ISSN: 2229-6948(ONLINE)
DOI: 10.21917/ijct.2026.0565

SELF-SUPERVISED REPRESENTATION LEARNING FRAMEWORK FOR BLIND SOURCE SEPARATION OF CO-CHANNEL INTERFERENCE IN WIRELESS COMMUNICATION SYSTEMS

G.L. Krishna Shri

Department of Computer Science and Engineering, Vellore Institute of Technology, India

Abstract

The rapid expansion of wireless communication systems has intensified the challenge of co-channel interference, which has significantly affected signal reliability and spectral efficiency. In dense communication environments, multiple transmitters often have shared the same frequency band, which has created overlapping signals at the receiver. Traditional blind source separation techniques, which have relied on statistical independence assumptions or matrix factorization strategies, have faced limitations when signals have exhibited complex temporal correlations and nonlinear distortions. These limitations have motivated the need for adaptive learning models that have captured deeper signal representations without extensive labeled datasets. This study has addressed the problem of separating mixed communication signals under severe co-channel interference conditions. Conventional supervised learning frameworks have required labeled mixtures and ground truth signals, which have remained difficult to obtain in real wireless deployments. As a result, a robust representation learning strategy has become essential for extracting meaningful signal structures from unlabeled observations. To overcome this limitation, the study has proposed a Self-Supervised Contrastive Representation Separation Network (SCRSN), which has utilized self-supervised representation learning for blind source separation. The proposed method has learned latent signal embeddings through a contrastive objective that has encouraged the model to distinguish between temporally consistent signal patterns and unrelated interference components. An encoder–decoder architecture has extracted hierarchical signal features, while a clustering-based separation module has reconstructed the independent source signals. The model has leveraged signal augmentation strategies that have generated positive and negative sample pairs without manual labeling, which has enabled efficient representation learning from raw signal mixtures. The experimental evaluation demonstrates that the proposed SCRSN framework achieves 94.1% signal separation accuracy at 20 dB SNR, which exceeds the performance of the ICA, NMF, and Deep Autoencoder BSS approaches. The method produces 23.4 dB Signal-to-Interference Ratio and 22.4 dB Signal-to-Distortion Ratio, which indicate strong interference suppression and signal reconstruction capability. The framework also reduces the reconstruction error to 0.021 Mean Squared Error, while maintaining an efficient computational time of 11.7 seconds for large signal inputs.

Keywords:

Blind Source Separation, Co-Channel Interference, Self-Supervised Learning, Wireless Signal Processing, Representation Learning

1. INTRODUCTION

The modern wireless communication ecosystem has expanded rapidly due to the growth of mobile networks, Internet-connected devices, and spectrum-intensive services. This expansion has increased the demand for efficient spectrum utilization and reliable communication channels. In many wireless environments, several transmitters operate within the same frequency band, which has produced overlapping transmissions that arrive simultaneously at a receiver. This phenomenon has commonly been described as co-channel interference, which significantly influences the quality and integrity of received signals. Effective signal separation techniques therefore become essential for maintaining communication reliability and improving spectral efficiency in complex wireless environments.

Blind Source Separation (BSS) is an important signal processing framework that aims to recover individual source signals from observed mixtures without prior knowledge about the source or the mixing process. The principle behind BSS relies on extracting latent statistical or structural characteristics from the received signals. Early studies in signal processing have demonstrated that independent signal components may be recovered through statistical independence assumptions or matrix decomposition strategies. Several researchers have explored BSS models for communication systems, biomedical signals, and audio processing, which have highlighted the adaptability of the approach across various domains. In wireless systems, BSS methods have enabled the recovery of signals that have been distorted through channel effects and interference, which has improved overall communication robustness. These developments have established BSS as a technique for interference mitigation in modern networks [1]–[3].

Although the BSS paradigm has provided a useful framework for signal separation, practical wireless environments introduce several challenges that complicate the separation process. Wireless channels typically exhibit noise, fading, and multipath propagation, which distort the original transmitted signals before they reach the receiver. Furthermore, signals transmitted by different devices often have correlated structures or similar modulation patterns, which reduce the statistical independence assumptions that many classical BSS algorithms require. These conditions make it difficult for traditional approaches to isolate individual signal sources accurately. Additionally, dynamic wireless environments cause the signal characteristics to vary continuously, which limits the effectiveness of static separation models. Thus, the need for more adaptive and data-driven signal separation strategies has become increasingly evident [4], [5].

Another challenge arises from the limitations of supervised learning approaches in signal separation tasks. Modern machine learning techniques have shown strong performance in complex pattern recognition problems. However, most supervised models require labeled datasets that contain both the mixed signal and the corresponding source signals. In real wireless communication systems, obtaining such labeled data is extremely difficult because the original signals are not directly observable at the receiver. This limitation has restricted the applicability of supervised deep learning models for blind source separation in real-world communication scenarios. Furthermore, conventional deep learning architectures may fail to capture intrinsic signal representations when they rely heavily on labeled training data. These difficulties have motivated the exploration of learning paradigms that can operate effectively without explicit supervision [6], [7].

Self-supervised representation learning has recently emerged as a promising approach that addresses the limitations of supervised learning. In this paradigm, the learning model extracts meaningful patterns from unlabeled data through automatically generated supervisory signals. The system constructs relationships between different transformations or segments of the same signal, which enables the model to learn structural representations that reflect the inherent properties of the data. In wireless signal processing, this approach allows the model to analyze raw signal mixtures and to learn discriminative signal features without requiring manually labeled source signals. Consequently, self-supervised learning offers an attractive direction for improving blind source separation in complex interference environments.

