Proceedings of CONF-CDS 2026 Symposium: Computer Vision-Based Multimodal Cognitive Load Estimation for Adaptive Media Communication
DOI: 10.54254/2755-2721/2026.GL35472

AI-Enhanced Noise Interference Mitigation and Signal Loss Recovery in Communication Transmission: A Case-Based Study

Jian Hu

School of Communication Engineering, Heriot-Watt University, Edinburgh, UK
jh2113@hw.ac.uk

Abstract. With the rapid development of information technology, communication systems are increasingly required to support stable, reliable, and real-time information transmission in complex environments. However, during practical communication transmission, signals are often affected by noise interference, channel fading, network jitter, multi-user interference, and packet loss. These problems may reduce speech intelligibility, increase bit error rates, cause video freezing, and weaken the continuity and reliability of communication services. Traditional enhancement techniques rely on stationary assumptions or fixed redundancy, and thus perform poorly under non-stationary noise or dynamically fluctuating network conditions. In recent years, deep learning-based methods have shown strong potential in learning nonlinear mappings from degraded to clean signals, offering better adaptability to diverse distortions. Nevertheless, the analysis shows that AI technology still faces practical deployment challenges, including high computational complexity, heavy dependence on large-scale labeled training data, and limited cross-scenario generalization capability, which require further optimization before such approaches can be reliably integrated into real-time audio/video communication systems.

Keywords: artificial intelligence, communication transmission, noise interference, signal loss, packet loss concealment

1. Introduction

With the rapid development of information and communication technologies, modern communication systems are applied in mobile networks, multimedia, and real-time audio/video services [1]. However, in practical environments, signals are inevitably affected by channel noise, fading, and packet loss, which degrade quality, increase latency and error rates [2]. These issues are critical in real-time systems where both reliability and low latency are required. Traditional techniques such as spectral subtraction, Wiener filtering, and MMSE estimators aim to suppress noise based on stationary or Gaussian assumptions [3, 4], while ARQ and FEC (standardized in RTP) are used for packet loss recovery [2].

However, these methods often degrade under non-stationary or dynamic conditions due to simplified assumptions. In recent years, deep learning has significantly advanced speech and signal

DOI: 10.54254/2755-2721/2026.GL35472

enhancement research. Neural network-based approaches, including convolutional, recurrent, and attention-based models, have demonstrated strong capability in learning complex mappings between noisy and clean signals [5]. Compared with traditional model-based methods, these data-driven approaches can better adapt to diverse noise distributions and channel variations. Based on real-time audio/video communication scenarios, this paper investigates the impact of noise interference and signal loss, and further explores AI-based optimization strategies for improving robustness and transmission quality.

2. Relevant theoretical foundation

2.1. Communication transmission

Communication transmission refers to the process of transferring information from a transmitter to a receiver through wired or wireless channels [1]. In practical communication systems, information such as audio, video, and digital data is transmitted through multiple stages including source encoding, channel encoding, modulation, transmission, demodulation, and decoding.

In ideal cases, the communication channel is assumed to be linear, time-invariant, and noise-free. However, real-world environments introduce various impairments such as thermal noise, electromagnetic interference, multipath fading, and bandwidth limitation, which lead to signal distortion and performance degradation [1]. To evaluate communication quality, common metrics include signal-to-noise ratio (SNR), bit error rate (BER), packet loss rate (PLR), and end-to-end delay [2]. These metrics jointly reflect the reliability, stability, and responsiveness of communication systems, especially in real-time scenarios such as voice calls and video conferencing where latency is critical [2].

2.2. Unified challenges and AI motivation

In practical communication systems, two major challenges significantly degrade performance: noise interference and signal loss.

Noise interference originates from environmental sound, device-level electronic noise, and channel-related disturbances [3]. It reduces signal clarity and increases decoding errors, particularly under low SNR conditions. Classical methods such as spectral subtraction and Wiener filtering attempt to suppress noise based on stationary assumptions [3], while MMSE-based estimators further improve performance under Gaussian noise models [4]. However, these methods are limited in non-stationary and highly dynamic environments due to inaccurate statistical modeling.

