Dendrogram-based Artificial Neural Network modulation classification for dual-hop cooperative relaying communications

H. Moulay, A.B. Djebbar, B. Dehri, Iyad Dayoub

Dendrogram-based Artificial Neural Network Modulation Classification for Dual-Hop Cooperative Relaying Communications

H. MOULAY[a], A. B. DJEBBAR[a,*], B. DEHRI and I. DAYOUB[b,c]

[a]TTNS Laboratory, Djillali Liabes University, Sidi Bel Abbes, Algeria
[b]IEMN, Département DOA (U.M.R 8520,CNRS), Université Polytechnique Hauts-de-France, le Mont Houy, Valenciennes, France
[c]INSA, F-59313, Université Hauts-de-France, Valenciennes, France

ARTICLE INFO

Keywords:
ANN
DANN
AMI
D-STBC
HOM
HOS

ABSTRACT

This paper contributes to the growing field of Artificial Neural Networks (ANNs) strategies of Automatic Modulation Identification (AMI) for Cognitive Radio (CR). Traditional AMI-based ANN methods suffer from many drawbacks such as overfitting due to ANN architectures with complex layers, low performance caused by falling into local minima, and the increased training time that keeps them from being used in real-time applications.

We propose generated Dendrogram to decompose the AMI task into sub-components using the ANN architectures for improved classification accuracy and less training time. In addition, one of the most significant challenges in AMI is identifying the modulation types and orders in cooperative systems due to the combination of several received signals propagating through different unknown channels. We propose dual-hop based on Amplify and Forward protocol (DH-AF) relaying system to generate the dataset used to train and test the proposed model. In addition, DH-AF relaying systems provide better coverage and signal reliability for existing wireless communication systems. We consider that the source communicates with the destination over a direct link ($S-D$) and indirect link via an intermediate relay ($S-\mathcal{R}-D$) using Distributed Space-Time Block Code (DSTBC). Then we used High Order Cumulants (HOC) and the High Order Moments (HOM) originating from the DSTBC-decoded signal as features, followed by the Dendrogram–based ANN (DANN) classifier. The simulation results confirm that the proposed method outperforms an ordinary ANN and other counterparts while taking less training time.

1. Introduction

Automatic Modulation Identification (AMI) is a procedure that identifies modulated signals without human input. It is commonly used in military and civilian domains [1]; it is an essential component of Cognitive Radio (CR) and aware systems that are able to sense the spectrum and convert to the optimal modulation schemes automatically [2]. In addition, it is utilized in 5G networks to monitor the spectrum and prevent unauthorized transmissions [3].

Two main algorithms are usually employed for AMI, Likelihood-Based (LB) and Feature-Based (FB). Although the LB approach can outperform the FB approach, it requires a high computational cost and depends on the preliminary data of the received signal [4].

In most cases, the receiver does not have any data concerning the modulation of the signals received; hence, researchers resort to FB approach to provide blind modulation detection and grant a reasonable tradeoff between low computational cost and high efficiency [5].

The performance of the FB approach depends on the Higher-Order Statistics (HOS), i.e., Higher-Order Moments (HOM) and Higher-Order Cumulants (HOC) derived from received signals.

Different classification frameworks are used in AMI, one of the most popular ones, Artificial Neural Networks (ANN) [4].

*Corresponding author
ahbou_djebbar@yahoo.fr (A.B. DJEBBAR)
ORCID(s): 0000-0002-2108-3234 (A.B. DJEBBAR)

Due to its excellent performance and compact architectures, ANN has been engaged in modulation identification tasks. In [6], the authors proposed a multi-stream network design for non-cooperative systems that raises the network width and enhanced the signal features, and in [7], the authors proposed a Convolutional Neural Network (CNN) for Multiple-Input and Multiple-Output (MIMO) systems.

In [8], the authors compared two types of supervised neural network approaches, i.e., Multilayer Perceptron (MLP) and Radial Basis Function (RBF)-based. However, a single learner suffers from the lack of hypothesis space or falling into local minima or maxima; consequently, it experiences unexpectedly poor performance [9].

We propose an AMI procedure that achieves significant accuracy for multi-class classification by merging numerous ANNs in a hierarchical clustering tree (Dendrogram) structure. We implement the proposed Dendrogram-based ANN (DANN) strategy to discriminate between modulation types and orders. In addition, the results of the DANN approach were benchmarked versus traditional ANN, Dendrogram-based Support Vector Machines (DSVM), and other proposed methods to determine its benefits.

Furthermore, to exploit the legacy spectrum more effectively under the umbrella of CR and meet the projected future mobile traffic demands as smartphones become more advanced and power-hungry, we employ dual-hop (DH) relaying systems using these smartphones [3].

DH relaying systems are integrated in various practical systems such as cellular phones and satellites to improve

throughput and energy efficiency [10]. Furthermore, present an alternative to the MIMO system by avoiding extra antenna that requires additional power cost and space in the radio devices [3].

Motivated by the previous advantages, we used the DH relaying system to generate the data for training and verifying the proposed model's efficiency. One of the most well-known relaying techniques is Amplify-and-Forward (AF) scheme because of its integrity and inexpensive implementation cost [10].