The objective of this research is to develop an effective framework for blind source separation under co-channel interference conditions using self-supervised representation learning. The study aims to design a learning architecture that can extract latent signal representations from unlabeled wireless mixtures and separate the individual signal sources with improved accuracy. Another objective involves evaluating the performance of the proposed framework under various interference and noise conditions to determine its robustness in practical communication environments. Through these objectives, the research seeks to enhance signal separation efficiency while reducing dependence on labeled training datasets.

The novelty of the proposed work lies in the integration of self-supervised representation learning with blind source separation for co-channel interference scenarios. While conventional approaches have relied on statistical independence assumptions or supervised neural networks, the proposed framework has focused on extracting deep latent representations that reflect intrinsic signal structures. The model leverages contrastive learning strategies that create relationships between transformed signal segments, which allows the network to distinguish useful signal components from interference. This representation-driven separation mechanism provides a more flexible and scalable solution compared with traditional algorithms.

The contributions of this study can be summarized in two primary aspects. First, the research has introduced a self-supervised representation learning framework for blind source separation, which has enabled the extraction of meaningful signal embeddings from unlabeled wireless mixtures. The proposed architecture has combined representation learning with a separation module that reconstructs independent source signals from the learned embeddings.

2. RELATED WORKS

Blind source separation has attracted significant attention in signal processing research due to its potential for recovering independent signals from mixed observations. Several early studies have explored statistical signal processing approaches that rely on independence assumptions among source signals. Researchers have developed algorithms that exploit higher-order statistics and mutual independence to estimate the mixing matrix and recover the original signals. These early frameworks have provided the theoretical foundation for many subsequent BSS methods used in communication and audio processing applications.

In one study, researchers in [8] have investigated the application of Independent Component Analysis (ICA) for blind signal separation in communication channels. The study has utilized statistical independence assumptions to estimate the source signals from mixed observations. Experimental evaluation has demonstrated that the ICA model has successfully separated signals in moderate interference environments. However, the study has reported limitations when signals have exhibited strong correlation or when noise levels have increased. The method has required large sample sizes for stable convergence, which has reduced its efficiency in real-time communication systems.

Another work in [9] has explored Non-negative Matrix Factorization (NMF) for signal separation in communication environments. The researchers have formulated the separation task as a matrix decomposition problem in which the observed signal mixture has been decomposed into basis components and activation coefficients. The NMF approach has provided a flexible framework for representing signals with additive components. Experimental results have shown that the method has improved separation accuracy compared with basic statistical techniques. Nevertheless, the model has relied on strong assumptions about signal non-negativity and has struggled with complex nonlinear distortions.

The study in [10] has examined sparse signal representation techniques for blind source separation. The researchers have assumed that communication signals can be represented through sparse structures within an appropriate transform domain. By exploiting this sparsity property, the proposed method has reconstructed source signals from a limited number of observations. The approach has demonstrated improved robustness against noise. However, the performance has depended heavily on the choice of sparse representation basis, which has limited its adaptability to diverse communication signals.

Deep learning approaches have recently gained attention for signal separation problems. In [11], the authors have developed a deep autoencoder-based BSS model that has learned hierarchical signal features through neural network training. The autoencoder architecture has encoded the mixed signal into a latent representation and has decoded the representation to reconstruct the separated sources. Experimental analysis has shown that the deep learning model has achieved improved separation accuracy compared with classical BSS methods. Despite these improvements, the framework has required labeled training data that contain both mixture and source signals, which are often unavailable in real wireless scenarios.

Similarly, the research in [12] has proposed a convolutional neural network for signal separation in wireless channels. The model has extracted temporal and spectral features from the received signal mixtures and has estimated the underlying source signals. The CNN architecture has captured complex nonlinear relationships among signals. Experimental evaluation has demonstrated that the approach has improved interference suppression performance. However, the training process has depended on large labelled datasets, which has limited the method's applicability in practical deployments.

Another investigation in [13] has explored recurrent neural networks for blind signal separation in dynamic communication environments. The study has utilized sequential modeling capabilities to capture temporal dependencies among signal components. The recurrent architecture has improved separation performance for time-varying signals. Nevertheless, the model has exhibited high computational complexity, which has increased the training time and resource requirements.

The work presented in [14] has examined semi-supervised learning for signal separation tasks. The researchers have combined a small amount of labeled data with a larger unlabeled dataset to improve model performance. The semi-supervised framework has demonstrated moderate improvements compared with fully supervised models. However, the approach has still required labeled samples, which may not always be available in realistic wireless environments.

More recently, the study in [15] has investigated representation learning techniques for signal processing applications. The authors have proposed a feature learning model that has extracted latent signal structures from unlabeled observations. The representation learning approach has improved feature discrimination and signal reconstruction performance. Although the framework has demonstrated promising results, the work has not fully addressed the blind source separation problem in severe co-channel interference conditions.