Signal loss is mainly caused by packet loss, channel fading, network congestion, and unstable wireless links [2]. In real-time communication, it leads to interruptions, freezing, and audio-video desynchronization. Traditional solutions include automatic repeat request (ARQ) and forward error correction (FEC), which improve reliability at the cost of increased latency or bandwidth consumption [2]. These trade-offs make them unsuitable for strict low-latency scenarios.

To address these limitations, artificial intelligence provides a data-driven optimization paradigm. Deep learning methods can learn noise distributions, channel dynamics, and packet loss patterns directly from data without relying on strict assumptions [5]. CNN-based models extract spectral features, recurrent networks model temporal dependencies, and attention mechanisms improve feature selection. Compared with traditional methods, AI-based approaches offer higher adaptability and robustness in complex communication environments.

DOI: 10.54254/2755-2721/2026.GL35472

3. Problem analysis

3.1. Noise interference

Communication signals are easily affected by various types of noise during transmission. Common noise sources include environmental noise, electronic equipment interference, narrowband interference, and non-stationary noise. Environmental noise usually comes from the external communication environment, such as traffic noise, human voices, or background sounds. Electronic interference may be related to communication devices, electromagnetic equipment, or adjacent channels. Narrowband interference is concentrated in a specific frequency range and may seriously affect certain communication bands. Compared with stable noise, non-stationary noise is more difficult to process because its intensity, frequency distribution, and temporal characteristics may change rapidly [6].

Noise interference can distort the received signal and reduce the distinction between the useful signal and interference components. In voice communication, noise may reduce speech intelligibility and make it difficult for users to understand the speaker. In data communication, noise may increase the bit error rate and reduce the accuracy of signal decoding. When the signal-to-noise ratio is low, the useful signal may be partly masked by noise, making traditional restoration more difficult. Traditional methods such as spectral subtraction, Wiener filtering, and minimum mean square error estimation can suppress noise under stable conditions, but they often perform poorly when noise changes dynamically. Therefore, noise interference remains a key factor that limits communication quality [7].

3.2. Signal loss

Signal loss is another major problem that affects the integrity and continuity of communication. It may be caused by channel fading, network jitter, multi-user interference, unstable links, network congestion, or packet loss during transmission. In wireless communication and real-time audio and video communication, the receiver may fail to obtain complete information when the terminal is moving, the network load fluctuates, or the channel condition deteriorates. This may result in interrupted voice, frozen images, missing data, or increased delay [8].

Signal loss not only reduces user experience but also affects the reliability and real-time performance of communication systems. Traditional methods usually use automatic repeat request or forward error correction to compensate for lost information. Automatic repeat request can improve transmission correctness by retransmitting lost data, but it increases latency. Forward error correction can recover errors or losses within a certain range, but its ability is limited when the packet loss rate is high or the channel changes sharply. In addition, redundant coding may increase bandwidth consumption. Therefore, in complex communication environments, signal loss requires more intelligent and adaptive recovery methods.

4. Case analysis of real-time audio and video communication

Real-time audio and video communication, widely used in online meetings, remote education, live streaming, mobile voice calls, and video conferencing, demands high continuity, low latency, and good clarity. The transmission process involves collection, encoding, network transmission, decoding, and playback; any interference at any stage degrades final quality. In this scenario, noise interference is mainly reflected in enhanced background noise, unclear speech, echo, equipment

DOI: 10.54254/2755-2721/2026.GL35472

noise, and video signal distortion. For example, during a mobile call or online meeting, traffic sounds, human voices, keyboard noises, and other disturbances mix with the useful speech, reducing intelligibility when the signal-to-noise ratio is low. Traditional denoising methods can suppress some stable noise, but their performance is often limited when dealing with sudden noise, complex background noise, or dynamically changing noise [9].

