

Deep Learning Approaches for Impulse Noise Mitigation and Classification in NOMA-Based Systems

MUHAMMAD HUSSAIN^{ID}, HINA SHAKIR^{ID}, AND HAROON RASHEED

Department of Electrical Engineering, Bahria University, Karachi 75260, Pakistan

Corresponding author: Hina Shakir (hinashakir.buke@bahria.edu.pk)

ABSTRACT The new emerging networks such as smart grids, smart homes and Internet of Things have enabled user accessibility across the globe and employ non-orthogonal multiple access (NOMA) scheme to accommodate huge number of connected devices. These devices which include smart meters, sensors and actuators etc. suffer from impulse noise (IN) while operating with power systems. Furthermore, NOMA scheme provides power domain multiple access (PDMA) which is found to be susceptible to IN. Based on the aforementioned IN intervention and its degrading effect on communication applications, novel mechanisms are desired to mitigate and classify the IN induced in the received signal. In this research work, novel IN mitigation and classification techniques are presented using deep learning methods for NOMA-based communication systems. The IN detection is performed by first identifying the IN occurrences using a deep neural network (DNN) which learns statistical traits of noisy samples followed by removal of harmful effect of IN in the detected occurrences. Using the proposed DNN, higher bit error rates (BER) were achieved when compared with the existing IN detection methods. The proposed method was further validated for high and low IN, and weak and strong IN occurrence probabilities. Moreover, another deep learning network is proposed in this research work to effectively distinguish between high IN and low IN in the noise contaminated NOMA symbols which can help improve the performance of IN detection models. Both of the deep learning methods proposed in this study show strong potential to address IN problem faced by the NOMA scheme.

INDEX TERMS Deep neural network (DNN), impulse noise (IN), multiuser communication, non-orthogonal multiple access (NOMA), power division multiple access (PDMA), successive interference cancellation (SIC).

I. INTRODUCTION

Non-orthogonal multiple access (NOMA) scheme is a new emerging technology for 5G which can fulfill the increasing demand of bandwidth in next generation networks. NOMA exhibits high spectral efficiency due to full bandwidth allocation to each user at the same time and frequency while differentiating users based on a power allocation strategy. In this scheme, the nearer user (lower-order user) receives less power than the farther user (higher-order user) from the Base Station (BS). At the transmitter side, BS superimposes signals of all users at the same time and frequency whereas at the receiver side, the nearer user handles signals of other users as interference thereby performing successive interference

cancellation (SIC) to extract its own signal. Following this strategy, the farthest user decodes its signal without SIC and considers signals of other users as noise because its own signal contributes most part of the received signal [1].

The research conducted in the recent past shows increased usage of NOMA with background additive white Gaussian noise (AWGN) [2]. The authors further investigate the deteriorating effect of impulse noise in wireless and powerline communications in smart grids, health care monitoring applications, intra-vehicle communications, industrial and mining environments. The sources of such noise were found to be automotive ignition, power lines, industrial equipment, household appliances, electronic devices, medical equipment etc. In the literature, impulse noise (IN) is defined as noise that occurs for short instances randomly in time but carries huge amount of power [3]. Despite a few research

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Gaggero^{ID}.

studies [4]–[6] which have discussed the strong impact of IN on NOMA-based communication systems, most researchers have included AWGN as a noise benchmark. As NOMA uses power domain multiple access, it is crucial to detect the occurrence of IN since it can change the power level of signal which in turn degrades NOMA performance. The classical IN detection approaches perform blanking and clipping by setting an optimum threshold which is susceptible to vary in response to channel conditions and therefore causes model mismatch. Any change in the detection threshold of impulses in traditional methods can cause ambiguity in the receiver threshold of signal detection, thus deteriorating the performance of NOMA. The unique traits of NOMA scheme which are non-orthogonal resource allocation and subsequent interference cancellation, have additionally increased the user's response to such IN interference [5]. The aforesaid issues demand more research work to be carried out to evaluate the impact and suppression strategies of such noise on NOMA-based systems. Moreover, investigation on efficient IN classification methods can significantly help develop better IN mitigation strategies.

In this research paper, two deep learning strategies have been proposed; one for IN detection and the other for high/low IN classification respectively to overcome the above discussed challenges in NOMA-based communication systems. The proposed IN mitigation approach first detects and then suppresses IN. For this purpose, the data samples with induced IN are identified using a deep learning network (DNN) and then such instances are decoded to remove the noise. The developed DNN scheme works independently of the traditional noise mitigation methods such as blanking, clipping etc. This ensures greater success in IN detection as compared to the traditional methods. The performance of the proposed DNN-based method is measured in terms of Bit error rate (BER) and compared in a NOMA-based system with the conventional approaches presented in the literature. Furthermore, the performance of DNN for IN mitigation is also measured for high and low IN, and weak and strong impulse occurrence probabilities considering a NOMA user pair scenario.

Although the occurrence of impulse is random in nature, high impulse occurrence is generally low but occasionally it becomes high. These two different cases can affect the proposed deep learning mitigation strategy for IN. Therefore, IN classification is proposed because it gives useful information regarding nature of impulse to improve the DNN performance for IN mitigation. The second proposed DNN for IN categorization takes contaminated data samples along with other relevant parameters which are described in the following section and classifies the IN samples as high IN or low IN.

The structure for both proposed deep networks comprises of input, output and hidden layers. These layers are fully connected with each other in a manner that every input combination produces the correct possible output of the two available outputs. The hidden layers work with non-linear activation

function to identify non-linear data. The input feature vectors fed to the two DNNs are chosen based on their characteristics to identify and classify the IN affected samples. For IN detection, the input features/parameters include current sample value, absolute median difference of a samples block and a statistic parameter well-known for data classification [7]. Conversely, the input parameters/features to DNN for IN classification include running sample, absolute median difference of a sample block and average sample occurrence probability with respect to its neighbors. The results achieved from the proposed deep learning approaches show that DNNs have successfully mitigated and classified the IN with superior performance over other existing methods.

The remainder of this paper is organized as follows. Section II discusses the effect of IN in NOMA-based system, available mitigation techniques in the literature and research contribution. The system model is described in Section III and existing IN suppression strategies are discussed in Section IV. Proposed DNNs for IN mitigation and classification are presented in Section V and VI respectively. Performance analyses of presented models are carried out in Section VII and conclusions are drawn in Section VIII.

II. BACKGROUND

A. EFFECT OF IMPULSE NOISE IN NOMA-BASED SYSTEMS

Extensive literature on NOMA, mostly in the context of next generation networks like smart grid or Internet of Things (IoTs), has presented several challenges in NOMA system implementation. The effect of IN on NOMA uplink [4] and downlink [6] system have also been analyzed. Research work in [4] has presented the outage performance of NOMA uplink systems in presence of IN. The vulnerability of NOMA system due to IN was validated by thorough Monte-Carlo simulations along with analytical results. In [6], authors presented the impact of IN on sum rate capacity of NOMA downlink systems. A particle IN scenario was used to quantify the actual loss due to IN. The authors further examined the performance of NOMA, in the presence of composite noise (Impulse with AWGN) over Rayleigh fading channel. Pairwise loss of bits (LoB) expression was derived and further applied to develop a union bound on the bit error rate (BER). The research work presented quantification of the variation in channel conditions in the presence of composite noise as experienced by NOMA users. In [5], authors investigated the performance degradation of NOMA-based IoT networks due to IN and proposed mitigation technique. A deep learning based multistage nonlinear solution was proposed to estimate the IN parameter for received orthogonal frequency division multiplexing (OFDM) symbols originating from power domain multiple-NOMA (PDM-NOMA) scheme.

