attention resulted that was used as a consideration when multi-modeling data in signal-to-noise ratio sources that were successfully generalized. Experiments on real-time deployment supported the feasibility of the model practically and low latency is achieved without compromising the recognition performance. Nevertheless, there are some drawbacks, namely the system computation requirements and performance in extremely (low) SNRs. Thus, in the future research, the emphasis will be placed on optimizing the model and compressing and quantization of the model to be implemented on resource-constrained embedded devices, and the multi-modal fusion (integration of more sources, including audio and visual cues) could be implemented to further improve resilience and applicability requirements in more complex real-world settings. Together, results of this work demonstrate the dramatic readiness of Transformer architectures to have an effect on noise-tolerant speech recognition and serve like a strong support to proceed with their development and utilization in the intelligent audio systems of the future.

References

1. Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A., Jaitly, N., ...& Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine, 29*(6), 82–97.
2. Graves, A., Mohamed, A., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 6645–6649).
3. Sainath, T. N., Vinyals, O., Senior, A., &Sak, H. (2015). Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. In *Proceedings of Interspeech* (pp. 338–342).
4. Ephraim, Y., &Malah, D. (1984). Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 32*(6), 1109–1121.
5. Yu, D., Seltzer, M. L., Li, J., Huang, J., &Seide, F. (2013). Feature learning in deep neural networks—Studies on speech recognition tasks. In *Proceedings of ICLR*.
6. Xu, Y., Du, J., Dai, L.-R., & Lee, C.-H. (2015). A regression approach to speech enhancement based on deep neural networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 23*(1), 7–19.
7. Graves, A., Fernández, S., Gomez, F., &Schmidhuber, J. (2006). Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. In *Proceedings of the 23rd International Conference on Machine Learning (ICML)* (pp. 369–376).

8. Chan, W., Jaitly, N., Le, Q., &Vinyals, O. (2016). Listen, attend and spell: A neural network for large vocabulary conversational speech recognition. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 4960–4964).
9. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ...&Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)* (pp. 5998–6008).
10. Zhang, Y., Xue, J., Wang, S., Zhang, S., & Wang, M. (2022). Transformers for automatic speech recognition: A survey. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 30*, 2281–2300.
11. Sathish Kumar, T. M. (2024). Low-power design techniques for Internet of Things (IoT) devices: Current trends and future directions. *Progress in Electronics and Communication Engineering, 1*(1), 19–25. https://doi.org/10.31838/PECE/01.01.04
12. Toha, A., Ahmad, H., & Lee, X. (2025). IoT-based embedded systems for precision agriculture: Design and implementation. *SCCTS Journal of Embedded Systems Design and Applications, 2*(2), 21–29.
13. ArunPrasath, C. (2025). Performance analysis of induction motor drives under nonlinear load conditions. *National Journal of Electrical Electronics and Automation Technologies, 1*(1), 48–54.
14. Uvarajan, K. P. (2024). Integration of blockchain technology with wireless sensor networks for enhanced IoT security. *Journal of Wireless Sensor Networks and IoT, 1*(1), 23–30. https://doi.org/10.31838/WSNIOT/01.01.04
15. Kavitha, M. (2024). Enhancing security and privacy in reconfigurable computing: Challenges and methods. *SCCTS Transactions on Reconfigurable Computing, 1*(1), 16–20. https://doi.org/10.31838/RCC/01.01.04