A Review on Deep Learning Enabled Massive MIMO Systems

Ass.Prof. Amrita Soni[1] , Tanmay Kanungo[2]
[1, 2] Ujjain Engineering College, India

Abstract- *Due to increasing number of users and multimedia applications, multiple input, multiple output (MIMO) systems have become indispensable from wireless technology. The benefit of MIMO systems is increase in the channel capacity at low bandwidths, however, the accurate reception of the MIMO signal at the receiver is extremely complex. This is due to the addition of noise and interference, as well as the different behavior of the wireless channel for the different channel paths. Due to the largeness and the complexity of the data, the technique used off late is the Deep Learning based approach for signal detection. This paper presents the basics of MIMO systems and insight into deep learning based MIMO channel estimation and signal detection techniques.*

Keywords- MIMO, Deep Learning, Channel Response, Artificial Neural Network (ANN), Deep Neural Network, Mean Square Error (MSE).

I. INTRODUCTION

In wireless communication, the receiver side BER strongly affected by channels noise, interference, distortion, synchronization error and wireless multipath fading channels, Multiple-input and multiple-output(MIMO) systems have resulted in higher spectral efficiency and capacity [1]. MIMO communications, the system is equipped with multiple antennas at both the transmitter and the receiver technique [2]. The multiple antenna scheme gives a more reliable performance through array gain, diversity and spatial multiplexing [3]. These concepts are briefly discussed below.

The growing demand of multimedia services and the progress of Internet related contents lead to increasing interest to high speed communications network [4]. The requirement for flexibility and wide bandwidth imposes the use of efficient transmission systems that would fit to the characteristics of wideband channels especially in wireless environment where the channel is very challenging process [5]. In wireless environment the signal is propagating since the transmitter to the receiver along number of different paths, collectively referred as multipath communication. While propagating the signal power drops of due to the following effects: a path loss, macroscopic fading and microscopic fading. The fading of the signal can be mitigated by different diversity methods [6]. For optimal detection of MIMO signals, deep learning and machine learning techniques are being explored.

II. MIMO SYSTEM MODEL

The basics of the MIMO system model are discussed below [7]-[8]:

Array Gain

In MIMO communications, array gain is the average increase in the SNR at the receiver that occurs from the coherent combining effect of the multiple antennas at the transmitter or receiver or both. Usually, multi antenna systems require a perfect knowledge of the channel either at the receiver or the transmitter or both to achieve an array gain. It is mathematically given by:

$$AG = \frac{SNR_{MIMO}}{SNR_{SISO}} \tag{1}$$

Here,
AG represents Array Gain
SNR stands for Signal to Noise Ratio
MIMO stands for multiple input multiple output
SISO stands for single input single output

Transmitter Array Gain: If the channel is known to the transmitter with multiple antennas, the transmitter will weigh the transmission with weights, depending on the channel coefficients, so that there is coherent combining at one antenna receiver. This is the MISO case. This type of array gain is called transmitter array gain. Mathematically, if the combined transmitted signal is given by:

$$S_{TX} = \sum_{i=1}^{i=n} S_i W_i \tag{2}$$

then the following ratio is called the transmitter array gain

$$AG_{TX} = \frac{S_{TX\,MIMO}}{S_{TX\,SISO}} \tag{3}$$

Here,
AG represents array gain

$S_{TX\ SISO}$ represents the transmitted SISO signal
$S_{TX\ MIMO}$ represents the transmitted MIMO signal
W_i represents the weights of the 'i' parallel transmission paths corresponding to the path co-efficient

Receiver Array Gain: If we have a single antenna at the transmitter and no knowledge of the channel and a receiver with multiple antennas, which has perfect knowledge of the channel, thenthe receiver can suitably weight the incoming signals so that they coherently add up at the output, thereby enhancing the signal. This is the SIMO case. This type of array gain is called receiver array gain.Similarly the receiver array gain is given by:

$$AG_{RX} = \frac{S_{RX\ MIMO}}{S_{RX\ SISO}} \quad (4)$$

Here,
AG represents array gain
$S_{RX\ SISO}$ represents the transmitted SISO signal
$S_{RX\ MIMO}$ represents the transmitted MIMO signal

Diversity Gain

In MIMO systems, the same information is transmitted from multiple transmit antennas and simultaneously received at multiple receive antennas. Since the fading for each link between a pair of transmit and receive antennas usually can be considered to be independent, the probability that the information is detected exactly is increased. Space time codes are designed to exploit following two resources.[11]

$$DOF = \min\,(N_T, N_R) \quad (5)$$

$$Diversity = N_T, N_R \quad (6)$$

Here,
DOF represents degree of freedom
N_T represents number of transmitters
N_R represents number of receivers

Multi Antenna Systems

Multi antenna systems can be classified as single-input multiple-output (SIMO), multiple-input single-output (MISO), and multiple-input multiple-output (MIMO) systems. In order to develop the input-output relations of SIMO, MISO, MIMO systems, we first describe the single-input single-output (SISO) system [14].