B. IN MITIGATION TECHNIQUES: MEMORYLESS NONLINEAR MITIGATION APPROACHES

The threshold-based IN scheme is described as a memoryless nonlinear mitigation approach and includes methods

TABLE 1. Comparison of proposed DNN with state-of-the-art methods.

Ref	Objective	Solution Approach	Technique	Best Performance
[8]	Performance improvement of blanking non-linearity element for OFDMA signals in impulsive noise channels	An iterative interference cancellation scheme with adaptive blanking threshold for each iteration	Iterative Blanking	Best performance with only three iterations over fixed blanking threshold
[9]	Selection clipping threshold with a priori knowledge approximate	An approximate knowledge of the arrival probability of impulses to render the clipping threshold, for OFDM transmissions	Probabilistic Clipping	Outperformed its conventional when coupled with a turbo decoder using adapted log-likelihood ratio metric
[10]	To boost signal-to-noise power ratio (SNR) at the OFDMA multicarrier receiver	A nonlinear estimators based on the multiple thresholding, with associated piecewise attenuation, or clipping, of the received signal amplitude.	Multiple-Threshold Blanking/Clipping	Estimators closely approached the performance of the optimal minimum mean square error Bayesian estimator.
[11]	Calculation of a blanking and a clipping threshold	Presented simple methods that obtain quasi optimal thresholds of a blanker and a soft limiter	Blanking/Clipping	A blanker and the soft limiter are shown suitable in strong and moderate impulsive environments
[12]	IN removal from OFDM-based PLC systems	Proposed a combined Time-domain /Frequency-domain technique for IN removal in OFDM-based PLC systems	Blanking, Clipping and Blanking/Clipping	Combined TD/FD technique performs better than practically used non-linearities to reduce the adverse effect of impulsive noise significantly
[13]	Interference reduction in high density network of independent small base stations	Presented a distributed power allocation algorithm based on multi-agent Q-learning	Machine Learning	Optimal joint power allocation for small base stations network.
[7]	IN removal from received symbols in OFDM-based Systems	A two-step technique in which a deep neural network (DNN) identifies the IN samples. The detected IN sample is suppressed in the second step.	Deep Neural Network	Performs better than the traditional IN mitigation approaches which include blanking and clipping
[5]	To obtain the optimum threshold of the corresponding NOMA user	A deep learning based multistage nonlinear processing approach to estimate the impulsive noise parameters (threshold) from the received OFDM symbol for OFDM-based PDM-NOMA systems.	Deep Neural Network	Optimization to simultaneously maximize the performance of all the users.
Proposed DNN	IN mitigation in NOMA-based system	A deep learning approach which detects IN based on learning about randomness of noise using statistical properties of input samples	Deep Neural Network	Outperformed the existing IN mitigation techniques. Also classified IN as high IN and low IN which was not attempted in the literature.

like blanking [8], clipping [9], and blanking/clipping [10]. In this mitigation approach, the short duration and the high amplitude of IN are investigated by using a threshold whose selection remains a challenging task. A threshold optimization technique is presented by the authors in [14] based on Neyman-Pearson criterion. In [11], the researchers provided an analytical equation for IN mitigation using blanking and clipping. A comparison of several analog domain processing techniques IN alleviation is presented in [15] which showed that selection of threshold value is the key parameter to improve performance of threshold-based nonlinear solutions. The model becomes mismatched when the threshold changes in response to channel conditions. As an implication, the performance of all the conventional threshold-based methods suffer from harsh impulsive environment.

C. IN MITIGATION TECHNIQUES: MACHINE LEARNING-BASED MITIGATION APPROACHES

Deep learning is a popular machine learning method which has been applied [13], [16], [17] for IN mitigation in applications pertaining to signals and images, and power/bandwidth

allocation in communication networks respectively. Due to its inherent ability to estimate the noise patterns and uncertainty in samples, deep learning network can be used for IN detection but requires appropriate processing strategies and network structures.

In an Orthogonal Frequency-division Multiple Access (OFDMA) system, the performance of conventional mitigation approaches can be degraded due to high Peak to Average Power Ratio (PAPR) of OFDM signals. Therefore, there is a trade-off between a false positive and true detection occurrence in the conventional threshold-based approaches. To address the aforesaid drawbacks, authors in [7] proposed a deep learning based IN mitigation technique for an OFDM based communication model. The presented IN elimination strategy was implemented in two parts where IN detection was followed by IN elimination. First, a DNN was developed to identify the IN effected signal samples. Then, in the second stage, identified IN effected signal was proposed to be either clipped or blanked. In [18], DNN processed the run-time sample value using median deviations filter output as an input feature for IN detection. In [19], Rank-Ordered Absolute

Differences (ROAD) statistic was used to determine whether the run-time sample is affected by IN. A comparison of the proposed DNN with the state-of-the-art methods has been made in Table 1.

Although DNN hasn't been fully utilized for IN migration in NOMA-base systems in the past, state-of-the-art of work has been presented in the literature in resource allocation, optimization, and receiver design for NOMA-base systems. In [20], authors presented end-to-end optimization of NOMA with help of a unified deep multi-task learning framework termed as DeepNOMA by dealing with the non-orthogonal overlapped transmissions as multiple distinctive but correlated learning tasks. DeepNOMA was composed of DeepMUD and DeepMAS modules, which corresponds to multiuser detection and signature mapping, respectively. To ensure fairness among tasks and to avoid getting stuck into local optima, the authors proposed a novel multi-task balancing loss function. In [21], for NOMA joint signal detection, authors developed a novel deep learning aided end-to-end receiver which simultaneously realizes the estimation, equalization, and demodulation. The authors proposed a deep learning-aided multi-user detection (DeepMuD) in [22] to allow the massive machine type communication by multi-user detection in uplink NOMA. The proposed Deep-MuD model greatly improves the error performance of the uplink NOMA and performs better than the conventional detectors. In [23], the authors presented a uniform signal processing architecture-based deep neural network. Due to the universal function approximation property of DNN, the end-to-end optimization of NOMA transceivers is achieved by the authors. The authors optimized the transceivers to match cross-layer behaviors and extract the user access behaviors out of the time-series signals with the help of DNN. The integration of neural computation and NOMA to achieve low-cost data transmission with high efficiency was further analyzed.

D. RESEARCH CONTRIBUTION

The article [5] presented a theoretical framework on performance degradation and mitigation of IN in the context of NOMA-based IoT networks. On the contrary, in our research work, a model is developed to analyze the effect of IN in NOMA-based systems and mitigation approaches are presented by providing mathematical and simulation justifications which is a superior contribution.