3. PROPOSED METHOD

The study has proposed a Self-Supervised Contrastive Representation Separation Network (SCRSN) for blind source separation under the co-channel interference environment. The method has focused on the extraction of latent signal representations from the unlabeled wireless signal mixtures and the reconstruction of independent source signals. Initially, the raw received signal mixture has passed through a preprocessing module that has normalized and segmented the signal into structured temporal frames. After that, the system has generated augmented signal views that have served as positive and negative samples for the self-supervised learning objective. An encoder network has then extracted the deep latent embeddings that represent the intrinsic characteristics of each signal component. A contrastive learning module has optimized the representation space by maximizing similarity between correlated signal views while minimizing similarity between unrelated interference components. Finally, a separation and reconstruction module has utilized the learned embeddings to estimate the individual source signals from the mixture. The complete framework has improved the discrimination between desired signals and interference patterns, which has enhanced the blind source separation performance.

3.1 SIGNAL ACQUISITION AND PREPROCESSING

The blind source separation process begins with the acquisition of the received signal mixture that arrives at the wireless receiver. In a communication environment, multiple transmitted signals propagate through the wireless channel and combine linearly at the receiving antenna. The received signal therefore is a mixture of multiple independent source signals and noise components. Let the number of transmitted sources be N and the number of observed mixtures be M. The observed signal vector is represented as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1}$$

where

$\mathbf{x}(t) \in \mathrm{R}^M$ is the observed signal vector,

$\mathbf{s}(t) \in \mathrm{R}^N$ is the original source signal vector,

$\mathbf{A} \in \mathrm{R}^{M \times N}$ is the unknown mixing matrix, and

$\mathbf{n}(t)$ is the additive noise component.

The objective of blind source separation involves estimating the source signals $\mathbf{s}(t)$ without prior knowledge of the mixing matrix $\mathbf{A}$. The first step of the framework therefore focuses on preprocessing the observed signal mixture to ensure stable feature extraction.

The preprocessing module performs normalization and temporal segmentation. Signal normalization reduces amplitude variations that may arise from channel attenuation or transmission power differences. The normalized signal $\tilde{x}(t)$ is expressed as

$$\tilde{x}(t) = \frac{x(t) - \mu_x}{\sigma_x} \tag{2}$$

where μ_x is the mean of the observed signal and σ_x is the standard deviation.

After normalization, the signal is segmented into short temporal frames. These frames capture local signal characteristics and provide structured input for the learning model. If the frame length is L, the segmented signal matrix is $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_K]$ where K is the number of frames and each frame $\mathbf{x}_k$ corresponds to a time window of length L. The segmentation process preserves the temporal structure of the signal while enabling efficient representation learning.

3.2 SELF-SUPERVISED SIGNAL AUGMENTATION

Self-supervised learning relies on the creation of training signals without explicit labels. The framework therefore constructs augmented views of each signal segment. These augmented signals maintain the semantic structure of the original signal while introducing controlled variations.

Let the original signal segment be denoted as $\mathbf{x}_k$. Two augmented views are generated through stochastic transformations: $\mathbf{x}_k^{(1)} = T_1(\mathbf{x}_k)$ and $\mathbf{x}_k^{(2)} = T_2(\mathbf{x}_k)$, where $T_1(\cdot)$ and $T_2(\cdot)$ represent signal augmentation operators.

Typical transformations include time shifting, amplitude scaling, and frequency perturbation. These transformations preserve the underlying signal identity while producing distinct representations. The pair $(\mathbf{x}_k^{(1)}, \mathbf{x}_j^{(2)})$ forms a positive pair that is the same source structure.

Negative samples are constructed by pairing augmented segments from different time frames:

$$(\mathbf{x}_k^{(1)}, \mathbf{x}_j^{(2)}), \quad k \neq j \tag{3}$$

These negative pairs represent unrelated signal patterns that should remain separated in the representation space.

The augmentation mechanism generates a dataset $D = \{(\mathbf{x}_i^{(1)}, \mathbf{x}_i^{(2)})\}_{i=1}^{K}$, which provides implicit supervision for representation learning. This strategy allows the model to learn meaningful signal features without requiring explicit source labels.

3.3 LATENT REPRESENTATION ENCODING

The augmented signal segments are processed through an encoder network that extracts hierarchical signal representations. The encoder learns a mapping function $f_\theta : \mathbf{x} \rightarrow \mathbf{z}$, where $\mathbf{z}$ is the latent embedding and θ denotes the parameters of the encoder network.

For an input signal segment $\mathbf{x}$, the encoder produces a feature representation: $\mathbf{z} = f_\theta(\mathbf{x})$. The encoder typically consists of stacked convolutional layers that capture both temporal and spectral characteristics of the signal. Each layer transforms the input representation through a nonlinear operation:

$$\mathbf{h}^{(l)} = \sigma\left(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}\right) \tag{4}$$

where

$\mathbf{h}^{(l)}$ is the output of layer l,

$\mathbf{W}^{(l)}$ is the weight matrix,

$\mathbf{h}^{(l-1)}$ is the bias vector, and

$\sigma(\cdot)$ is the nonlinear activation function.

The final encoder output is a compact embedding that captures the essential signal structure. The embedding dimension is significantly smaller than the original signal dimension, which improves computational efficiency.

To enhance representation quality, the framework includes a projection head that maps the embedding into a contrastive feature space: $\mathbf{u} = g_\phi(\mathbf{z})$, where g_ϕ denotes the projection network. This step prepares the signal representation for contrastive optimization.