We propose in our work a Dendrogram-based on ANNs to create a deeper architecture by decomposing the AMI assignment into sub-components instead of using deeper layers that require more significant memory and an increased training time.

The literature on AMI for DH-AF relaying systems is limited compared to non-cooperative systems since pattern recognition becomes complicated due to the combination of several received signals propagating via different unknown channels. Therefore, we used DH-AF relaying systems to generate the data used in training and testing the proposed classifier to model the real-world data by modeling channel impairments under various Signal-to-Noise Ratio (SNR) levels.

We summarize the contribution of our work in the following steps:

- We leverage the spread-out antenna (installed in smartphones) over a wireless network to generate a DH-AF relaying system data used in training and testing the proposed model.
- After features extraction using HOM and HOC, we apply Principal Component Analysis (PCA) to reduce computational cost.
- We propose a novel classification approach DANN, constructed by ANNs following a Decision Tree (DT) structure through Ascendant Hierarchical Clustering (AHC) to identify the modulation type.
- We evaluate the proposed approach by comparing it to DSVM and to traditional ANN and other related works to validate the results.

The remainder of the paper is outlined as follows. In section 2, we exhibit the considered system model. In section 3, we present the necessary background and previous works on AMI. The proposed approach is described in section 4. We calculate the computational complexity of the used and the proposed methods in section 5. In section 6, we present simulation parameters and evaluate the results; and we conclude in section 7.

Notations: the upper (lower) bold letters indicates matrix (column vector). $(.)^{-1}$, $(.)^{T}$, $(.)^{*}$, and $(.)^{\mathcal{H}}$, denote inverse, transpose, conjugate, and hermitian operations, respectively. $\vartheta_{\mathbf{\Theta}}$=card($\mathbf{\Theta}$) denotes the cardinal of modulations pool $\mathbf{\Theta}$. $\mathcal{CN}(0,\sigma^2)$ denotes zero mean white complex Gaussian (ZMWCG) with variance σ^2.

2. System Model

DH-AF relaying systems are presented to meet MIMO's challenges, where instead of using collocated antennas, we follow a distributed pattern. In other words, each mobile node participates in a distributed antennas array [10], assigning an individual antenna along with its hardware, processing, and power resources to serve the communication between the source and destination nodes. Thus in our work, we consider DH-AF relaying system with one antenna at the source and relay nodes, whereas the destination node can be occupied with one or N_r antennas, with transmission diversity, mainly achieved by adopting the Distributed Space-Time Block Code (DSTBC) using AF protocol to reduce the system complexity and power cost [1].

As shown in Figure 1, the considered DH-AF relaying system comprises a direct link connecting the source and destination ($\mathcal{S}-\mathcal{D}$), an indirect link established with the assistance of relay ($\mathcal{S}-\mathcal{R}-\mathcal{D}$) using AF protocol [10].

In fact, broadcasting data is separated into two phases, each lasting two-time slots. In the first phase, the source node broadcasts the information to its destination. Fortunately, the relay node also receives this information due to the broadcasting nature of wireless channels. During the second phase, the process of amplifying and dispatching the received information to the destination is accomplished by the relay, while the source node remains silent to form the Alamouti STBC matrix $\mathbf{S}_i = \begin{bmatrix} s_{2i} & -s_{2i+1}^{*} \\ s_{2i+1} & s_{2i}^{*} \end{bmatrix}$ [11].

We assume that the channels linking the nodes present a flat quasi-static Rayleigh fading. In addition, these channels are supposed constant during a block of two symbols and may randomly vary from block to another. The noise terms are modeled as $\mathcal{CN}(0,N_0)$ zero-mean, complex Gaussian random variables with equal variance N_0 [1].

We suppose the broadcast of vectors of N_s data symbols, $\boldsymbol{s_i} = [s_{iN_s}, s_{iN_s+1}, \ldots, s_{iN_s+N_s-1}]^T$ that are randomly and independently drawn from the pool of constellations $\mathbf{\Theta} = \{\text{BPSK}, \text{8PSK}, \text{8PAM}, \text{16QAM}\}$. Where in our case N_s is fixed to 2 (N_s=2) to meet with Alamouti structure.

In the initial transmission phase, the source $\mathcal{S}$ conveys its first Alamouti encoded row vector $[s_{2i}, -s_{2i+1}^{*}]$ with transmitted power P_s to the destination and relay nodes. The received signal by the n_r-th antenna of the destination node, through direct link ($\mathcal{S}-\mathcal{D}$), is given by

$$\begin{cases} y_{sd,2i}^{(n_r)} = \sqrt{P_s} h_{sd,i}^{(n_r)} s_{2i} + \eta_{sd,2i}^{(n_r)} \\ y_{sd,2i+1}^{(n_r)} = \sqrt{P_s} h_{sd,i}^{(n_r)} (-s_{2i+1}^{*}) + \eta_{sd,2i+1}^{(n_r)} \end{cases} \tag{1}$$