The model presented in [5] determines only the blanking/clipping threshold using a deep neural network (DNN) for NOMA user pair and then uses the computed threshold (a classic mitigation technique) for IN mitigation. In comparison, our proposed DNN approach (a novel approach) has identified the IN contaminated bits in the NOMA user pair and removed the harmful effect of IN. Additionally, the performance of proposed DNN approach is evaluated in terms of low/high impulse and strong/weak likelihood of IN. An analysis of DNN performance for a NOMA user pair is also shown in which one performs SIC (user 1) and the other

one does not (user 2). Moreover, classification of IN as low and high IN using a DNN is also performed. The presented classification method aims to improve the performance of proposed mitigation approach.

III. SYSTEM MODEL

The system model for IN detection and classification in NOMA symbols using a DNN is shown in Fig. 1. The communication channel is contaminated by Rayleigh fading and IN. In the pre-processing step, conversion of NOMA noisy/noiseless symbols into sample bits is carried out which is followed by DNN training. The trained DNN then performs either the detection or the classification of IN on test NOMA symbols.

A. IN MODELS

The proposed mitigation system employs the following IN models:

1) BERNOULLI-GAUSSIAN MODEL

According to Bernoulli-Gaussian model, level of noise can be expressed as:

$$\xi = n_G + bn_I \quad (1)$$

where n_G and n_I represent AWGN with mean zero and variance σ_G^2 and σ_I^2 , respectively [24], [25]. Here for Bernoulli random sequence, b is the arrival rate of IN with probability p , independent of n_G and n_I . The noise is independent and identically distributed (i.i.d.) random variable, with probability density function (pdf) given by:

$$P_n(x) = (1 - p)G(x, 0, \sigma_G^2) + pG(x, 0, \sigma_G^2 + \sigma_I^2) \quad (2)$$

where $G(v, 0, \sigma_G^2)$ is a Gaussian pdf with mean μ_x and variance σ_x^2 . The average noise power ξ_0 is given by:

$$\xi_0 = E[n^2] \quad (3)$$

$$\xi_0 = E[n_G^2] + E[b^2]E[n_I^2] \quad (4)$$

$$\xi_0 = \sigma_G^2 + p \sigma_I^2 \quad (5)$$

2) LAPLACIAN-GAUSSIAN MODEL

In comparison to Bernoulli noise model, pdf of Laplacian noise with zero mean and variance $2c^2$ has a heavier tail and moves slower toward zero [26], [27]. Therefore, Laplacian-Gaussian model can be chosen to represent IN if large number of impulses with short amplitude occur. Otherwise, Bernoulli-Gaussian model is a better choice. The average noise power in Laplacian-Gaussian noise model can be written as:

$$\xi_0 = E[n_G^2] + E[r^2]E[n_I^2] \quad (6)$$

$$\xi_0 = \sigma_G^2 + p 2c^2 \quad (7)$$

where r is rate of arrival of IN, defined as Bernoulli random variable with pdf of Laplacian-Gaussian model.

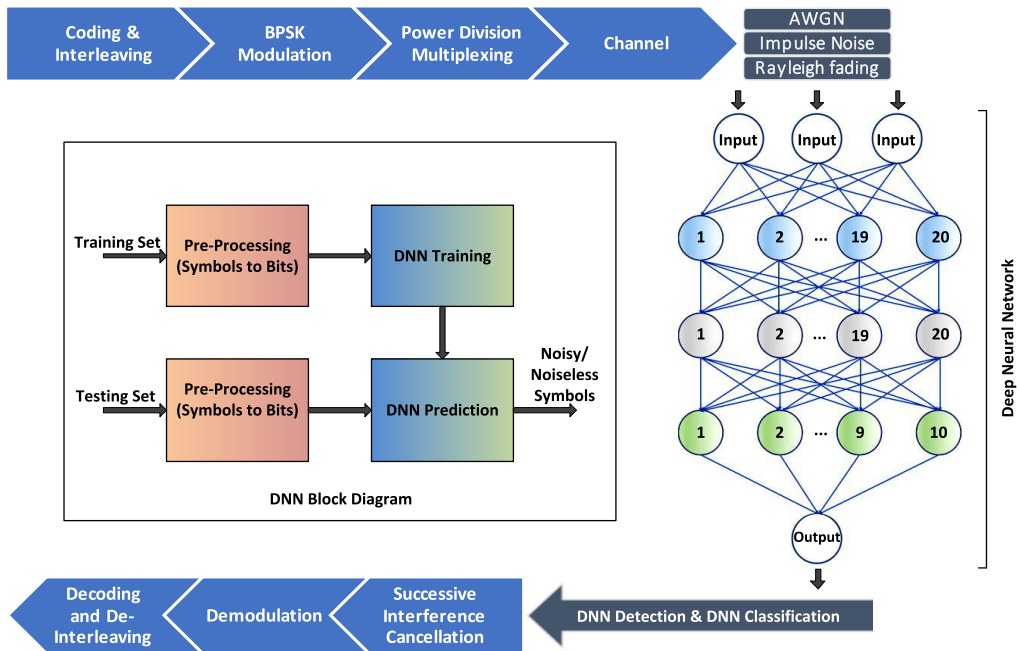


FIGURE 1. System model.

B. IN ANALYSIS IN NOMA-BASED SYSTEMS

For non-orthogonal downlink transmission, BS simultaneously transmits signals to multiple user equipments (UEs) at the same time and frequency. Each user utilizes a fraction of total power P , therefore multiple users can use the same frequency bandwidth and they are distinguished by the power level that is allocated by the BS. The decoder detects the signals of a user by successful interference cancellation (SIC) while the other users are treated as noise [28]. For two user model, received symbol for user 1 (lower-order user) in the presence of IN can be written as:

$$r_1 = s_1\sqrt{\alpha_1 P}h_1 + s_2\sqrt{\alpha_2 P}h_1 + \xi \tag{8}$$

And received symbol for user 2 (higher-order user) in the presence of IN can be written as:

$$r_2 = s_2\sqrt{\alpha_2 P}h_2 + s_1\sqrt{\alpha_1 P}h_2 + \xi \tag{9}$$

where s_1 and s_2 are the transmitted symbols, α_1 and α_2 are the power allocation coefficients and h_1 and h_2 are the channel coefficients for user 1 and user 2 respectively.

For an i_{th} user, the received symbol in presence of IN can be written as:

$$r_i = s_i\sqrt{\alpha_i P}h_i + \sum_{j=1, j \neq i}^{i-1} s_j\sqrt{\alpha_j P}h_i + \xi \tag{10}$$

where s_i and s_j are the transmitted symbols for i_{th} and j_{th} users respectively and h_i is the channel coefficient for the i_{th} user.