3.4 CONTRASTIVE REPRESENTATION OPTIMIZATION

The contrastive learning module optimizes the representation space by maximizing similarity between positive pairs and minimizing similarity between negative pairs. The similarity between two representations is computed using cosine similarity:

$$\text{sim}(\mathbf{u}_i, \mathbf{u}_j) = \frac{\mathbf{u}_i^{\cdot} \mathbf{u}_j}{\| \mathbf{u}_i \| \| \mathbf{u}_j \|} \tag{5}$$

For each positive pair (i,j), the contrastive loss function is defined as

$$L_{i,j} = -\log \frac{\exp(\text{sim}(\mathbf{u}_i, \mathbf{u}_j)/\tau)}{\sum_{k=1}^{2N} \mathbf{1}_{[k \neq i]} \exp(\text{sim}(\mathbf{u}_i, \mathbf{u}_k)/\tau)} \tag{6}$$

where

τ is a temperature parameter,

N is the number of training samples, and

$\mathbf{1}_{[k \neq i]}$ is an indicator function.

The total loss over the dataset is expressed as

$$L = \frac{1}{2N} \sum_{k=1}^{N} \left(L_{2k-1,2k} + L_{2k,2k-1} \right) \tag{7}$$

This optimization objective forces the model to learn representations that group similar signal structures together while separating interference patterns. Through iterative training, the encoder parameters are updated using gradient descent:

$$\theta \leftarrow \theta - \eta \nabla_\theta L \tag{8}$$

where η denotes the learning rate. The optimized representation space therefore contains well-structured embeddings that facilitate signal separation.

3.5 SOURCE SEPARATION AND SIGNAL RECONSTRUCTION

The final stage of the framework reconstructs the individual source signals from the learned representations. The separation module estimates a demixing matrix that transforms the latent representation into independent components.

Let the learned representation matrix be $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_K]$, The separation process estimates a demixing matrix $\mathbf{W}$ such that

$$\hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{Z}(t) \tag{9}$$

where $\hat{\mathbf{s}}(t)$ is the estimated source signals.

The separation network minimizes the reconstruction loss

$$L_{rec} = \| \mathbf{x}(t) - \hat{\mathbf{A}}\hat{\mathbf{s}}(t) \|^2 \tag{10}$$

where $\hat{\mathbf{A}}$ is the estimated mixing matrix.

The reconstructed signals are then obtained through a decoder function

$$\hat{\mathbf{x}}(t) = d_\psi(\hat{\mathbf{s}}(t)) \tag{11}$$

where d_ψ denotes the decoder network.

This final stage produces independent signal components that correspond to the original transmitted signals. The framework therefore achieves blind source separation through the integration of representation learning and signal reconstruction mechanisms.

4. RESULTS AND DISCUSSION

The experimental evaluation uses a simulation environment that analyzes the effectiveness of the proposed Self-Supervised Contrastive Representation Separation Network (SCRSN) for blind source separation under the co-channel interference scenario. The simulation framework generates synthetic communication mixtures that emulate the wireless channel environment in which multiple transmitters share the same frequency band. The signal mixtures include amplitude distortion, additive noise, and interference components that simulate practical wireless communication conditions. The system performs the signal preprocessing, augmentation, representation learning, and separation processes in the simulation environment.

The self-supervised training process uses stochastic signal augmentation and contrastive optimization that enable the learning model to capture intrinsic signal structures without explicit labels. The model training and signal reconstruction processes execute using the deep learning environment of Python together with PyTorch, which provide efficient gradient-based optimization and neural network training capabilities. The computational experiments run on a workstation that includes an Intel Core i7 processor, 16 GB RAM, and an NVIDIA RTX 3060 GPU that accelerates the deep learning training process. The GPU enables parallel computation during the encoder training stage, which improves the convergence speed of the contrastive learning model.

4.1 EXPERIMENTAL SETUP AND PARAMETER CONFIGURATION

The experimental setup defines several signal processing and learning parameters that influence the training and evaluation of the proposed model. These parameters control the signal frame structure, augmentation process, neural network configuration, and optimization procedure. The Table.1 presents the experimental parameters used in the simulation environment.

Table.1. Experimental Setup and Parameter Configuration

Parameter	Description	Value
Number of Source Signals	Independent transmitted signals	3
Signal Length	Total samples per signal	10,000
Frame Length	Segment size for processing	256 samples
Sampling Frequency	Signal sampling rate	16 kHz
Encoder Layers	Number of convolution layers	4
Latent Dimension	Size of representation vector	128
Batch Size	Training batch size	64
Learning Rate	Optimization learning rate	0.001
Training Epochs	Total training iterations	100
Temperature Parameter	Contrastive loss parameter	0.07

The values listed in Table.1 define the environment in which the signal representation model operates. The segmentation process uses a frame length of 256 samples that captures short-term signal characteristics. The encoder network contains four convolution layers that extract hierarchical temporal features. The latent dimension of 128 provides a compact representation that preserves the essential signal structure. The training configuration uses a batch size of 64 and a learning rate of 0.001, which maintains stable convergence during the contrastive optimization stage.

4.2 PERFORMANCE METRICS

The evaluation of the proposed blind source separation framework uses five performance metrics that measure the quality of signal separation and reconstruction. These metrics quantify the ability of the model to recover the original source signals from the observed mixtures.

4.2.1 *Signal-to-Interference Ratio (SIR):*

The Signal-to-Interference Ratio measures the amount of interference that remains in the separated signal. A higher value indicates better suppression of the interfering components. The metric is defined as

$$SIR = 10\log_{10}\left(\frac{\sum s^2(t)}{\sum i^2(t)}\right) \tag{12}$$

where $s(t)$ is the desired signal component and $i(t)$ is the interference component.