Single Input Single Output System

The schematic block diagram of a SISO system is depicted in Figure 1.

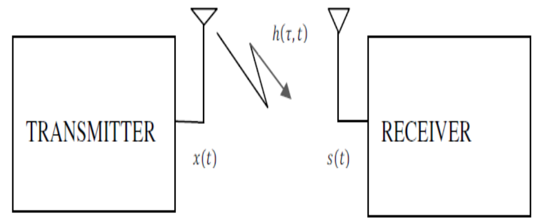

Fig.1: Block diagram of SISO system

The input-output relationship of the SISO system is given by

$$s(t) = \int_0^{\tau total} h(\tau,t)x(t-\tau)d\tau = h(\tau,t) * x(t) \quad (7)$$

Where, x(t) is the transmitted signal, s(t) is the received signal at time$t, \tau total$ is the duration of the impulse response, and denotes the convolution operation

Single Input Multiple Output System

The schematic block diagram of a SIMO system with single transmit antenna and receive antennas is depicted in Figure 2.

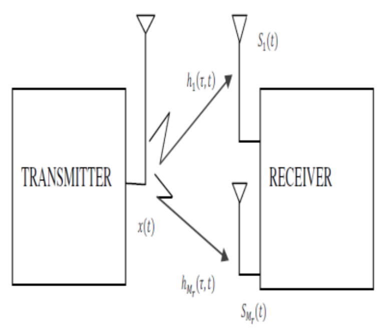

Fig.2: Block diagram of a SIMO system

The received signals at the respective receive antennas are given by

$$S_1(t) = h_1(\tau,t) * x(t) \quad (8)$$

$$S_2(t) = h_2(\tau,t) * x(t) \quad (9)$$

$$S_{Mr}(t) = h_{Mr}(\tau,t) * x(t) \quad (10)$$

Therefore, the input-output relationship of a SIMO system can be expressed as

$$S(t) = H(\tau, t) * x(t) \qquad (11)$$

Where

$S(t) = [S_1(t) S_2(t) \dots \dots \dots . S_{Mr}(t)]^T$ in the received signal vector

$h(t) = [h_1(t) h_2(t) \dots \dots \dots . h_{Mr}(t)]^T$ in the channel vector

Multipath Environment

In wireless environment, transmitted signal follow several propagation paths. Many of these paths, having reflected from surrounding objects, reach the receiver with different propagation delays. This multipath leads to delay spread, inter symbol interference (ISI), fading and random phase distortion [9]. The corresponding channel impulse response is shown in Figure 3.