Considering the NOMA scheme, users' signals which are closer to BS are assumed to be interference and farther

ones to BS are treated as noise. The second term in (10) $\sum_{j=1, j \neq i}^{i-1} s_j\sqrt{\alpha_j P}h_i = \chi$ represents the level inter-cell interference symbols for UE_i . Also, $\chi_0 = \sum_{j=1, j \neq i}^{i-1} \alpha_j P|h_i|^2$ is the average power of inter-cell interference which is considered as noise. α is power allocation coefficient, $\alpha_1 + \alpha_2 + \alpha_3 + \dots + \alpha_i = 1$ and h is the fading coefficient of wireless channel. Now, the above equation becomes:

$$r_i = s_i h_i + \chi + \xi \tag{11}$$

The channel is represented as a Rayleigh fading channel. The pdf of Rayleigh fading channel with random variable $v = h$ is given by:

$$f_h(v) = \frac{v}{\sigma_h^2} \exp\left(\frac{-v^2}{2\sigma_h^2}\right) \tag{12}$$

and its m_{th} moment is:

$$E[h^m] = (2\sigma_h^2)^{m/2} \Gamma\left(\frac{m}{2} + 1\right) \tag{13}$$

We can define average received Signal-to-noise ratio (SNR) as:

$$\bar{\gamma} = E[\gamma] = \frac{\alpha_i E_b E[h^2]}{\chi_0 + \xi_0} \tag{14}$$

where E_b is the bit energy and $E[h^2]$ is 2_{nd} moment. Let γ denote the instantaneous output SNR, defined as follows:

$$\gamma = \frac{\alpha_i E_b h_i^2}{\chi_0 + \xi_0} \tag{15}$$

Through two degrees of freedom, $|h|^2$ is chi-square distributed with the condition that random variable $|h|$ is Rayleigh distributed. Also, γ is chi square distributed if $|h|^2$ is chi square distributed. Hence, the pdf of γ can be expressed as:

$$f(\gamma) = \frac{1}{\bar{\gamma}} \exp\left(-\frac{\gamma}{\bar{\gamma}}\right), \quad \bar{\gamma} > 0 \quad (16)$$

The BER is given by:

$$P_e = \int_0^\infty P_e(e/u) f_\gamma(u) du \quad (17)$$

Substituting the values of $P_e(e/u)$ and $f_\gamma(u)$:

$$P_e = \int_0^\infty \left\{ (1-p) \frac{\chi_0 + 2\sigma_G^2}{\alpha_i E_b} Q\left(\sqrt{\frac{h^2 \alpha E_b}{\chi_0 + 2\sigma_G^2}}\right) \exp^{u\gamma} + p \frac{\chi_0 + 2\sigma_G^2 + 2\sigma_I^2}{\alpha_i E_b} Q\left(\sqrt{\frac{h^2 \alpha E_b}{\chi_0 + 2\sigma_G^2 + 2\sigma_I^2}}\right) \exp^{u\gamma} \right\} d\gamma \quad (18)$$

IV. EXISTING IN SUPPRESSION STRATEGIES

A. BLANKING

In blanking method, the received signal is substituted by null if the signal is above blanking threshold T_b for IN cancellation at the receiver. No change is accrued in the phase and other parameters and only amplitude is changed [12].

$$r_i = \begin{cases} r_i & |r_i| < T_b \\ 0 & |r_i| > T_b \end{cases} \quad (19)$$

where T_b is known as the blanking threshold.

B. CLIPPING

The clipping method performs IN cancellation by clipping the received signal at T_c level if signal is above the threshold T_c . A clipping algorithm is used for cancellation of IN at the receiver where IN is being added by the channel during communication. In clipping, no change is accrued in the phase and other parameters and change occurs in the amplitude only [12].

$$r_i = \begin{cases} r_i & |r_i| < T_c \\ T_c e^{j\arg(r_i)} & |r_i| > T_c \end{cases} \quad (20)$$

where T_c is known as the clipping threshold.

C. BLANKING/CLIPPING

The blanking/clipping hybrid method clips the received signal at T_1 level if signal amplitude lies between the threshold T_1 & T_2 . The received signal is substituted by null if signal is above the threshold T_2 and remains unchanged if it is lower than the threshold T_1 . In this method, no change is accrued in the phase and other parameters and only amplitude

is changed [12].

$$r_i = \begin{cases} r_i & |r_i| < T_1 \\ T_1 e^{j\arg(r_i)} & T_1 < |r_i| \leq T_2 \\ 0 & |r_i| > T_2 \end{cases} \quad (21)$$

where the blanking threshold (T_2) is greater than the clipping threshold (T_1).

V. PROPOSED DNN FOR IN DETECTION

In order to detect IN, a deep learning strategy is devised as a deep neural network (DNN) which searches out the instances of IN corrupted samples. Theoretically, deep learning is a subset of machine learning which learns and improves its own performance using machine learning algorithms. In this section, the layout of proposed DNN, its input features, output and signal detection approach are presented.

A. DNN LAYOUT

The proposed DNN is designed with three fully connected hidden layers $U^{[1]}$, $U^{[2]}$ and $U^{[3]}$, each of which consists of n_1 , n_2 and n_3 neurons respectively. These layers map the input to the appropriate outputs. The number of layers and neurons are chosen by performing repeated experiments such that minimum training loss is achieved. The input to the DNN is a set of three features denoted by $z = [z^1, z^2, z^3]^T$ which are discussed in the next subsection and the output is represented by o/p' . An activation function af at each layer enables connection from the preceding layer to the following layer using parameter matrix and bias vector. The parameter matrix W and bias vector q connect the hidden layers with each other and to the output layer using the number of neurons (n_r) chosen for the R_{th} layer. The output has one layer which generates a binary sequence of '0' and '1'. The relationship between the hidden layers and the output layer o/p' is given by:

$$U^{[1]} = af^{[1]}(W^{[1]}z + q^{[1]}) \quad (22)$$

$$U^{[2]} = af^{[2]}(W^{[2]}U^{[1]} + q^{[2]}) \quad (23)$$

$$U^{[3]} = af^{[3]}(W^{[3]}U^{[2]} + q^{[3]}) \quad (24)$$

$$o/p' = af^{[4]}(W^{[4]}U^{[3]} + q^{[4]}) \quad (25)$$

$$ReLU(z) = \max(z, 0) \quad (26)$$

$$Sgmd(z) = \frac{1}{1 + e^{-z}} \quad (27)$$

Here $ReLU(z)$ represents Rectified Linear Unit as an activation function which is used by the hidden layers to give 0 when the input z is less than 0 and gives value equal to z otherwise. The activation function applied by the output is Sigmoid function denoted by $Sgmd(z)$ and is rounded off to give either 0 or 1 as the output. The training performance of DNN is measured using a cost function which calculates the average error associated with a predicted value of output for a training dataset. The error cost function is directly proportional to the difference of true output and predicted

output and is given as [29].

$$Error(W, q) = -\frac{1}{k} \left[\sum_{p=1}^k o/p_j \log(o/\hat{p}_j) + (1 - o/p_j) \log(1 - o/\hat{p}_j) \right] + \frac{\lambda}{2k} \sum_{r=1}^{R-1} \sum_{p=1}^{n_r} \sum_{q=1}^{n_r+1} W_{pq}^2 \quad (28)$$

Here k are training samples, λ is the regularization parameter and n_r is the number of neurons in R_{th} layer. The DNN uses backward propagation algorithm and is trained with Adam’s Optimizer [30] to optimize the network performance. While training, the DNN computes the values of parameter matrices W and the bias vectors q to reduce the error function $Error(W, q)$.