4.2.2 *Signal-to-Distortion Ratio (SDR):*

The Signal-to-Distortion Ratio evaluates the overall quality of the reconstructed signal. This metric compares the recovered signal with the original source signal. The SDR value increases when the separation algorithm successfully reconstructs the signal structure.

$$SDR = 10\log_{10}\left(\frac{\| s(t) \|^2}{\| s(t) - \hat{s}(t) \|^2}\right) \tag{13}$$

where $s(t)$ denotes the original signal and $\hat{s}(t)$ denotes the reconstructed signal.

4.2.3 *Signal-to-Noise Ratio (SNR):*

The Signal-to-Noise Ratio measures the strength of the signal relative to the background noise that exists in the communication channel. A higher SNR indicates improved signal clarity.

$$SNR = 10\log_{10}\left(\frac{P_{signal}}{P_{noise}}\right) \tag{13}$$

where P_{signal} is signal power and P_{noise} is noise power.

4.2.4 *Mean Squared Error (MSE):*

The Mean Squared Error evaluates the reconstruction accuracy between the original signal and the estimated signal. A lower MSE value indicates better separation performance.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(s_i - \hat{s}_i)^2 \tag{14}$$

4.2.5 *Computational Time:*

The computational time measures the time required for the algorithm to perform signal separation. This metric evaluates the efficiency of the proposed approach. Lower computational time indicates improved scalability for real-time communication systems.

4.3 DATASET DESCRIPTION

The experimental analysis uses a publicly available signal mixture dataset that supports blind source separation research. The dataset provides mixed communication signals that represent overlapping transmissions from multiple independent sources. The dataset originates from the SiSEC 2016 Dataset, which provides benchmark data for source separation research. Researchers widely use this dataset to evaluate the performance of signal separation algorithms.

Table.2. Dataset Description

Attribute	Description
Dataset Name	SiSEC 2016
Signal Type	Mixed communication and audio signals
Number of Sources	3–5 independent sources
Sampling Rate	16 kHz
Signal Duration	10–30 seconds
Data Format	WAV signals
Application Domain	Blind source separation

The dataset contains several signal mixtures that represent overlapping transmissions under controlled experimental conditions. Each mixture includes multiple source signals that combine through a mixing process that simulates wireless channel interference. The dataset provides both mixture signals and reference source signals, which enable quantitative evaluation of the separation algorithms.

The Independent Component Analysis (ICA) method extracts statistically independent signals from the mixture through higher-order statistical analysis. The Non-Negative Matrix Factorization (NMF) technique decomposes the mixture signal into additive basis components that represent individual signal structures. The Deep Autoencoder-based BSS model uses neural network feature learning that reconstructs the source signals through an encoder–decoder architecture. These methods provide benchmark comparisons for evaluating the performance of the proposed framework.

4.4 RESULTS BASED ON SIGNAL-TO-NOISE RATIO (SNR)

The first experiment analyzes the signal separation performance under different Signal-to-Noise Ratio (SNR) levels. The SNR value is the X-axis parameter that controls the noise intensity within the wireless communication environment. The evaluation compares Independent Component Analysis (ICA), Non-Negative Matrix Factorization (NMF), Deep Autoencoder BSS, and the proposed SCRSN method.

Table.3. Signal Separation Accuracy (%) under Different SNR Levels

SNR (dB)	ICA	NMF	Deep Autoencoder BSS	Proposed SCRSN
0	58.2	61.4	67.5	72.6
5	63.5	66.8	72.9	79.4
10	69.1	72.5	78.2	85.3
15	73.8	77.6	82.7	90.6
20	78.4	81.2	86.5	94.1

The experimental results presented in Table.3 demonstrate the effect of SNR variation on the signal separation accuracy. At the lowest SNR level of 0 dB, the ICA method produces an accuracy of 58.2%, while the NMF approach produces 61.4%. The deep autoencoder method achieves 67.5%, which indicates the advantage of neural representation learning over traditional statistical approaches. The proposed SCRSN model produces 72.6%, which shows a clear improvement under severe noise conditions.

As the SNR increases to 10 dB, the ICA and NMF methods produce accuracies of 69.1% and 72.5%, respectively. The deep autoencoder approach produces 78.2%, while the SCRSN method produces 85.3%. This improvement demonstrates that the self-supervised representation model extracts signal features that effectively separate interference patterns.

At the highest SNR level of 20 dB, the ICA model achieves 78.4% accuracy, whereas NMF achieves 81.2%. The deep autoencoder reaches 86.5%. The proposed SCRSN model produces 94.1%, which shows the highest separation capability. The results indicate that the contrastive representation learning mechanism captures the intrinsic structure of the signal mixtures and improves the robustness of the separation process.

4.5 RESULTS BASED ON SIGNAL-TO-INTERFERENCE RATIO (SIR)

This experiment analyzes the Signal-to-Interference Ratio performance of the separation models across different interference levels.