The received signal at the relay, through ($\mathcal{S}-\mathcal{R}$) link, is given by:

$$\begin{cases} y_{sr,2i} = \sqrt{P_s} h_{sr,i} s_{2i} + \eta_{sr,2i} \\ y_{sr,2i+1} = \sqrt{P_s} h_{sr,i} (-s_{2i+1}^{*}) + \eta_{sr,2i+1} \end{cases} \tag{2}$$

where $h_{sd,i}^{(n_r)}$ and $h_{sr,i}$ are channels between the source and the destination, and between the source and the relay, respectively and $\boldsymbol{\eta}_{sd,i}^{(n_r)} = [\eta_{sd,2i}^{(n_r)}, \eta_{sd,2i+1}^{(n_r)}]$ and $\boldsymbol{\eta}_{sr,i} = [\eta_{sr,2i}, \eta_{sr,2i+1}]$

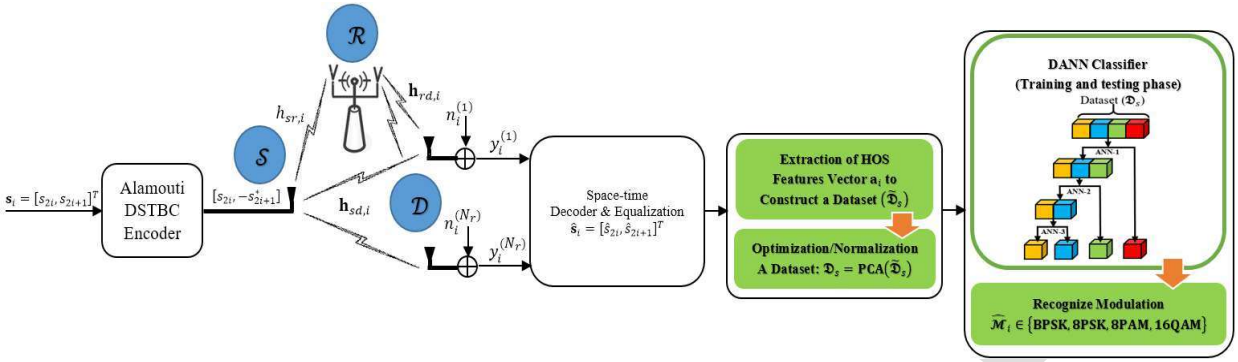

Figure 1: Baseband equivalent model of cooperative DH-AF relaying system.

are the corresponding AWGN noises at destination and relay nodes, respectively.

Alternatively, in the second phase, the relay emulates MIMO system in a distributed fashion by constructing the second row vector of the Alamouti matrix by applying conjugation to $y_{sr,2i}$ and sign reversion as well as conjugation to $y_{sr,2i+1}$, followed by permutation between the positions of the resulting symbols, then amplifying the signal by a factor $\mathcal{A}_f$; the received vector at the destination is modeled as

$$\mathcal{A}_{f,i}\sqrt{P_r}[-(y_{sr,2i+1})^*, (y_{sr,2i})^*] \tag{3}$$

where P_r and $\mathcal{A}_{f,i}=\sqrt{P_s(P_s|h_{sr,i}|^2+N_0)^{-1}}$ are the power and the amplification factors at $\mathcal{R}$ respectively.

The received signal provided from $\mathcal{R}$ to $\mathcal{D}$, via the indirect link in vector form, is

$$\begin{cases} y^{(n_r)}_{rd,2i} = h^{(n_r)}_{srd,i}s^*_{2i} + \eta^{(n_r)}_{srd,2i} \\ y^{(n_r)}_{rd,2i+1} = h^{(n_r)}_{srd,i}(s_{2i+1}) + \eta^{(n_r)}_{srd,2i+1} \end{cases} \tag{4}$$

where the equivalent channel $h^{(n_r)}_{srd,i}$ and noise are defined by

$$\begin{cases} h^{(n_r)}_{srd,i} = \mathcal{A}_{f,i}\sqrt{P_rP_s}h^{(n_r)}_{rd,i}h^*_{sr,i} \\ \eta^{(n_r)}_{srd,2i} = \eta^{(n_r)}_{rd,2i} + \mathcal{A}_{f,i}\sqrt{P_r}h^{(n_r)}_{rd,i}\eta^*_{sr,2i} \\ \eta^{(n_r)}_{srd,2i+1} = \eta^{(n_r)}_{rd,2i+1} - \mathcal{A}_{f,i}\sqrt{P_r}h^{(n_r)}_{rd,i}\eta^*_{sr,2i+1} \end{cases} \tag{5}$$