B. I/P FEATURES USED IN DNN FOR IN DETECTION

Correct I/P feature extraction and feature relationship with the output can help eliminate the problem of over-fitting which is often faced while training the DNN. The input features should be chosen so that redundancy is removed from the learning patterns of DNN, enabling better classification of sample as corrupted or un-corrupted. The noise patterns in the data samples can be identified if the test sample bit is analyzed along with its neighboring samples. For this purpose, two features have been chosen. The input layer comprises of the input sample value and the following two features:

- 1) ROAD Statistic Value: ROAD statistic is widely used to detect noise in 2-D images. It gives a high value for noisy samples and produces a low value otherwise [19]. In this research work, ROAD score is calculated using the following steps for 1-D sample data bits stored in a vector with dimensions of $1 \times 2n$: (a) The deviation magnitude between the current sample s_i and the remaining samples in the neighborhood (both right and left neighbors of data samples) is denoted by $Absd(i)$ and is calculated as:

$$Absd(i) = |s_i - [s_{i-n}, \dots, s_{i-1}, s_{i+1}, \dots, s_{i+n}]| \quad (29)$$

- (b) Sort $Absd(i)$ values in increasing order:

$$arr(i) = sort(Absd(i)) \quad (30)$$

- (c) The ROAD feature is computed by adding up the first n values of $arr(i)$:

$$ROAD = \sum_{i=1}^n arr(i). \quad (31)$$

- 2) Difference Median Output: The Difference median output is represented by err_i as follows:

$$err_i = s_i - median([s_{i-n}, \dots, s_i, \dots, s_{i+n}]) \quad (32)$$

where the variable $median$ represents a standard moving average filter taking median as the central tendency of any $2n + 1$ samples at a time.

C. IN SIGNAL DETECTION

DNN is a conceptual approach that has been widely used to represent complex non-linear systems with the help of proper training. To learn about IN instances in the incoming sample values, the chosen DNN inputs include incoming sample value, difference median output feature and ROAD statistic feature which are discussed above. The value of both input features is found to be high for an input IN sample and low for a input noiseless sample. Based on the learning of statistical traits of noisy samples, the DNN output is a binary sequence of 0’s and 1’s where 1 represents a noisy sample detection and 0 represents an un-corrupted sample.

VI. PROPOSED DNN FOR IN CLASSIFICATION

Due to varying yet randomly changing amplitudes of the impulse, it becomes a difficult task to classify a noisy sample as high or low impulse. To tackle this problem, a DNN is proposed to classify the IN into high and low impulse categories. Similar to the IN detection DNN in Section V, classification DNN comprises of 3 hidden layers with n_1, n_2 and n_3 neurons, one input layer and one output layer. ReLU function and Sigmoid Layer were chosen as activation functions in the hidden layers and output layer respectively. The DNN used backward propagation algorithm for training and Adam’s Optimizer for optimizing the network performance. The input consists of input sample and two computed features helpful in IN classification. The output consists of one layer which generates a binary sequence of ‘0’ and ‘1’. A ‘0’ at location k as an output showed that the received samples k was corrupted with low IN whereas a ‘1’ output indicated that the received sample contains high IN. The relationships among input, hidden and output layers have already been explained in Section V.

A. I/P FEATURES USED IN DNN FOR IN CLASSIFICATION

Although the arrival of IN is random and so is its amplitude but in theory, it follows binomial distribution. According to the distribution pdf, occurrence of high amplitude samples is less than the low amplitude samples in NOMA symbols. The high and low noise patterns in the corrupted samples can be identified if the current sample is analyzed along with its neighboring samples. For this purpose, following two significant features which are helpful in IN classification were chosen as input features along with the input sample value:

- 1) Difference median output: The Difference median output is represented in Section V (32) is denoted by err_i . The value of err_i feature is proven to be high for a high IN sample and low for low IN data.
- 2) Average Occurrence Probability output: The average value of occurrence probability of 64 samples (current sample and the next 63 samples) [5] is chosen as the second input feature for proposed DNN because it’s a strong indicator of high or low impulsivity. Since a high impulse arrival is less frequent than the low

TABLE 2. Simulation parameters.

DNN Simulation Parameters	
η	0.01
λ	0.1
n	5
n_1	20
n_2	20
n_3	10
NOMA Simulation Parameters	
α_1	0.333
α_2	0.667
T_b	1.6
T_c	1.5
T_1	1.5
T_2	1.8

impulse, the chances of its occurrence are low. This phenomenon is reflected in small values of the average probability output for high IN occurrence and vice versa.

VII. RESULT AND DISCUSSION

As discussed earlier, the presented DNN methods are implemented for NOMA downlink system in presence of IN over Rayleigh fading channel. Since the emphasis of this research work lies in IN detection and classification, BPSK modulation technique has been chosen for computational convenience. The proposed DNN based IN mitigation performance is compared with blanking, clipping and blanking/clipping methods respectively for a range of BER values. Furthermore, the performance of deep learning approach for IN classification is also evaluated.

A. SIMULATION PARAMETERS

The key simulation parameters which were common in training both the DNNs were set to the following optimal values:

Learning Rate η , a hyper-parameter for fine-tuning the DNN = 0.01; $\lambda = 0.1$; Number of sample bits (n) = 5; The number of neurons are set as $n_1 = 20$, $n_2 = 20$ and $n_3 = 10$ in the first, second and third hidden layers respectively. These simulation parameters are also reported in Table 2 for a better understanding.

The size of $W^{[1]}$, $W^{[2]}$, $W^{[3]}$ and $W^{[4]}$ parameter matrices are (20×3) , (20×20) , (10×20) and (10×1) respectively. The dimensions of bias vectors $q^{[1]}$, $q^{[2]}$, $q^{[3]}$ and $q^{[4]}$ are (20×1) , (20×1) , (10×1) and (1×1) respectively. Xavier initialization [31] is used to select the initial values of all the above chosen parameters since gradient descent method was observed to suffer from poor performance when randomly initialized.

B. DNN PERFORMANCE FOR IN DETECTION

The proposed DNN is trained using 10^6 NOMA-based BPSK symbols defined in (8) and (9) which are contaminated by IN represented by Gaussian-Bernoulli noise model in (1). In Fig. 2, BER performance of proposed DNN-based IN mitigation approach is evaluated for average probability of impulse occurrence.