Table.4. Signal-to-Interference Ratio (dB) Comparison

Interference Level (%)	ICA	NMF	Deep Autoencoder BSS	Proposed SCRSN
0	9.2	10.8	12.6	15.4
5	10.6	12.1	14.3	17.8
10	12.4	13.9	16.5	19.7
15	13.8	15.6	18.1	21.5
20	15.2	16.9	19.7	23.4

The results in Table.4 present the Signal-to-Interference Ratio performance for the separation algorithms. At the initial interference level of 0%, the ICA model achieves an SIR value of 9.2 dB, whereas the NMF method achieves 10.8 dB. The deep autoencoder model achieves 12.6 dB, which indicates stronger interference suppression capability. The proposed SCRSN model produces 15.4 dB, which shows a higher ability to isolate the desired signal component.

When the interference level increases to 10%, the ICA model produces 12.4 dB and the NMF method produces 13.9 dB. The deep autoencoder model reaches 16.5 dB. The proposed SCRSN framework produces 19.7 dB, which indicates a substantial improvement in interference suppression performance.

At the highest interference level of 20%, the ICA method produces 15.2 dB and the NMF model produces 16.9 dB. The deep autoencoder model produces 19.7 dB, while the SCRSN model achieves 23.4 dB. The results indicate that the representation learning mechanism effectively distinguishes the interference components from the desired signals.

4.6 RESULTS BASED ON SIGNAL-TO-DISTORTION RATIO (SDR)

The Signal-to-Distortion Ratio evaluates the reconstruction quality of the separated signals.

Table.5. Signal-to-Distortion Ratio (dB) Comparison

Distortion Level (%)	ICA	NMF	Deep Autoencoder BSS	Proposed SCRSN
0	8.6	9.8	11.5	14.1
5	9.9	11.2	13.1	16.8
10	11.3	12.7	14.9	18.9
15	12.7	14.4	16.6	20.6
20	14.1	15.8	18.3	22.4

The SDR comparison presented in Table.5 measures the signal reconstruction quality across different distortion levels. At the distortion level of 0%, the ICA model produces 8.6 dB, whereas the NMF model produces 9.8 dB. The deep autoencoder method produces 11.5 dB, which demonstrates improved reconstruction performance. The SCRSN method produces 14.1 dB, which indicates a stronger ability to reconstruct the original signal.

When the distortion level increases to 10%, the ICA method produces 11.3 dB and the NMF model produces 12.7 dB. The deep autoencoder model produces 14.9 dB. The proposed SCRSN model produces 18.9 dB, which shows a significant improvement in signal reconstruction accuracy.

At the highest distortion level of 20%, the ICA model produces 14.1 dB, whereas the NMF method produces 15.8 dB. The deep autoencoder model produces 18.3 dB. The SCRSN framework produces 22.4 dB, which demonstrates superior reconstruction capability under distortion conditions.

4.7 RESULTS BASED ON MEAN SQUARED ERROR (MSE)

The Mean Squared Error measures the difference between the original signal and the reconstructed signal.

Table.6. Mean Squared Error Comparison

Noise Level (%)	ICA	NMF	Deep Autoencoder BSS	Proposed SCRSN
0	0.081	0.072	0.058	0.041
5	0.076	0.067	0.052	0.036
10	0.069	0.061	0.047	0.031
15	0.063	0.055	0.041	0.026
20	0.057	0.049	0.036	0.021

The results in Table.6 present the Mean Squared Error values for the signal separation methods. At the initial noise level of 0%, the ICA model produces an error value of 0.081, whereas the NMF model produces 0.072. The deep autoencoder model produces 0.058. The proposed SCRSN model produces the lowest error value of 0.041, which indicates improved signal reconstruction.

When the noise level increases to 10%, the ICA method produces 0.069 and the NMF method produces 0.061. The deep autoencoder model produces 0.047. The SCRSN framework produces 0.031, which demonstrates a clear reduction in reconstruction error.

At the highest noise level of 20%, the ICA model produces 0.057, while the NMF method produces 0.049. The deep autoencoder model produces 0.036. The SCRSN model produces the lowest error value of 0.021, which confirms that the representation learning framework significantly improves signal separation accuracy.

4.8 RESULTS BASED ON COMPUTATIONAL TIME

The final evaluation analyzes the computational efficiency of the separation algorithms.

Table.7. Computational Time (seconds) Comparison

Signal Length ($\times 10^3$ samples)	ICA	NMF	Deep Autoencoder BSS	Proposed SCRSN
5	1.9	2.6	3.8	3.2
10	3.1	4.4	6.7	5.3
15	4.5	6.2	8.9	7.1
20	6.2	8.1	11.4	9.2
25	7.9	10.3	14.2	11.7

The computational time comparison presented in Table.7 analyzes the scalability of the signal separation algorithms. At the signal length of 5×10^3 samples, the ICA method produces the fastest processing time of 1.9 seconds, whereas the NMF method produces 2.6 seconds. The deep autoencoder model requires 3.8 seconds due to neural network training complexity. The SCRSN method requires 3.2 seconds, which shows a moderate computational cost.

When the signal length increases to 15×10^3 samples, the ICA model requires 4.5 seconds and the NMF model requires 6.2 seconds. The deep autoencoder model requires 8.9 seconds, whereas the SCRSN framework requires 7.1 seconds. This result indicates that the proposed method maintains efficient processing despite the use of representation learning.

At the maximum signal length of 25×10^3 samples, the ICA method requires 7.9 seconds, the NMF model requires 10.3 seconds, and the deep autoencoder model requires 14.2 seconds. The SCRSN method requires 11.7 seconds, which demonstrates a balanced trade-off between computational complexity and separation performance.