With matrix notation, the resulting signals sent through direct and indirect links at the destination node can be rewritten as follow:

$$\underbrace{\begin{bmatrix} y^{(1)}_{sd,2i} + y^{(1)}_{rd,2i+1} \\ (y^{(1)}_{rd,2i})^* + (y^{(1)}_{sd,2i+1})^* \\ \vdots \\ y^{(N_r)}_{sd,2i} + y^{(N_r)}_{rd,2i+1} \\ (y^{(N_r)}_{rd,2i+1})^* + (y^{(N_r)}_{sd,2i+1})^* \end{bmatrix}}_{\mathbf{y}_i} = \mathbf{H}_i\mathbf{s}_i + \boldsymbol{\eta}_i \tag{6}$$

where the equivalent channel matrix $\mathbf{H}_i$ is defined by

$$\mathbf{H}_i = \begin{bmatrix} \sqrt{P_s}h^{(1)}_{sd,i} & \mathcal{A}_{f,i}\sqrt{P_sP_r}h^*_{sr,i}h^{(1)}_{rd,i} \\ \mathcal{A}_{f,i}\sqrt{P_sP_r}h_{sr,i}(h^{(1)}_{rd,i})^* & -\sqrt{P_s}(h^{(1)}_{sd,i})^* \\ \vdots & \vdots \\ \sqrt{P_s}h^{(n_r)}_{sd,i} & \mathcal{A}_{f,i}\sqrt{P_sP_r}h^*_{sr,i}h^{(n_r)}_{rd,i} \\ \mathcal{A}_{f,i}\sqrt{P_sP_r}h_{sr,i}(h^{(n_r)}_{rd,i})^* & -\sqrt{P_s}(h^{(n_r)}_{sd,i})^* \end{bmatrix} \tag{7}$$

Since $\mathbf{H}_i$ is complex and orthogonal we have

$$\mathbf{H}_i^{\mathcal{H}}\mathbf{H}_i = \left(\sum_{n_r=1}^{N_r} P_s|h^{(n_r)}_{sd,i}|^2 + A^2_{f,i}P_sP_r|h^2_{sr,i}||h^{(n_r)}_{rd,i}|^2\right)\mathbf{I}_{N_s} \tag{8}$$

For signal recovery, Zero-Forcing (ZF) equalization matrix $\mathbf{G}_i$ is applied to the received signal as follows [12]

$$\hat{\mathbf{s}}_i = \mathbf{G}_i\mathbf{y}_i \tag{9}$$

where $\mathbf{G}_i$ is defined by

$$\mathbf{G}_i = \left(\mathbf{H}_i^{\mathcal{H}}\mathbf{H}_i\right)^{-1}\mathbf{H}_i^{\mathcal{H}} \tag{10}$$

3. Previous Work in AMI

Various techniques have been deployed in modulation detection, such as classification and clustering, to improve the accuracy of detection. This section reviews the different studies in this field and presents an overview of distinct kinds of learning in the context of AMI.

Classification and clustering are two different techniques used for supervised and unsupervised learning. Identifying the input instances according to the prior knowledge of labels is called classification, while grouping them based on their similarity is known as clustering; consequently, there is no need for training or testing datasets in clustering [13].

Previous research aims to design a robust and quickly deployable AMI algorithm that can be a powerful tool and deliver a good performance in real-time scenarios. For this purpose, [11] proposed an Automatic Modulation

Classification (AMC) for STBC-MIMO systems with a Deep Learning (DL). Sparse Auto-Encoders (SAE) and RBF Network (RBFN) were used as classifiers. Limited-memory Broyden-Fletcher-Goldfarb-Shannon (L-BFGS) was used for optimization, and optimal Maximum-Likelihood (ML) was used for performance evaluation. Moreover, the performance evaluation carried out perfect Channel State Information (CSI), and channel estimation errors impact the classification results. The authors in [12] used DL for modulation classification based on CNN for MIMO systems using ZF equalization and HOS as features under perfect CSI conditions.

In [14], the authors proposed an identification of real signals by a smart embedded devices using Amplitude Shift Keying (ASK) and Frequency Shift Keying (FSK) modulations where the signal is identified based on several classifiers, ANNs, SVM, and DT. While the authors in [15] proposed ANN-based Adaptive Modulation and Coding (AMC) for link adaptation in MIMO-Orthogonal Frequency Division Multiplexing (MIMO-OFDM) systems to optimize the best modulation and coding scheme under the Packet Error Rate constraint, they compared the performance of the ANN with the K-Nearest Neighbors (KNN) classifier. As a result, ANN exhibited better performance and robustness than the KNN.

Another recent work in [16] argued that exploiting the complementary information from I/Q multi-channel, I-channel, and Q-channel data and utilizing spatial and temporal properties in the signal may increase the AMI. They integrated one-Dimensional (1-D) convolutional, two-Dimensional (2-D) convolutional, and long short-term memory (LSTM) layers to extract the features from individual and combined I/Q samples.

Despite the positive effects and the wide spread of DL outlined above, negative issues also need to be considered, such as generalization errors and stalling in local minima caused by complex and more profound network architecture that does not suit the signal recognition. In addition, the most significant disadvantage is the training time that precludes them from being used in real-time applications. Hence, researchers often resort to clustering to decrease complexity and the training time.

In [17], the authors proposed AMI based on a set of unique properties of modulations. The separation between FSK and MSK modulations from PSK and QAM modulations was accomplished by estimating the In-phase and Quadrature (I/Q) constellation diagram of the received signal using k-means and k-centers. The modulation identification was made by calculating the correlation between the received signal and the data recorded in the database using an ML classifier.