Fig. 2 shows a comparison of BER performance of DNN based mitigation with blanking, clipping and blanking/clipping for IN detection. In general, BER performance is observed to improve with an increase in the SNR. It is evident that DNN outperforms conventional methods in terms of BER improvement. The results depicted in Fig. 2 show that blanking and clipping are very susceptible to noise at low levels of SNR and finding a suitable threshold to distinguish between desired and contaminated signals has been a tough task. On the contrary, the proposed DNN has successfully detected IN even at low levels of signal. This phenomenon can be seen at 5 dB SNR, where proposed DNN method has detected approximately 0.1 Mbits more true symbols out of 1 Mbits compared to conventional methods. At high SNRs, the DNN identified 700 more true symbols compared to conventional methods. This is because the conventional methods perform better detection at high SNR, therefore the performance difference between proposed DNN and conventional methods decrease.

It is noteworthy that the performance degradation level depends upon the user order. Higher-order users perform less SIC operation than the others and is strengthened by better received power. However, they experience interference from the lower-order users. In principle, a power coefficient α determines the transmit signal power of the lower-order user (user 1) as α_1 and that of the higher-order user (user 2) as $\alpha_2 = 1 - \alpha_1$. Therefore, they are quite sensitive to any additional noise/interference. In addition, it is observed that low IN power and low IN occurrence rate which do not affect the performance of the time division multiple access (TDMA) users have a degrading impact on NOMA users. Specifically, the performance of the higher-order NOMA users is affected and this degradation is expected to increase as the number of UEs increases. Considering the above mentioned scenario, we now show that the proposed DNN based approach can efficiently eliminate the IN and has retrieved the theoretical performance of NOMA user pair specifically for the higher-order user which is more sensitive to noise. In Fig. 3, the performance of NOMA user pair using conventional method (blanking/clipping) and DNN based IN mitigation technique has been illustrated. The comparison reveals that at 5 dB SNR, performances of user 1 and user 2 are approximately same hence a comparison is insignificant. At 30 dB, DNN identified 700 more true symbols for user 1 and 2000 more true symbols for user 2 compared to conventional method. Due to more contaminated symbols in user 2 signal, the BER performance of user 2 is poorer than user 1. Due to this reason, DNN gets a better chance of performing IN mitigation for user 2.

Fig. 3 shows the consistent performance of DNN for user 1 for IN mitigation compared to user 2. User 1 utilizes SIC to eliminate inter-user interference therefore, it has to counter IN noise only. In contrast, user 2 is inherently more effected by inter-user interference and IN, therefore user 2 has variant BER performance with different values of SNRs.

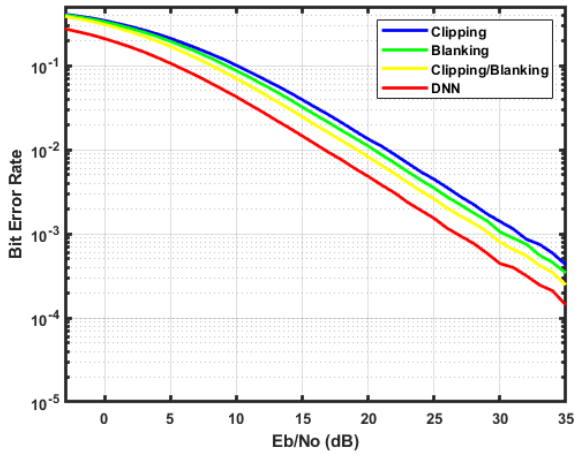


FIGURE 2. Comparison of different IN mitigation techniques.

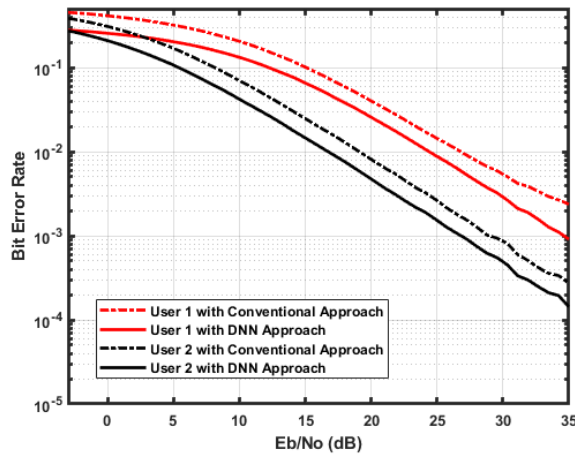


FIGURE 3. Performance of NOMA user pair using proposed DNN-based technique.

The DNN response for low and high impulse scenario is evaluated by testing the DNN on NOMA user pair in Fig. 4. A relatively highly IN (50 dB) as well as a lower IN (20 dB) scenarios have been considered. Overall, there is less effect of high or low IN on user 1 and user 2 for SNRs less than 10 dB. As the SNR increases, the performance of DNN is observed to be improved. This is due to the fact that at high SNR, ratio of signal versus IN is improved leading the signal level to become traceable for successful IN detection. In Fig. 5, DNN responses for different value of parameter ‘p’ are evaluated by testing the DNN on NOMA user pair. The test scenario includes high value of parameter ($p = 0.5$) for strong likelihood of impulse occurrence and low value of parameter ($p = 0.2$) for weak likelihood of impulse occurrence. It is observed that the effect of impulse probability parameter is very low because the DNN gets trained more efficiently on high or low impulse as compared to occurrence probability of impulse. The training of DNN is performed on 1 Mbits of

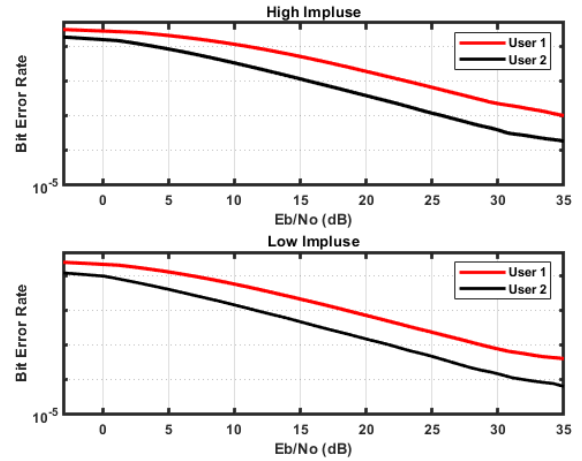


FIGURE 4. DNN performance for high and low IN scenarios for NOMA user pair.

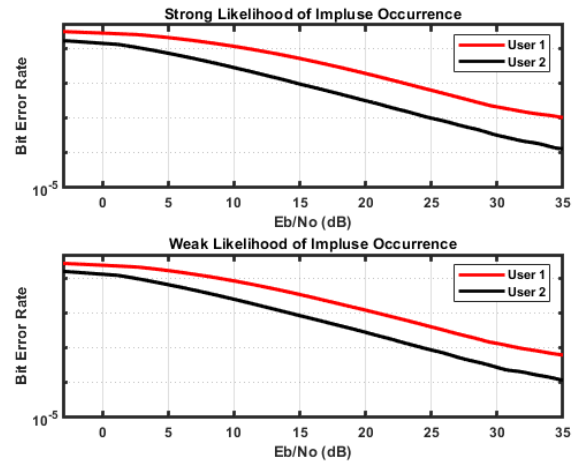


FIGURE 5. DNN performance for strong and weak likelihood of IN occurrence for NOMA user pair.

data samples so that the effect of probability is minimum on the result.