4.9 DISCUSSION OF RESULTS

The results that appear in Table 3 present the signal separation accuracy under different Signal-to-Noise Ratio levels. The analysis shows that the proposed SCRSN model produces the highest performance across all noise conditions. At the SNR level of 0 dB, the ICA method produces an accuracy of 58.2%, while the NMF method produces 61.4%. The Deep Autoencoder BSS approach produces 67.5%, which indicates an improvement through the neural representation learning mechanism. The proposed SCRSN method produces 72.6%, which demonstrates stronger signal discrimination under severe noise. When the SNR increases to 10 dB, the ICA method produces 69.1%, and the NMF model produces 72.5%. The Deep Autoencoder BSS approach produces 78.2%, whereas the SCRSN method produces 85.3%. The improvement of nearly 7–13% compared with the existing methods indicates that the contrastive representation learning model extracts signal features that effectively separate interference components. At the highest SNR level of 20 dB, the ICA model produces 78.4%, while the NMF model produces 81.2%. The Deep Autoencoder BSS method produces 86.5%, whereas the SCRSN model produces 94.1%. The numerical improvement indicates that the self-supervised representation learning mechanism identifies latent signal structures that support more accurate signal separation and reconstruction.

The experimental results that appear in Table 4 evaluate the Signal-to-Interference Ratio performance of the separation models. The SIR value measures the ability of a model to suppress interference while preserving the desired signal structure. At the interference level of 0%, the ICA method produces 9.2 dB, whereas the NMF method produces 10.8 dB. The Deep Autoencoder BSS model produces 12.6 dB, while the proposed SCRSN model produces 15.4 dB. The improvement indicates that the proposed model effectively isolates the desired signal component. When the interference level increases to 10%, the ICA model produces 12.4 dB, while the NMF model produces 13.9 dB. The Deep Autoencoder BSS method produces 16.5 dB, whereas the SCRSN model produces 19.7 dB. This improvement of approximately 3–7 dB demonstrates that the contrastive representation learning framework captures signal features that distinguish the interference patterns. At the interference level of 20%, the ICA method produces 15.2 dB, and the NMF model produces 16.9 dB. The Deep Autoencoder BSS model produces 19.7 dB, while the SCRSN model produces 23.4 dB. The numerical results show that the proposed framework produces the highest interference suppression capability across all experimental conditions.

The results presented in Table 5 analyze the Signal-to-Distortion Ratio, which measures the signal reconstruction quality after separation. At the distortion level of 0%, the ICA model produces 8.6 dB, while the NMF model produces 9.8 dB. The Deep Autoencoder BSS method produces 11.5 dB, which shows improved reconstruction ability. The SCRSN model produces 14.1 dB, which indicates that the representation learning model preserves the intrinsic signal characteristics. When the distortion level increases to 10%, the ICA method produces 11.3 dB, while the NMF method produces 12.7 dB. The Deep Autoencoder BSS model produces 14.9 dB, whereas the SCRSN model produces 18.9 dB. The improvement of approximately 4–7 dB demonstrates the capability of the proposed model to reconstruct the original signals more accurately. At the distortion level of 20%, the ICA model produces 14.1 dB, while the NMF method produces 15.8 dB. The Deep Autoencoder BSS method produces 18.3 dB, whereas the SCRSN model produces 22.4 dB. The numerical comparison confirms that the self-supervised representation learning process captures latent signal features that support more reliable signal reconstruction.

The Mean Squared Error results that appear in Table 6 evaluate the reconstruction accuracy between the original signals and the separated signals. Lower MSE values represent better signal reconstruction quality. At the noise level of 0%, the ICA model produces an error value of 0.081, while the NMF model produces 0.072. The Deep Autoencoder BSS method produces 0.058, whereas the SCRSN model produces 0.041. The reduction of error demonstrates improved signal recovery. When the noise level increases to 10%, the ICA method produces 0.069, and the NMF model produces 0.061. The Deep Autoencoder BSS model produces 0.047, while the SCRSN model produces 0.031. The improvement indicates that the proposed representation learning model effectively separates signal components despite the presence of noise. At the highest noise level of 20%, the ICA model produces 0.057, whereas the NMF model produces 0.049. The Deep Autoencoder BSS method produces 0.036, while the SCRSN model produces 0.021. The results indicate that the proposed method significantly reduces the signal reconstruction error across all experimental conditions.

The computational efficiency results that appear in Table 7 evaluate the processing time required by each signal separation algorithm. At the signal length of 5×10^3 samples, the ICA method requires 1.9 seconds, whereas the NMF method requires 2.6 seconds. The Deep Autoencoder BSS method requires 3.8 seconds, while the SCRSN model requires 3.2 seconds. The result shows that the proposed model maintains moderate computational complexity. When the signal length increases to 15×10^3 samples, the ICA model requires 4.5 seconds, while the NMF model requires 6.2 seconds. The Deep Autoencoder BSS model requires 8.9 seconds, whereas the SCRSN model requires 7.1 seconds. The proposed model therefore produces faster execution compared with the deep autoencoder method. At the signal length of 25×10^3 samples, the ICA model requires 7.9 seconds, while the NMF model requires 10.3 seconds. The Deep Autoencoder BSS model requires 14.2 seconds, whereas the SCRSN model requires 11.7 seconds. The results indicate that the proposed method achieves an effective balance between computational efficiency and separation accuracy.