Signal loss is also common, caused by network congestion, wireless channel fluctuation, terminal movement, or multi-user access, leading to packet loss or transmission delay. In practice, voice may become intermittent, delayed, or silent, while video may freeze, blur, show mosaic artifacts, or lose synchronization with audio. Since real-time communication is sensitive to delay, traditional retransmission is not always suitable. Noise interference and signal loss often co-occur, jointly reducing clarity and continuity. Traditional methods are useful under stable conditions but are less adaptive in complex, dynamic, and latency-sensitive environments. Therefore, AI technology is needed to learn noise characteristics, channel changes, and packet loss patterns, thereby improving anti-interference ability and signal recovery capability [10].

5. Application and effectiveness analysis of AI-based optimization

5.1. AI-based optimization framework

For real-time audio and video communication, an AI-based collaborative optimization framework can be built to address both noise interference and signal loss. The general idea is to preprocess the collected audio, video, or communication signals first, extract representative features, and then use deep learning models to complete noise suppression and signal recovery. In the feature extraction stage, Mel-frequency cepstral coefficients, spectrograms, time-domain features, and frequency-domain features can be used to describe signal states. These representations help the model distinguish useful signal components from noise. In the model processing stage, CNN, BiLSTM, attention mechanisms, and generative models can be combined to learn noise variation patterns and signal loss behaviors. CNN can extract local patterns from spectrograms or signal features. BiLSTM can model both forward and backward temporal dependencies in communication signals. Attention mechanisms allow the model to focus more on important signal components. Generative models can assist in reconstructing missing or degraded signals. Through this process, AI methods can improve signal clarity, reduce noise impact, and enhance communication continuity [11].

5.2. Deep learning-based noise suppression

In noise suppression, deep learning models enhance complex noise processing. CNNs identify local spectral structures and noise distributions across frequency regions, while BiLSTM captures temporal evolution of speech signals. Their combination enables analysis of both local characteristics and long-term dependencies, beneficial for real-time audio communication where noise varies over time. Attention mechanisms further improve performance by assigning higher weights to informative signal regions, focusing on useful speech content and reducing irrelevant interference. Unlike traditional filtering, AI-based methods learn noise-signal differences from training data, adapting better to non-stationary, sudden, and background noises. This improves speech clarity, signal-to-noise ratio, and user intelligibility in real-time communication. However, actual effectiveness depends on training data quality and diversity; insufficient noise types may degrade performance in new environments [12].

DOI: 10.54254/2755-2721/2026.GL35472

5.3. Sequence prediction-based signal recovery

For signal recovery, LSTM and other sequence prediction models can be used to predict and compensate for missing data packets or signal fragments. Real-time audio and video data have temporal continuity. The current speech or video frame is usually related to previous and subsequent frames. Therefore, AI models can infer the possible features of missing parts based on existing signal fragments. This can reduce voice interruption, video freezing, and short-term silence caused by packet loss. AI-based recovery can also be combined with adaptive forward error correction. Traditional forward error correction uses redundant coding to improve recovery ability, but a fixed redundancy ratio may not be efficient in all network conditions. With AI-based prediction, the system can estimate the current network status and dynamically adjust the redundancy level. When packet loss is high, the system can increase redundancy to improve recovery ability. When the channel is stable, it can reduce redundancy to save bandwidth. Compared with pure retransmission-based methods, AI-assisted recovery is more suitable for real-time communication because it can reduce waiting time and maintain smoother interaction [13].

6. Model optimization and improvement

Although AI technology can improve noise suppression and signal recovery, further optimization is still necessary for practical application. First, model lightweighting should be considered. Methods such as model pruning, parameter sharing, knowledge distillation, and network compression can reduce the number of parameters and computational complexity. This can make AI models more suitable for mobile terminals and edge devices. Lightweight models are especially important for real-time communication because the system must process signals quickly without causing additional delay.

Second, edge computing can be introduced to improve deployment efficiency. Instead of sending all signal processing tasks to cloud servers, part of the computation can be performed near the user side. This can reduce transmission pressure, lower communication delay, and improve service responsiveness. Edge nodes can process noise suppression or preliminary signal recovery tasks, while cloud servers can be used for model updating and more complex computation [14].