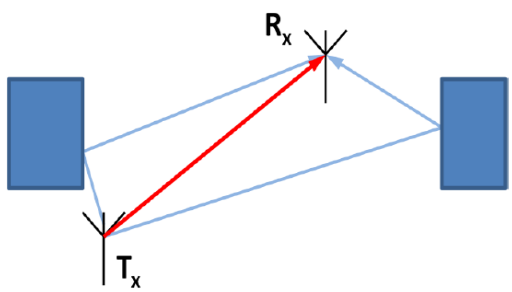

Fig.3: Multipath Environment in Wireless Communications

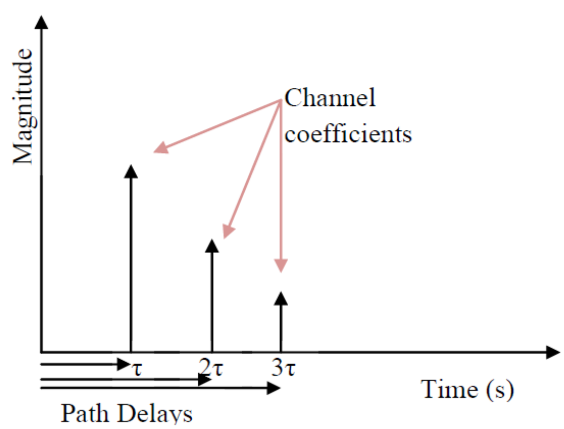

Fig.4: Channel Impulse response in Multipath Environment

Delayed copies of the transmitted signal interfere with subsequent signals, causing Inter Symbol Interference (ISI). Therefore transmitted symbol rate is limited by the delay spread of the channel. Multipath Propagation causes MIMO channels to be frequency selective which cannot have flat frequency response. To combat this effect MIMO is combined with OFDM. OFDM transforms the frequency selective fading channel into parallel flat fading sub channels, but the length of CP inserted should be greater than the channel length.

For a 2 x 2 MIMO-OFDM configuration the received OFDM symbols are given the equations 2. Here the term Xi represents the transmitted symbol from the ith transmitting antenna, the term Yj represents the received symbol from the jth receiving antenna, the term Ni represents the noise component present in the ith symbol and the term Hij represents the channel coefficient corresponding to the ith transmitting antenna and jth receiving antenna.

$$R_x.Ant\ 1{:}\ Y_1(K) = H_{11}(K)X_1(K) + H_{12}(K)X_2(K) + N_1(K) \qquad (12)$$

$$R_x.Ant\ 2{:}\ Y_2(K) = H_{21}(K)X_1(K) + H_{22}(K)X_2(K) + N_2(K) \qquad (12)$$

$$\begin{bmatrix} Y_1(K) \\ Y_2(K) \end{bmatrix} = \begin{bmatrix} H_{11}(K) & H_{12}(K) \\ H_{21}(K) & H_{22}(K) \end{bmatrix} \begin{bmatrix} X_1(K) \\ X_2(K) \end{bmatrix} + \begin{bmatrix} N_1(K) \\ N_2(K) \end{bmatrix} \qquad (13)$$

In a Direct path environment, where the reflected waves from multiple objects are absent, in channel looks to be flat and contains a single co-efficient in its impulse response. This type of channel is can be modeled using Additive White Gaussian Noise (AWGN) channel [10]. The model does not account for the phenomena of fading, interference, frequency selectivity, and nonlinearity or dispersion. In a multipath environment, the channel is always frequency selective type. This type of channel can be modeled using Rayleigh random distribution. Often the space time block coding or STBC model is used for MIMO encoding [11]. The STBC stands for space time block coding. The space–time encoding mapping of Alamouti's two-two-branch transmits diversity technique can be represented by the coding matrix:

Coding matrix:

$$X_1 = \begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix} \qquad (13)$$

In the coding matrix X1, the subscript index gives the transmit rate compared to a SISO system. For Alamouti's scheme, the transmission rate is 1. The rows of the coding matrix represent the transmit antennas while its columns correspond to different time instances [12]. It is clear that the encoding is done in both the space and time domains. Let us denote the transmit sequence from antennas one and two by x^1 and x^2 , respectively.

$$x^{t1} = [\, x_1, -x*_2] \qquad (15)$$

$$x^{t2} = [\, x_2, x*_1] \qquad (16)$$

The key feature of the Alamouti scheme is that the transmit sequences from the two transmit antennas are orthogonal, since the inner product of the sequences x^1 and x^2 is zero, i.e.

$$x^{t1} . x^{t2} = x_1 x*_2 - x*_2 x_1 \quad (17)$$

The code matrix has the following property:

$$X . X^H = \begin{bmatrix} |x_1|^2 + |x_2|^2 & 0 \\ 0 & |x_1|^2 + |x_2|^2 \end{bmatrix} = (|x_1|^2 + |x_2|^2) I_2 \quad (18)$$

here I_2 is a $2 X 2$ identity matrix

At the receive antenna, the received signals over two consecutive symbol periods, denoted by r_1 And r_2 for time t and $t + T$, respectively, can be expressed as

$$r_1 = h_1 x_1 + h_2 x_2 + n_1 \quad (19)$$

$$r_2 = -h_1 x*_1 + h_2 x*_2 + n_2 \quad (20)$$

Where n_1 and n_2 are independent complex variables with zero mean and power spectral density $N0/2$ per dimension, representing additive white Gaussian noise samples at time t and $t + T$, respectively.