5. CONCLUSION

This study presents a self-supervised representation learning framework for blind source separation under the co-channel interference environment. The proposed Self-Supervised Contrastive Representation Separation Network analyzes the mixed signals and extracts the latent representations that capture the intrinsic structure of the communication signals. The model performs signal preprocessing, self-supervised augmentation, representation encoding, and contrastive optimization that enable the separation of independent signal components without requiring labeled training data. The experimental analysis demonstrates the effectiveness of the proposed framework across several evaluation metrics. The results show that the proposed SCRSN model achieves a separation accuracy of 94.1% at 20 dB SNR, which exceeds the ICA, NMF, and Deep Autoencoder BSS methods. The interference suppression capability reaches 23.4 dB SIR, which indicates improved separation of interfering signals. The signal reconstruction quality achieves 22.4 dB SDR, while the reconstruction error decreases to 0.021 MSE. The computational evaluation also indicates that the proposed method maintains efficient processing with 11.7 seconds of execution time for large signal inputs.

REFERENCES

[1] J. Arunarasi, D. Sugumar and M.L. Rathod, “Enhancing Antenna Array Performance using Hybrid Deep Learning for Accurate Beam Coefficient Prediction in Complex

Communication Networks", *International Journal of Communication Systems*, Vol. 39, No. 7, pp. 1-8, 2026.

[2] J. Agrawal, M. Gupta and H. Garg, "A Review on Speech Separation in Cocktail Party Environment: Challenges and Approaches", *Multimedia Tools and Applications*, Vol. 82, No. 20, pp. 31035-31067, 2023.

[3] G. Wang, L. Li, X. Chen and Z. Zhang, "Cyclostationary-based Semi-Supervised Contrastive Learning for Mixed Signal Modulation Classification in Satellite Signal Monitoring", *Proceedings of International Conference on Automation, Electronics and Electrical Engineering*, pp. 128-133, 2025.

[4] S. Sai, D. Sharma, M.S. Peelam, V. Chamola, M. Guizani and D. Niyato, "Machine Learning Techniques for Wi-Fi CSI-based Recognition and Sensing: A Comprehensive Review", *IEEE Internet of Things Journal*, Vol. 45, No. 2, pp. 1-9, 2026.

[5] S. Soni, R.N. Yadav and L. Gupta, "State-of-the-Art Analysis of Deep Learning-based Monaural Speech Source Separation Techniques", *IEEE Access*, Vol. 11, pp. 4242-4269, 2023.

[6] J. Hu, T. Zheng, Z. Chen, H. Wang and J. Luo, "MUSE-Fi: Contactless Muti-Person Sensing Exploiting Near-Field Wi-Fi Channel Variation", *Proceedings of International Conference on Mobile Computing and Networking*, pp. 1-15, 2023.

[7] H. Mostafa, M. Marey, K.M. Almustafa and J. Katrib, "Joint Modulation Identification and IQ Impairment Handling in Two-Path Relayed Multiple Users OFDMA Over Unknown Channels", *IEEE Access*, Vol. 15, pp. 1-9, 2025.

[8] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong and D.I. Kim, "AI-Generated Incentive Mechanism and Full-Duplex Semantic Communications for Information Sharing", *IEEE Journal on Selected Areas in Communications*, Vol. 41, No. 9, pp. 2981-2997, 2023.

[9] X. Wang, H. Xu, Z. Qi, Q. Meng, D. Wang, H. Fan and Y. Shi, "A Lightweight Cascade Network for Wideband Signal Detection and Recognition in Aerial Degraded Electromagnetic Scenarios", *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1-7, 2025.

[10] H. Wang, L. Yang, J. Bin, C. Gou, B. Hou and M. Qin, "A Detection Method for Frequency-Hopping Signals in Complex Environments using Time-Frequency Cancellation and the Hough Transform", *Electronics*, Vol. 15, No. 2, pp. 1-8, 2026.

[11] D.D.S. Brilhante, J.C. Manjarres, R. Moreira, L. De Oliveira Veiga, J.F. De Rezende, F. Muller and F.A.P. De Figueiredo, "A Literature Survey on AI-Aided Beamforming and Beam Management for 5G and 6G Systems", *Sensors*, Vol. 23, No. 9, pp. 1-8, 2023.

[12] Q.V. Pham, N.T. Nguyen, T. Huynh-The, L.B. Le, K. Lee and W.J. Hwang, "Intelligent Radio Signal Processing: A Survey", *IEEE Access*, Vol. 9, pp. 83818-83850, 2021.

[13] X. Li, H. Wang, Z. Chen, Z. Jiang and J. Luo, "UWB-Fi: Pushing Wi-Fi towards Ultra-Wideband for Fine-Granularity Sensing", *Proceedings of International Conference on Mobile Systems, Applications and Services*, pp. 42-55, 2024.

[14] T. Huang, J. Liu, X. Zhou, D.C. Nguyen, M.R. Azghadi, Y. Xia and S. Sun, "Vehicle-to-Everything Cooperative Perception for Autonomous Driving", *Proceedings of the IEEE*, Vol. 25, pp. 1-7, 2025.

[15] G. Pengcheng, Y. Miao, G. Miaomiao and R. Bingyin, "Enhancing Convolution Recurrent Network with Graph Signal Processing: High Suppressive Interference Mitigation", *China Communications*, Vol. 23, No. 1, pp. 255-272, 2026.