Likewise, in [13] the authors discussed modulation identification using different clustering algorithms based on I/Q samples including hierarchical clustering for the first time. The modulation order was recovered by clustering the I/Q samples, and then the centroids of these clusters were obtained using the relevant clustering algorithms. Finally, the modulation type is estimated by comparing the obtained centroids to the modulation constellation diagram using the minimum distance classifier. However, based on the results observed on the above mentioned algorithms, modulation identification using clustering requires a large sample size, many algorithms, and many stages; in addition, it cannot yield high accuracy under low SNR values.

Since clustering analysis can locate the distribution characteristics and benefit connections among data attributes, combining it with neural networks results in a powerful tool to defeat the drawbacks of traditional ANN.

More recently, the authors in [18] proposed modulation classifiers based on ResNet-50 and Inception ResNet-V2 deep learning model using transfer learning. The authors used constellation. However, they did not use clustering to recover the constellation. Instead, a color image has been produced using filters consisting of three masks based on the signals' constellation density. Modulation classification has been done in three stages by training both models with generated color images.

Unfortunately, the literature on combining clustering with ANN in AMI is rarely analyzed and suffers from many gaps and shortcomings. To overcome this drawbacks, this paper proposes merging multiple ANNs in the Hierarchical Clustering (HC) framework to establish the robustness and high accuracy for modulation classification.

4. Proposed Method

As with any standard AMI system, the proposed system comprises two subsystems, features extraction and modulations classification. The features extraction subsystem consists of two main steps: preprocessing step and the features selection step. The processing of the received signal improves the cognition of communication systems by smoothing the features extraction phase if it is performed well. Otherwise, it severely degrades the accuracy of the classifier [19]. After the preprocessing phase, the features extraction begins.

Previous works [19, 20, 21] proved that adopting HOC and HOM as features of interest for signal classification can produce good performance with extreme computation simplicity.

The features used in this paper are formed by sequences of HOM and HOC up to sixth order to construct $\mathcal{T}$-dimension HOS features vectors. Afterward, we will use it as input to the classifier after applying PCA.

Table 1 presents several theoretic values of HOM and HOC for various modulation types [22, 7].

4.1. Classification tools

4.1.1. ANN

ANN is a Machine Learning (ML) technique capable of pattern recognition; in fact, its computational model is inspired by the human brain's structure and behavior.

It comprises three layers: the input layer that associates each neuron with one feature in the features vector $\mathbf{a}_i$ =

Table 1
Some statistical moments and cumulants for several modulation types.

	M_{60}	M_{61}	M_{63}	C_{60}	C_{61}	C_{63}
BPSK	1	1	1	16	16	16
8PSK	0	0	1	0	0	4
8PAM	3.62	3.62	3.62	7.19	7.19	7.19
16QAM	0	0.38	2.08	0	1.8	1.8

$[a_{i,1}, a_{i,2}, \ldots, a_{i,\mathcal{T}}]^T$ of the training data. The hidden layer applies a given transformation to the input values via weighted connections. It connects input nodes to output nodes or hidden nodes, or it connects hidden nodes to other hidden nodes, depending on the complexity of the model [8], where the choice of the structure defines the results.

In our work, we adopt multilayer Feed-Forward Neural Networks, where each layer directly supplies the next layer, feeding the inputs forward through the network. The output of the m-th node in the first hidden layer is [8]:

$$c_{i,m}^{(1)} = g^{(1)}\left(\sum_{t=1}^{\mathcal{T}} w_{m,t}^{(1)} a_{i,t} + w_{m,0}^{(1)} a_{i,0}^{(1)}\right) \tag{11}$$

where $a_{i,0}^{(1)}$ denotes the bias of the input layer, and $w_{m,t}^{(1)}$ denotes a weight in the first hidden layer, connecting t-th input to m-th hidden node, where $w_{m,0}^{(1)}$ is the bias weight for the first hidden layer. $g^{(1)}$ is an activation function.

The output of k-th node in the J-th hidden layer is mathematically described as follows [8]

$$b_{i,k}^{(J)} = g^{(J)}\left(\sum_{p=1}^{\mathcal{P}} w_{k,p}^{(J)} c_{i,p}^{(J-1)} + w_{k,0}^{(J)} c_{i,0}^{(J)}\right) \tag{12}$$

We consider, $c_{i,0}^{(J)}$, and $\mathcal{P}$ is the number of nodes in the $(J-1)$-th hidden layer.

Finally, the output layer produces an output value called label N_i corresponding to the modulation type and order prediction chosen from the candidate pool $\boldsymbol{\Theta}$. In the output layer, the output of the n-th node, using the sigmoid activation function $\tilde{g}$, can be written as [8]

$$y_{i,n} = \tilde{g}\left(\sum_{k=1}^{\mathcal{K}} \tilde{w}_{n,k} b_{i,k}^{(J)} + \tilde{w}_{n,0} \tilde{b}_{i,0}\right) \tag{13}$$

4.1.2. Dendrogram-based ANN

If $\boldsymbol{\Theta}$ is the target modulation pool consisting of $\vartheta_{\boldsymbol{\Theta}}$ modulations, then we aim to select $\mathcal{T}$-dimension HOS features vectors $\mathbf{a}_i = [a_{i,1}, a_{i,2}, \ldots, a_{i,\mathcal{T}}]^T$ to differentiate between each modulation type and order, then we apply features normalization using PCA.