III. MACHINE LEARNING ASSISTED MIMO SYSTEMS

Due to the largeness and complexity of the MIMO data, off late machine learning based techniques are being used for signal detection. Machine learning can be crudely understood as the design of automated computational systems which mimic the human behaviour and can be trained in the sense that they can learn from data fed to the system. Primarily machine learning is categorized into three major categories which are [13]-[15]:

1) Unsupervised Learning: In this approach, the data set is not labelled or categorized prior to training a model. This typically is the most crude form of training wherein the least amount of apriori information is available regarding the data sets.
2) Supervised Learning: In this approach, the data is labelled or categorized or clustered prior to the training process. This is typically possible in case the apriori information us available regarding the data set under consideration.
3) Semi-Supervised Learning: This approach is a combination of the above mentioned supervised and unsupervised approaches. The data is demarcated in two categories. In one category, some amount of the data is labelled or categorized. This is generally not the larger chunk of the data. In the other category, a larger chunk of data is un-labeled and hence the data is a mixture of both labelled and un-labeled data groups.

Some other allied categories of machine learning are:

4) Reinforcement Learning
5) Transfer Learning
6) Adversarial Learning
7) Self-Supervised learning etc.

While these learning algorithms can be studied separately, however they are essentially the modified versions of unsupervised, supervised and semi-supervised learning architectures. A more advanced and useful category of machine learning is deep learning which is the design of deep neural nets with multiple hidden layers.

Machine learning based classifiers are typically much more accurate and faster compared to the conventional classifiers. They render more robustness to the system as they are adaptive and can change their characteristics based on the updates in the dataset [16]. The common classifiers which have been used for the classification of pests are:

Regression Models: In this approach, the relationship between the independent and dependent variable is found utilizing the values of the independent and dependent variables. The most common type of regression model can be thought of as the linear regression model which is mathematically expressed as [17]:

$$y = \theta_1 + \theta_2 x \quad (21)$$

Here,
x represents the state vector of input variables
y represents the state vector of output variable or variables.
ϴ1 and ϴ2 are the co-efficients which try to fit the regression learning models output vector to the input vector.

Often when the data vector has large number of features with complex dependencies, linear regression models fail to fit the input and output mapping. In such cases, non-linear regression models, often termed as polynomial regression is used. Mathematically, a non-linear or higher order polynomial regression models is described as:

$$y = \theta_0 + \theta_1 x^3 + \theta_2 x^2 + \theta_3 x \qquad (22)$$

Here,
x is the independent variable
y is the dependent variable
$\theta_1, \theta_2 \ldots .. \theta_n$ are the co-efficients of the regression model.

Typically, as the number of features keep increasing, higher order regression models tend to fit the inputs and targets better. A typical example is depicted in figure 2

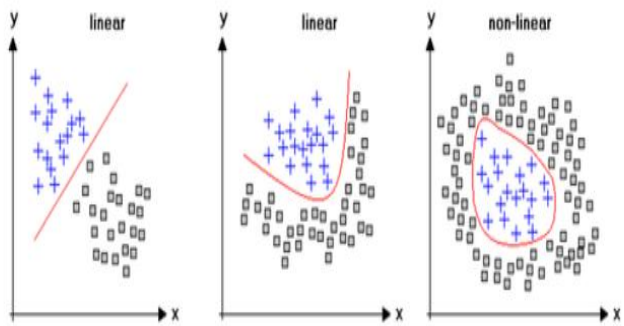

Fig.5: Linear and Non-Linear Regression fitting [14]

Support Vector Machine (SVM): This technique works on the principle of the hyper-plane which tries to separate the data in terms of ‘n’ dimensions where the order of the hyperplane is (n-1). Mathematically, if the data points or the data vector **‘X’** is m dimensional and there is a possibility to split the data into categories based on **‘n’** features, then a hyperplane of the order **‘n-1’** is employed as the separating plane. The name plane is a misnomer since planes corresponds to 2 dimensions only but in this case the hyper-plane can be of higher dimensions and is not necessarily a 2-dimensional plane. A typical illustration of the hyperplane used for SVM based classification is depicted in figure 3.