Third, transfer learning and federated learning can be used to improve cross-scenario adaptability. Transfer learning allows a model trained in one communication environment to adapt more quickly to another environment with limited additional data. Federated learning enables multiple devices or users to jointly train a model without directly sharing private data. This is valuable for communication applications because real user data may involve privacy and security concerns. Through these optimization methods, AI-based communication enhancement can become more practical, efficient, and reliable [15].

7. Conclusion

This paper presents a study on the application of artificial intelligence techniques in mitigating noise interference and signal loss in communication transmission systems. However, several limitations should be acknowledged. First, the study is primarily based on a literature review and theoretical analysis, without large-scale experimental validation or real-world deployment results. Therefore, quantitative performance comparisons, such as improvements in SNR, BER, or latency reduction under different conditions, are not fully provided. Second, the discussion of AI-based methods remains conceptual, focusing on general model architectures (e.g., CNN, LSTM, and attention

DOI: 10.54254/2755-2721/2026.GL35472

mechanisms) rather than implementing a specific end-to-end system or optimized framework. Third, the evaluation of different methods is not supported by unified benchmark datasets or standardized experimental settings, which limits the depth of comparative analysis.

In future work, the study will focus on building a more complete experimental framework for communication signal enhancement. Specifically, real-world datasets containing diverse noise types and channel conditions will be introduced to enable quantitative evaluation. In addition, an end-to-end AI-based communication enhancement model will be implemented and tested under controlled simulation and practical network environments. Performance metrics such as signal-to-noise ratio improvement, packet loss concealment accuracy, and end-to-end latency will be systematically measured.

Furthermore, future work will explore lightweight model design for deployment on edge devices, aiming to balance model performance and computational efficiency. Optimization strategies such as model compression and adaptive inference mechanisms will also be investigated to improve real-time applicability in resource-constrained communication systems.

References

[1] Proakis, J. G., & Salehi, M. (2008). *Digital communications* (5th ed.). McGraw-Hill.
[2] Schulzrinne, H., Casner, S., Frederick, R., & Jacobson, V. (2003). RTP: A transport protocol for real-time applications (RFC 3550). IETF. https://www.rfc-editor.org/rfc/rfc3550
[3] Boll, S. F. (1979). Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. https://doi.org/10.1109/TASSP.1979.1163209
[4] Ephraim, Y., & Malah, D. (1984). Speech enhancement using MMSE short-time spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. https://doi.org/10.1109/TASSP.1984.1164453
[5] Valin, J.-M. (2018). Hybrid DSP/deep learning speech enhancement. *Interspeech*. https://arxiv.org/abs/2008.04259
[6] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x
[7] Haykin, S. (2009). *Communication systems* (5th ed.). Wiley.
[8] Oppenheim, A. V., & Schafer, R. W. (2010). *Discrete-time signal processing* (3rd ed.). Pearson.
[9] Loizou, P. C. (2013). *Speech enhancement: Theory and practice*. CRC Press.
[10] Benesty, J., Sondhi, M., & Huang, Y. (Eds.). (2007). *Springer handbook of speech processing*. Springer.
[11] Pascual, S., Bonafonte, A., & Serrà, J. (2017). SEGAN: Speech enhancement generative adversarial network. *Interspeech*. https://arxiv.org/abs/1703.09452
[12] Choi, H.-S. et al. (2019). Phase-aware speech enhancement with deep complex U-Net. *ICLR*. https://arxiv.org/abs/1903.03107
[13] Hu, Y. et al. (2020). DCCRN: Deep complex convolution recurrent network. *Interspeech*. https://arxiv.org/abs/2008.00264
[14] Pal, M. et al. (2024). Speech enhancement deep-learning architecture for edge processing. https://arxiv.org/abs/2405.16834
[15] Mohamed, M. M. et al. (2020). Deep speech packet loss concealment: A mini-survey. https://arxiv.org/abs/2005.07794