After splitting the data into 90% training subset and 10% testing subset, we decompose the original multiclass problem using a class hierarchy from which a Dendrogram of classifiers is constructed. The Dendrogram is built by applying an agglomerative hierarchical clustering algorithm. The agglomerative approach works in a bottom-up manner, starting by assigning each element to a single cluster (singleton) and then iteratively merging pairs of clusters until obtaining only one cluster. Clusters are joined based on a similarity function between them. We adopted linkage based on correlation (the shortest 1-"Sample correlation" between clusters) [23].

Thus, the clusters formed between the leaves (the classes) and the root constitute the Dendrogram's internal nodes. The generated Dendrogram provides insight into how close each class is to other classes in the features space.

Once the Dendrogram is produced, we train binary ANNs for every non-leaf node (or parent node) in the class hierarchy (including the root) to distinguish between its child nodes. Values closer to "0" in the linkage indicate that the two classes are nearly identical, while values closer to "1" indicate that the two classes are entirely distinct. As a result, we obtain a set of ANNs arranged in a tree. Then, to classify a new instance, the tree of classifiers is traversed in a top-down manner applying the classifiers from the root until a leaf is reached (see Figure 2).

The steps of the proposed DANN approach are summarized in Algorithm 1.

Algorithm 1 The proposed DANN-based AMC

Input: Sampled received signal $\mathbf{y}_i$. The algorithm requires a perfect CSI. $\vartheta_{\boldsymbol{\Theta}}$: Number of classes in the candidate pool $\boldsymbol{\Theta}$.
Output: $\hat{\mathcal{M}}_i \in \boldsymbol{\Theta}$.
for $iter = 1$ to $\vartheta_{\boldsymbol{\Theta}}$.
for $i = 1$ to M_c (Monte Carlo trials per modulation) **do**.
  Based on the sampled received signal $\mathbf{y}_i$, estimate $\hat{\mathbf{s}}_i$ using equation (9).
  Extract features using the sequences of HOM and HOC to construct $\mathcal{T}$ dimension HOS features vectors $\mathbf{a}_i = [a_{i,1}, a_{i,2}, \ldots, a_{i,\mathcal{T}}]^T$.
  end for i.
  end for $iter$.
  Construct a dataset $\tilde{D}_s$ (set of classified instances) given in the form of a matrix with $\mathcal{T}$ rows and N_{tot} columns ($N_{tot} = \vartheta_{\boldsymbol{\Theta}} \times M_c \times N_s$).
  Normalize predictor variables to produce a new subset D_s=PCA($\tilde{D}_s$).
  Split the entire dataset into two subsets (90% for the training subset and 10% for the testing subset).
  Defining the tree with its binary branchings (structure of the Dendrogram) by computing the '1-Sample correlation' values of the features for each class.
  Build the model DANN by training binary ANN at each parent node.
  Using the testing subset, predict the modulation type and order ($\hat{\mathcal{M}}_i$).

Table 2
Configuration of the classifiers used in the proposed work.

Method	Type	Input	Hidden Layer1 Neurons	Hidden Layer2 Neurons	Penalty factor	RBF	Learning rate	Output
DANN	Feed-Forward Neural Network	$\mathcal{T}$-neurons	10	15	/	/	0.01	ϑ_{Θ}-neurons
DSVM	Binary SVM	$\mathcal{T}$	/	/	2	RBF-Kernel ($\sigma = 10$)	/	$N_i \in \{0, ..., \vartheta_{\Theta} - 1\}$
ANN	Feed-Forward Neural Network	$\mathcal{T}$-neurons	10	15	/	/	0.01	ϑ_{Θ}-neurons

Following the previous steps yields the Dendrogram shown in Figure 2. At the root node, the first binary decision appears for 16QAM versus the rest of the modulation {BPSK, 8PSK, 8PAM}. Hence, the 16QAM class is a terminal node (leaf) and will be considered a negative class during the training of ANN-1, while the remaining combined three classes are positive. Similarly, the second binary classifier in the tree ANN-2 considers 8PAM as a negative class and elements of {BPSK, 8PSK} as positive. Finally, ANN-3 discriminates the elements of BPSK from those of 8PSK ones.

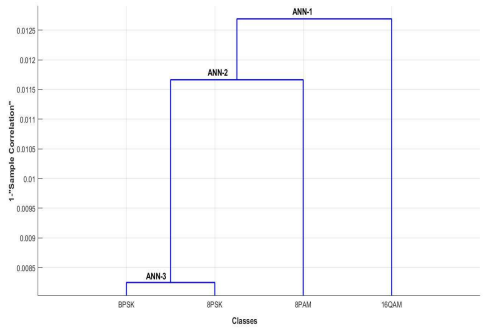

Figure 2: Dendrogram constructed through AHC illustrates the multiple ANNs produced for the candidate pool $\boldsymbol{\Theta}$.