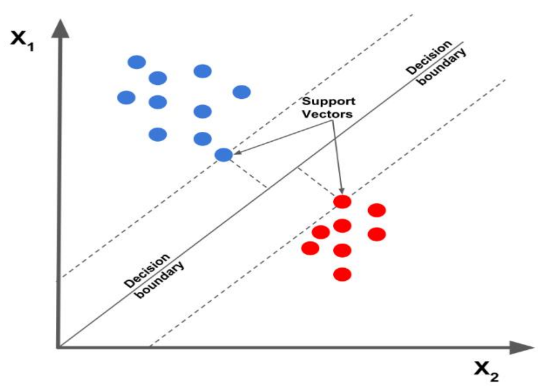

Fig.6: Separation of data classes using SVM [17]

The selection of the hyperplane H is done on the basis of the maximum value or separation in the Euclidean distance d given by:

$$d = \sqrt{x_1^2 + \cdots \ldots \ldots . x_n^2} \qquad (23)$$

Here,
x represents the separation of a sample space variables or features of the data vector,
n is the total number of such variables
d is the Euclidean distance

The (n-1) dimensional hyperplane classifies the data into categories based on the maximum separation. For a classification into one of **‘m’** categories, the hyperplane lies at the maximum separation of the data vector **‘X’.** The categorization of a new sample **‘z’** is done based on the inequality:

$$d_x^z = Min(d_{C1}^z, d_{C2}^z \ldots d_{C2=m}^z) \qquad (24)$$

Here,
d_x^z is the minimum separation of a new data sample from **‘m’** separate categories
$d_{C1}^z, d_{C2}^z \ldots d_{C2=m}^z$ are the Euclidean distances of the new data sample **‘z’** from m separate data categories.

Neural Networks: Owing to the need of non-linearity in the separation of data classes, one of the most powerful classifiers which have become popular is the artificial neural network (ANN). The neural networks are capable to implement non-linear classification along with steep learning rates. The neural network tries to emulate the human brain’s functioning based on the fact that it can process parallel data streams and can learn and adapt as the data changes. This is done through the updates in the weights and activation functions. The mathematical model of the neural network is depicted in figure 7.

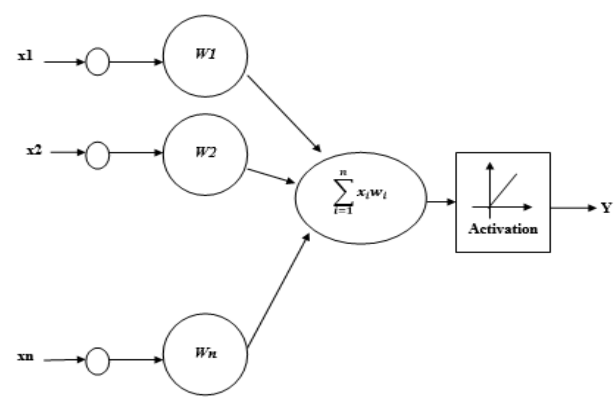

Fig.7: Mathematical Model of Single Neuron [13]

The mathematical equivalent of an artificial neuron is depicted in figure 4 where the output can be given by:

$$y = f(\sum_{i=1}^{n} x_i w_i + b) \qquad (25)$$

Here,
x denote the parallel inputs
y represents the output
w represents the bias
f represents the activation function

The neural network is a connection of such artificial neurons which are connected or stacked with each other as layers. The neural networks can be used for both regression and classification problems based on the type of data that is fed to them. Typically the neural networks have 3 major conceptual layers which are the input layer, hidden layer and output layer. The parallel inputs are fed to the input layer whose output is fed to the hidden layer. The hidden layer is responsible for analysing the data, and the output of the hidden layer goes to the output layer. The number of hidden layers depends on the nature of the dataset and problem under consideration. If the neural network has multiple hidden layers, then such a neural network is termed as a deep neural network. The training algorithm for such a deep neural network is often termed as deep learning which is a subset of machine learning. Typically, the multiple hidden layers are responsible for computation of different levels of features of the data. Several categories of neural networks such as convolutional neural networks (CNNs), Recurrent Neural Network (RNNs) etc.

IV. RELATED WORK

The section presents the related work done in the domain, with the approach used and the findings. The related work is summarized in table 1.