C. DNN-BASED CLASSIFICATION OF IN

The training set consists of 10^6 NOMA symbols with a range of E_b/N_0 that can enable the performance evaluation under various noise interference scenarios. The performance of DNN is evaluated for IN classification as high impulse and low impulse noise for 100 highly noise contaminated NOMA symbols at 5 dB SNR. A numerical value of 2 is considered as normalized threshold level for IN classification. Since IN arrival is random, the proposed DNN has classified the IN based on the probability of arrival and the median amplitude deviation from the neighboring samples as input features.

In Fig 6, high impulse detection using DNN is shown. High impulses are represented by red stems and low impulses which are removed using DNN are represented by blue circles on the x-axis. The x-axis represents 100 random IN

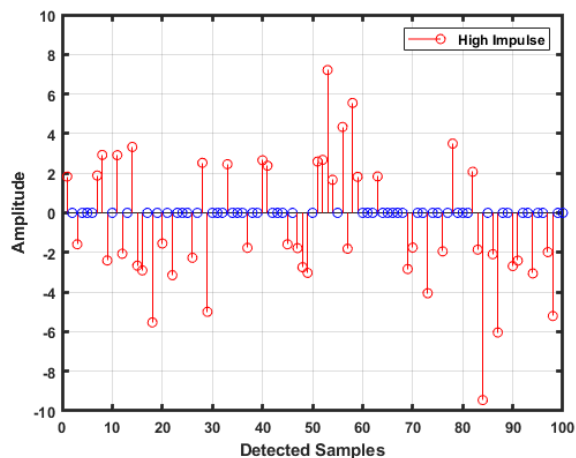


FIGURE 6. High impulse detection using DNN.

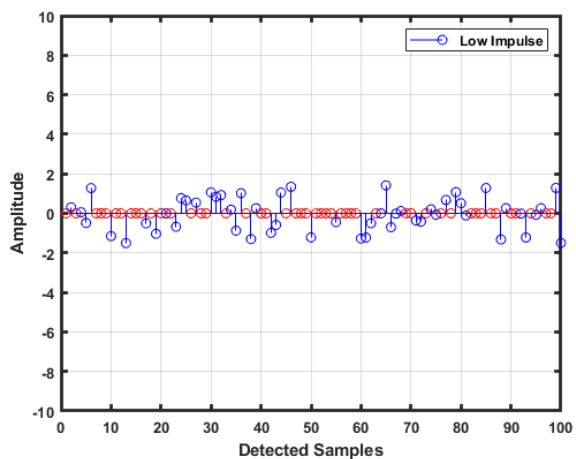


FIGURE 7. Low impulse detection using DNN.

samples and the y-axis represents the normalized amplitude of impulses. In Fig. 7, low impulse noise is represented by blue stems, and high impulses which are removed using DNN are represented by red circles on the x-axis. It is evident from Fig. 6 that all identified high impulses are higher than the normalized threshold level of 2. In Fig. 7, all identified low impulses are shorter than the normalized threshold level of 2 which shows the robust performance of DNN. The classification of impulse gives information about the nature of impulses which depends upon the type of source of impulse noise generation. It is clearly visible from Fig. 6 and Fig. 7 that most of the impulses are either low or high amplitude impulses and only two impulses are in the very high amplitude category. This type of information is valuable to set parameters of DNN for the IN detection.

The DNN has identified high IN in the incoming noisy NOMA symbols with an accuracy of 99% and low impulses with an accuracy of 87% respectively. However, DNN classified a few noiseless NOMA symbols as low IN instances.

This has been due to the random IN peakedness which is similar to the amplitude of noiseless NOMA symbols and cannot be accurately determined. Overall, the proposed DNN classified sufficient number of IN instances accurately. The IN classification in NOMA symbols using deep learning approach has been carried out for the first time according to authors' best knowledge in this research work and is a significant contribution of this research work.

VIII. CONCLUSION

In this research work, we have presented deep learning approaches for IN mitigation and classification for NOMA symbols. The proposed method for IN mitigation has shown superior performance when compared with several existing mitigation techniques. Another deep learning approach developed for IN classification has been attempted for the first time to classify IN as high impulse or low impulse noise with fairly good accuracy which can be used to improve the performance of IN detection models. The achieved results using deep learning methods show that deep learning has strong potential to solve several complex communication problems and can be explored further.

REFERENCES

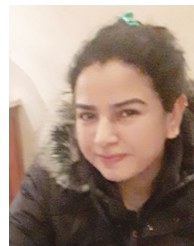
- [1] M. Aldababsa, C. Goztepe, G. K. Kurt, and O. Kucur, "Bit error rate for NOMA network," *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1188–1191, Jun. 2020.
- [2] M. Hussain and H. Rasheed, "Nonorthogonal multiple access for next-generation mobile networks: A technical aspect for research direction," *Wireless Commun. Mobile Comput.*, vol. 2020, Nov. 2020, Art. no. 8845371.
- [3] S. Sharma, V. Bhatia, and A. K. Mishra, "Wireless consumer electronic devices: The effects of impulsive radio-frequency interference," *IEEE Consum. Electron. Mag.*, vol. 8, no. 4, pp. 56–61, Jul. 2019.
- [4] B. Selim, M. S. Alam, G. Kaddoum, and B. L. Agba, "Effect of impulsive noise on uplink NOMA systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3454–3458, Mar. 2020.
- [5] B. Selim, M. S. Alam, J. V. C. Evangelista, G. Kaddoum, and B. L. Agba, "NOMA-based IoT networks: Impulsive noise effects and mitigation," *IEEE Commun. Mag.*, vol. 58, no. 11, pp. 69–75, Nov. 2020.
- [6] M. Hussain and H. Rasheed, "Performance of orthogonal beamforming with NOMA for smart grid communication in the presence of impulsive noise," *Arabian J. Sci. Eng.*, vol. 45, no. 8, pp. 6331–6345, Aug. 2020.
- [7] R. Barazideh, S. Niknam, and B. Natarajan, "Impulsive noise detection in OFDM-based systems: A deep learning perspective," in *Proc. IEEE 9th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2019, pp. 0937–0942.
- [8] C.-H. Yih, "Iterative interference cancellation for OFDM signals with blanking nonlinearity in impulsive noise channels," *IEEE Signal Process. Lett.*, vol. 19, no. 3, pp. 147–150, Mar. 2012.
- [9] D. F. Tseng, Y. S. Han, W. H. Mow, L. C. Chang, and A. J. H. Vinck, "Robust clipping for OFDM transmissions over memoryless impulsive noise channels," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1110–1113, Jul. 2012.
- [10] N. Rožić, P. Banelli, D. Begušić, and J. Radić, "Multiple-threshold estimators for impulsive noise suppression in multicarrier communications," *IEEE Trans. Signal Process.*, vol. 66, no. 6, pp. 1619–1633, Mar. 2018.
- [11] H. Oh and H. Nam, "Design and performance analysis of nonlinearity pre-processors in an impulsive noise environment," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 364–376, Jan. 2017.
- [12] K. Al-Mawali, A. Z. Sadik, and Z. M. Hussain, "Joint time-domain/frequency-domain impulsive noise reduction in OFDM-based power line communications," in *Proc. Australas. Telecommun. Netw. Appl. Conf.*, Dec. 2008, pp. 138–142.