4.1.3. DSVM

Like DANN, DSVM is a hierarchical cluster employing SVMs through binary decisions to differentiate between modulation type and order. To guarantee an unbiased and coherent comparison, we used the same structure used for DANN by following the same binary decisions and using clustering based on correlation. A previous study using sleep stage classification revealed that combining multiple deep SVMs might reach more increased classification accuracy [23]; however, the advantage of fusion SVMs in Dendrogram is unknown in modulation classification.

Based on the configurations reported in Table 2, DANN and ANN have the same parameters; however, ANN is trained on the entire training dataset while DANN decomposes the AMI assignment into sub-components; hence it is trained on smaller datasets. Following the same strategy as DANN, DSVM also decomposes the AMI assignment into sub-components using binary SVMs to perform the classification based on the parameters mentioned in Table 2.

5. Computational complexity

The computational complexities at each SNR of the used and proposed methods, in this paper, are summarized in Table 3.

Table 3
Computational complexity of ANN and SVM based approaches.

	Training	Testing
ANN	$\mathcal{O}(\mathcal{T} N_{tr} \vartheta_{\Theta}) P^{N_{HL}} I)$	$\mathcal{O}(\mathcal{T} N_{te} \vartheta_{\Theta})$
SVM	$\mathcal{O}(\mathcal{T} (N_{tr} \vartheta_{\Theta})^2)$	$\mathcal{O}(\mathcal{T} N_{te} \vartheta_{\Theta})$
DANN	$\mathcal{O}(\mathcal{T} 2N_{tr}(\vartheta_{\Theta}\text{-}1) P^{N_{HL}} I)$	$\mathcal{O}(\mathcal{T} 2N_{te}(\vartheta_{\Theta}\text{-}1))$
DSVM	$\mathcal{O}(\mathcal{T}(2N_{tr}(\vartheta_{\Theta}\text{-}1))^2)$	$\mathcal{O}(\mathcal{T} 2N_{te}(\vartheta_{\Theta} - 1))$

In Table 3, N_{tr} is the number of training instances, N_{te} is the number of testing instances, I is the number of iterations used in ANN and DANN, and N_{HL} is the number of hidden layers, where each containing P neurons. $Csys_{tr} = \mathcal{O}(N_{tr} \vartheta_{\Theta} N_r N_s)$ and $Csys_{te} = \mathcal{O}(N_{te} \vartheta_{\Theta} N_r N_s)$ are the DH-AF relaying signals generation computational load without classification tools for training and testing steps respectively. Consequently, the computational complexity using different classifiers for training and testing steps are obtained by adding $Csys_{tr}$ to each element of the first column and by

adding $Csys_{te}$ to each element of the second column of the Table 3, respectively.

6. Simulation results

The proposed classification methods mentioned above were tested through a set of experiments employing Monte-Carlo simulations.

The proposed approach was confirmed for the digital modulation pool $\mathbf{\Theta}$={BPSK,8PSK,8PAM,16QAM}. M_c = 1000 Monte Carlo trials per modulation were conducted for each SNR value.

We consider that the SNR across source–relay link (SNR_{sr}) and the SNR across relay–destination links (say SNR_{rd}) are fixed to $SNR = SNR_{rd} = SNR_{sr}$.

We generated a random message and Rayleigh fading channel from i.i.d. zero-mean independent complex Gaussian random variables with the same variance in each run, and the number of samples were varied between $N = 512$, and 2048 samples.

We consider a single relay for all the experiments, and the STBC encoder type Alamouti with (2×1) ($N_r = 1$) and (2×2) ($N_r = 2$) antennas configurations are inspected.

6.1. Evaluation Measures for Classifiers Performances

We employed several performance metrics to evaluate the proposed method, namely the Accuracy (ACC), Sensitivity (SEN), Specificity (SPE), Precision (PRE), and F-SCORE. Equation (14) states that they are derived by true and false, positive and negative rates (TP, TN, FP, and FN) extracted from the confusion matrix.

$$\begin{cases} ACC = ((TN + TP))/((TP + TN + FP + FN)) \\ SEN = TP/((TP + FN)) \\ SPE = TN/((TN + FP)) \\ PRE = TP/((TP + FP)) \\ F - SCORE = 2x(SENxPRE)/(SEN + PRE) \end{cases} \tag{14}$$

We compute the Probability of Correct Classification P_{cc} across all classes to evaluate the classifier's performance. It is employed frequently as an individual metric of global performance in AMI. P_{cc} is the average of the diagonal components of the confusion matrix for each value of SNR, it is defined as:

$$P_{cc} = \frac{N_c}{N_{tot}} x100 \tag{15}$$

where $N_c = \sum_{\hat{\mathcal{M}}_n \in \mathbf{\Theta}} N_{\mathcal{M}_n}$, with $N_{\mathcal{M}_n}$ represents the number of trials when the modulation is correctly recognized. $\mathcal{M}_n$ and $\hat{\mathcal{M}}_n$ are the correct and estimated modulations types and orders, respectively.