Table 1. Comparative Analysis of Existing Work

Authors	**Year and Publication**	**Approach Used**	**Findings**
Guo et al.	Elsevier, 2023	Deeplearning based joint channel estimation and feedback framework for massive MIMO systems.	Normalized MSE (NMSE) of 10^{-3} achieved for SNR of 30dB
Arvinte et al.	IEEE 2023	A Generative Adversarial Network (GAN) based Deep Learning model for MIMO channel estimation	Normalized MSE (NMSE) of -40dB achieved for SNR of 38dB
Liu et al.	IEEE 2022	A skip-connection attention (SC-attention) based deep learning model for Channel Estimation of MIMO systems.	Normalized MSE (NMSE) of 10^{-2}achieved for SNR of 15dB
Nguyen et al.	IEEE, 2021	Use of Deep Learning to Sphere Decoding for Large MIMO Systems	Bit Error Rate (BER) of 10^{-2}achieved for an SNR value of 16dB. Exponential computational complexity achieved for increase in BER.
Sohrabi et al.	IEEE 2021	Application of Deep Learning for Distributed Channel Feedback and Multiuser Precoding.	Sum rate of up to 16bps/Hz achieved for 8 multiple MIMO paths for FDD massive MIMO system.
Lin et al.	IEEE 2021	Estimation of channel state information (CSI) and antenna selection based on deep learning.	NMSE of 10^{-2} achieved for an SNR value of 25Db. Achievable sum rate of 9bits/s/Hz at SNR of 25Db.
Mashaddi et al.	IEEE 2021	Deep Learning-Based Pilot Design and Channel Estimation for MIMO-OFDM Systems	Linear minimum mean square error (LMMSE) estimation applied at decoder with an NMSE of -24 achieved. System outperforms NN+FFT and NN+Attention+FFT in terms of NMSE.
Balevi et al.	IEEE, 2020	Channel Estimation for MIMO systems with previously untrained deep neural network.	Normalized MSE (NMSE) of 10^{-5} achieved.

Baek et al.	IEEE, 2019	Deep Learning Based Approach for detection of conventional MIMO systems.	Bit Error Rate (BER) of 10^{-5} achieved for an SNR value of 16Db.
He et al.	IEEE 2020	Joint MIMO channel estimation and signal detection (JCESD) employed using Deep Learning.	Bit Error Rate (BER) of 10^{-4} achieved for an SNR value of 35Db
Wang et al.	IEEE 2020	Cooperative Modulation classification for MIMO systems using Deep Learning.	Automatic Modulation Classification (AMC) receiver designed to assist handovers. Classification accuracy of above 90% achieved for SNR of 10Db.
Yuan et al.	IEEE 2019	MIMO channel estimation based on machine learning, NARX and CNN models used.	Least NMSE of -20Db achieved among the machine learning models used.
Gecgel et al.	IEEE 2019	A machine learning based approach for transmitter selection for MIMO-GSM systems.	Decision tree and multilayer perceptron algorithms are adopted as transmit antenna selection approaches with a BER of 10^{-4} achieved for an SNR of 25dB
Mashaddi et al.	Journal of Indian Institute of Science Springer.	A deep learning approach for massive MIMO channel state acquisition and feedback.	The work tries to reduce training overhead for CSI acquisition andfeedback overhead, NMSE of -20 achieved for CSI overhead of 80%.

V. RESEARCH GAP

The research gap identified in the existing work can be summarized as:

Back propagation based ML or Deep Learning algorithms have found limited application with the majority of the approaches applying feed forward deep learning models. Dimensional reduction tools such as principal component analysis have not been used extensively. This may lead to overfitting of the data keeping in mind the enormity of the data that is transmitted in MIMO systems. Scanty moving regression and moving average approaches have been adopted. Applying extensive moving approach may help in finding the recent trends in the channel state information (CSI) and result in computation of much more accurate channel statistics enabling lesser BER of the system.

Moreover, a relatively high normalized error rate of 10^{-3}achieved for relatively high SNR values of around 30dB. A steeper reduction of error rate should be aimed for

VI. CONCLUSION

It can be concluded that MIMO systems and moreover massive MIMO systems are the only way out to

provide limited bandwidth yet bandwidth hungry applications. Massive MIMO may serve as a technique to increase the system capacity and date rate. However, MIMO systems suffer from distortion effects. This is due to the addition of noise and interference, as well as the different behavior of the wireless channel for the different channel paths. Due to the largeness and the complexity of the data, the technique used off late is the Deep Learning based approach for signal detection. This paper presents a comprehensive review of machine learning and deep learning assisted massive MIMO systems along with the salient features of the approaches. It is expected that the paper will serve as an effective tool to find future directions of research.