- [13] R. Amiri, H. Mehrpouyan, D. Matolak, and M. Elkashlan, "Joint power allocation in interference-limited networks via distributed coordinated learning," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Aug. 2018, pp. 1–5.
- [14] G. Ndo, P. Siohan, and M. H. Hamon, "Adaptive noise mitigation in impulsive environment: Application to power-line communications," *IEEE Trans. Power Del.*, vol. 25, no. 2, pp. 647–656, Apr. 2010.
- [15] S. V. Zhidkov, "Analysis and comparison of several simple impulsive noise mitigation schemes for OFDM receivers," *IEEE Trans. Commun.*, vol. 56, no. 1, pp. 5–9, Jan. 2008.
- [16] G. Kaliraj and S. Baskar, "An efficient approach for the removal of impulse noise from the corrupted image using neural network based impulse detector," *Image Vis. Comput.*, vol. 28, no. 3, pp. 458–466, Mar. 2010.
- [17] R. Amiri, H. Mehrpouyan, L. Fridman, R. K. Mallik, A. Nallanathan, and D. Matolak, "A machine learning approach for power allocation in HetNets considering QoS," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–7.
- [18] H. Kong and L. Guan, "A noise-exclusive adaptive filtering framework for removing impulse noise in digital images," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 45, no. 3, pp. 422–428, Mar. 1998.
- [19] R. Garnett, T. Huegerich, C. Chui, and W. He, "A universal noise removal algorithm with an impulse detector," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1747–1754, Nov. 2005.
- [20] N. Ye, X. Li, H. Yu, L. Zhao, W. Liu, and X. Hou, "DeepNOMA: A unified framework for NOMA using deep multi-task learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2208–2225, Apr. 2020, doi: [10.1109/TWC.2019.2963185](https://doi.org/10.1109/TWC.2019.2963185).
- [21] Y. Xie, K. C. Teh, and A. C. Kot, "Deep learning-based joint detection for OFDM-NOMA scheme," *IEEE Commun. Lett.*, vol. 25, no. 8, pp. 2609–2613, Aug. 2021, doi: [10.1109/LCOMM.2021.3077878](https://doi.org/10.1109/LCOMM.2021.3077878).
- [22] A. Emir, F. Kara, H. Kaya, and H. Yanikomeroglu, "DeepMuD: Multi-user detection for uplink grant-free NOMA IoT networks via deep learning," *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 1133–1137, May 2021, doi: [10.1109/LWC.2021.3060772](https://doi.org/10.1109/LWC.2021.3060772).
- [23] N. Ye, J. An, and J. Yu, "Deep-learning-enhanced NOMA transceiver design for massive MTC: Challenges, state of the art, and future directions," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 66–73, Aug. 2021, doi: [10.1109/MWC.001.2000472](https://doi.org/10.1109/MWC.001.2000472).
- [24] A. Dubey, D. Sharma, R. K. Mallik, and S. Mishra, "Modeling and performance analysis of a PLC system in presence of impulsive noise," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2015, pp. 1–5.
- [25] S. Ghadimi and J. Hussain, "Effect of impulse noise on wireless relay channel," *Wireless Sensor Netw.*, vol. 4, no. 6, pp. 167–172, 2012.
- [26] S. Prakash, A. Bansal, and S. K. Jha, "Performance analysis of narrowband PLC system under Gaussian Laplacian noise model," in *Proc. Int. Conf. Electr., Electron., Optim. Techn. (ICEEOT)*, Mar. 2016, pp. 3597–3600.
- [27] S. Niranjayan and N. C. Beaulieu, "Analysis of wireless communication systems in the presence of non-Gaussian impulsive noise and Gaussian noise," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2010, pp. 1–6.
- [28] M. Hussain and H. Rasheed, "A computational power allocation scheme for fair NOMA downlink system," *J. Inf. Commun. Technol. Robotic Appl.*, vol. 9, no. 1, pp. 73–79, Jan./Jun. 2018.
- [29] P. Liu, K.-K. R. Choo, L. Wang, and F. Huang, "SVM or deep learning? A comparative study on remote sensing image classification," *Soft Comput.*, vol. 21, no. 23, pp. 7053–7065, 2017.
- [30] H. Tao and X. Lu, "On comparing six optimization algorithms for network-based wind speed forecasting," in *Proc. 37th Chin. Control Conf. (CCC)*, Jul. 2018, pp. 8843–8850.
- [31] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.



MUHAMMAD HUSSAIN received the B.E. degree in electronic engineering and the master's degree in communications engineering from the NED University of Engineering and Technology, Karachi, Pakistan. He is currently pursuing the Ph.D. degree in electrical engineering with Bahria University, Karachi.

He is currently a Senior Assistant Professor with Bahria University. His research interests include analysis and optimization of wireless communications for the next generation, such as 5G mobile/smart grid communication networks with a focus on multiple access, spectrum and power allocation, and interference/noise modeling.



HINA SHAKIR received the B.E. degree in electrical engineering from the NED University of Engineering and Technology, Karachi, Pakistan, and the M.S. degree in computer science from SZABIST, Karachi. She is currently pursuing the Ph.D. degree in electrical engineering with Bahria University, Karachi.

She is an Assistant Professor with the Electrical Engineering Department, Bahria University. Her research interests include bio-medical image analysis, inverse problems, optimization techniques, machine learning, and digital signal processing.



HAROON RASHEED received the B.E. and M.E. degrees in electrical engineering from the NED University of Engineering and Technology, and the M.S. and Ph.D. degrees in telecommunications from the Asian Institute of Technology, Thailand.

He is working as a Professor and the Head of the Electrical Engineering Department, Bahria University, Karachi, Pakistan. He has been working as a Research Associate in Global Center of Excellence (GCOE) Program at Tohoku University, Sendai, Japan, and CWC, Oulu, Finland. He is the author of several research publications in the refereed journals and IEEE international conferences. He has supervised several research thesis of M.S. and Ph.D. students and is a HEC approved Ph.D. Supervisor. His research interests include broadband wireless communications, including smart grid, network coding, channel fading, cognitive cooperative networks, game theory, and optimization techniques and advanced signal processing techniques in biomedical. He is a member of IEEE Executive Committee Member Karachi Chapter. He is also a Program Committee Member of FIT Conference, China Com, and NCMCS Conference. He was a recipient of the Research Grant from ICT Research and Development for his Project Energy Detector for Spectrum Sensing. He served as the Program Chair of 17th IEEE International Multi-Topic Conference (INMIC). He is an Editor of *The Journal of Bahria University Information and Communication Technologies (BUJICT)* and *Universal Journal of Communications and Networks* (Horizon Publications, USA) and a Technical Reviewer of *Wireless Personal Communications* journal.

• • •