REFERENCES

[1] M. A. Jensen and J. W. Wallace, "A review of antennas and propagation for MIMO wireless communications," in IEEE Transactions on Antennas and Propagation, vol. 52, no. 11, pp. 2810-2824, Nov. 2004, doi: 10.1109/TAP.2004.835272.

[2] N. Kim, Y. Lee and H. Park, "Performance Analysis of MIMO System with Linear MMSE Receiver," in IEEE Transactions on Wireless Communications, vol. 7, no. 11, pp. 4474-4478, November 2008, doi: 10.1109/T-WC.2008.070785.

[3] A. H. Mehana and A. Nosratinia, "Diversity of MMSE MIMO Receivers," in IEEE Transactions on Information Theory, vol. 58, no. 11, pp. 6788-6805, Nov. 2012, doi: 10.1109/TIT.2012.2204954.

[4] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," in IEEE Transactions on Signal Processing, vol. 54, no. 3, pp. 884-893, March 2006, doi: 10.1109/TSP.2005.863008.

[5] A. Ghazal et al., "A Non-Stationary MIMO Channel Model for High-Speed Train Communication Systems," 2012 IEEE 75th Vehicular Technology Conference (VTC Spring), 2012, pp. 1-5, doi: 10.1109/VETECS.2012.6240333.

[6] H. Shin and M. Z. Win, "MIMO Diversity in the Presence of Double Scattering," in IEEE Transactions on Information Theory, vol. 54, no. 7, pp. 2976-2996, July 2008, doi: 10.1109/TIT.2008.924672.

[7] M. Damrath, H. B. Yilmaz, C. Chae and P. A. Hoeher, "Array Gain Analysis in Molecular MIMO Communications," in IEEE Access, vol. 6, pp. 61091-61102, 2018, doi: 10.1109/ACCESS.2018.2875925.

[8] Q. He, Z. Wang, J. Hu and R. S. Blum, "Performance Gains From Cooperative MIMO Radar and MIMO Communication Systems," in IEEE Signal Processing Letters, vol. 26, no. 1, pp. 194-198, Jan. 2019, doi: 10.1109/LSP.2018.2880836.

[9] F. Rusek and A. Prlja, "Optimal Channel Shortening for MIMO and ISI Channels," in IEEE Transactions on Wireless Communications, vol. 11, no. 2, pp. 810-818, February 2012, doi: 10.1109/TWC.2011.121911.110809.

[10] A. Abdelaziz and C. E. Koksal, "Fundamental limits of covert communication over MIMO AWGN channel," 2017 IEEE Conference on Communications and Network Security (CNS), 2017, pp. 1-9, doi: 10.1109/CNS.2017.8228657.

[11] N. Varshney, A. V. Krishna and A. K. Jagannatham, "Selective DF Protocol for MIMO STBC Based Single/Multiple Relay Cooperative Communication: End-to-End Performance and Optimal Power Allocation," in IEEE Transactions on Communications, vol. 63, no. 7, pp. 2458-2474, July 2015, doi: 10.1109/TCOMM.2015.2436912.

[12] S. Yan, N. Yang, R. Malaney and J. Yuan, "Transmit Antenna Selection with Alamouti Coding and Power Allocation in MIMO Wiretap Channels," in IEEE Transactions on Wireless Communications, vol. 13, no. 3, pp. 1656-1667, March 2014, doi: 10.1109/TWC.2013.013014.131248.

[13] H. Wang, C. Ma and L. Zhou, "A Brief Review of Machine Learning and Its Application," 2009 International Conference on Information Engineering and Computer Science, 2009, pp. 1-4, doi: 10.1109/ICIECS.2009.5362936.

[14] X. Gao, L. Dai, Y. Sun, S. Han and I. Chih-Lin, "Machine learning inspired energy-efficient hybrid precoding for mmWave massive MIMO systems," 2017 IEEE International Conference on Communications (ICC), 2017, pp. 1-6, doi: 10.1109/ICC.2017.7997065.

[15] J. Huang et al., "A Big Data Enabled Channel Model for 5G Wireless Communication Systems," in IEEE Transactions on Big Data, vol. 6, no. 2, pp. 211-222, 1 June 2020, doi: 10.1109/TBDATA.2018.2884489.

[16] W. Kim, J. Park, J. Yoo, H. J. Kim and C. G. Park, "Target Localization Using Ensemble Support Vector Regression in Wireless Sensor Networks," in IEEE Transactions on Cybernetics, vol. 43, no. 4, pp. 1189-1198, Aug. 2013, doi: 10.1109/TSMCB.2012.2226151.

[17] Martin Hagan, “ANN Design” 2nd Edition, Cengage Publications.

[18] E. Balevi, A. Doshi and J. G. Andrews, "Massive MIMO Channel Estimation With an Untrained Deep Neural Network," in IEEE Transactions on Wireless Communications, vol. 19, no. 3, pp. 2079-2090, March 2020, doi: 10.1109/TWC.2019.2962474.

[19] M. -S. Baek, S. Kwak, J. -Y. Jung, H. M. Kim and D. -J. Choi, "Implementation Methodologies of Deep Learning-Based Signal Detection for Conventional MIMO Transmitters," in IEEE Transactions on Broadcasting, vol. 65, no. 3, pp. 636-642, Sept. 2019, doi: 10.1109/TBC.2019.2891051.

[20] N. T. Nguyen, K. Lee and H. DaiIEEE, "Application of Deep Learning to Sphere Decoding for Large MIMO Systems," in IEEE Transactions on Wireless Communications, vol. 20, no. 10, pp. 6787-6803, Oct. 2021, doi: 10.1109/TWC.2021.3076527.

[21] M. B. Mashhadi and D. Gündüz, "Pruning the Pilots: Deep Learning-Based Pilot Design and Channel Estimation for MIMO-OFDM Systems," in IEEE Transactions on Wireless Communications, vol. 20, no. 10, pp. 6315-6328, Oct. 2021, doi: 10.1109/TWC.2021.3073309.

[22] H. He, C. Wen, S. Jin and G. Y. Li, "Model-Driven Deep Learning for MIMO Detection," in IEEE Transactions on Signal Processing, vol. 68, pp. 1702-1715, 2020, doi: 10.1109/TSP.2020.2976585.

[23] Y. Wang, J. Wang, W. Zhang, J. Yang and G. Gui, "Deep Learning-Based Cooperative Automatic Modulation Classification Method for MIMO Systems," in IEEE Transactions on Vehicular Technology, vol. 69, no. 4, pp. 4575-4579, April 2020, doi: 10.1109/TVT.2020.2976942.

[24] N. T. Nguyen, K. Lee and H. DaiIEEE, "Application of Deep Learning to Sphere Decoding for Large MIMO Systems," in IEEE Transactions on Wireless Communications, vol. 20, no. 10, pp. 6787-6803, Oct. 2021, doi: 10.1109/TWC.2021.3076527.

[25] B. Lin, F. Gao, S. Zhang, T. Zhou and A. Alkhateeb, "Deep Learning-Based Antenna Selection and CSI

Extrapolation in Massive MIMO Systems," in IEEE Transactions on Wireless Communications, vol. 20, no. 11, pp. 7669-7681, Nov. 2021, doi: 10.1109/TWC.2021.3087318.

[26] J. Yuan, H. Q. Ngo and M. Matthaiou, "Machine Learning-Based Channel Estimation in Massive MIMO with Channel Aging," 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2019, pp. 1-5, doi: 10.1109/SPAWC.2019.8815557.

[27] S. Gecgel, C. Goztepe and G. Karabulut Kurt, "Transmit Antenna Selection for Large-Scale MIMO GSM With Machine Learning," in IEEE Wireless Communications Letters, vol. 9, no. 1, pp. 113-116, Jan. 2020, doi: 10.1109/LWC.2019.2944179.

[28] MB Mashhadi, D Gündüz, Deep learning for massive MIMO channel state acquisition and feedback, Journal of Indian Institute of Science, Springer, 2020, vol.100, pp. 369-382.

[29] J. Guo, T. Chen, S. Jin, G.Y. Li, X. Wang, X. Hou, “Deep learning for jointchannel estimation and feedback in massive MIMO systems”, Digital Communications and Networks, Elsevier 2023.

[30] M. Arvinte and J. I. Tamir, "MIMO Channel Estimation Using Score-Based Generative Models," in IEEE Transactions on Wireless Communications, 2023 vol. 22, no. 6, pp. 3698-3713.

[31] M. Liu, X. Li, B. Ning, C. Huang, S. Sun and C. Yuen, "Deep Learning-Based Channel Estimation for Double-RIS Aided Massive MIMO System," in IEEE Wireless Communications Letters, 2022, vol. 12, no. 1, pp. 70-74.