The performance of the DANN strategy is assessed by operating several validation metrics in section (6.2), including sensitivity, specificity, and precision (Figure 3). The simulation results are compared to other methods such as DSVM and ANN in section (6.3), in terms of Accuracy and F-score (Figure 7), and in terms of probability of correct classification (Figure 4 and Figure 5) and by calculations of confusion matrices (Figure 6).

In Section (6.4), we conducted an impartial comparison between the performances of the DANN method with popular existing methods in other related works (see Table 4).

6.2. DANN performance

Figure 3 depicts the PRE, SEN, and SPE of DANN. All modulations are tracked with the SPE above 0.94. The highest performances were observed for BPSK and 8PAM with PRE and SEN of 0.99 and SPE of 1. On the other hand, 8PSK and 16QAM are detected with lower SPE of 0.94 and 0.95, respectively. In addition, they experience a drop in the PRE (0.83 for 8PSK and 0.84 for 16QAM) and in the SEN (0.84 for 8PSK and 0.82 for 16QAM).

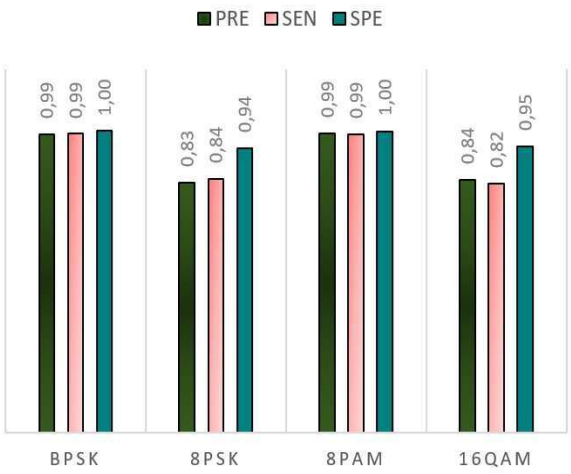

Figure 3: Class-specific metrics for DANN performance, at SNR=0dB, $N_r = 2$, and $N = 512$.

The explanation for the drop in PRE and SEN for 8PSK and 16QAM modulations can be derived from the confusion matrix of the DANN classifier (see Figure 6a).The values of the off-diagonal elements quantify the misclassification rate. Hence, the DANN confusion matrix shows that 16.13% of 8PSK modulation was incorrectly predicted as 16QAM, and 17.65% of 16QAM modulation was incorrectly predicted as 8PSK. This intense confusion between 8PSK and 16QAM accounts for their low SEN.

6.3. Comparison between DANN, DSVM and ANN

We compare, in this subsection, the proposed DANN using a DH-AF relaying system against DSVM and ANN algorithms in respect of P_{cc} as a function of SNR by varying the number of samples N and the number of receiving antennas N_r.

Figure 4 plots the P_{cc} of different classifiers for different values of N, the number of receiving antennas for this set

of experiments is fixed to $N_r = 2$. At low SNR values, we observe progress in classification accuracy as N increases from 512 to 2048 samples, the proposed classifier achieved the most significant improvement with 12% P_{cc}, because small sample size of N makes the test data-limited and does not appear to achieve completely error-free performance particularly at low SNR values.

However, the impact of N on P_{cc} diminishes when the SNR≥0dB, due to the time and space diversity gain achieved by DH-AF relaying system that reduces the errors rate of the received signal, thus improving the quality of the dataset for the training process.

As presented in Figure 4, the proposed DANN approach yields a promising performance compared to the other techniques. The results indicate that DANN grants more than 90% P_{cc} under low SNR rates, as it upgrades the performance when SNR=-5dB, where we can see that the P_{cc} of DANN is about 4% and 5% higher than ANN and DSVM, respectively when $N = 512$, while it exhibits a 7% and 10% higher P_{cc} than ANN and DSVM at the same SNR value when $N = 2048$.

In addition, it attains 100% P_{cc} when the SNR surpasses 5dB. This improvement is induced through training sub-classifiers individually on its subset of children labels, which improves the prediction performed using this hierarchical architecture.

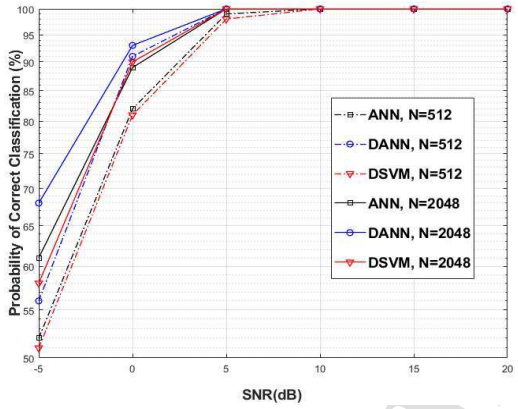

Figure 4: Classification performance per SNR(dB) of cooperative DH-Relaying system with $N_r = 2$, and different symbol length $N = 512, 2048$.

Figure 5 depicts the P_{cc} versus SNR for the DH-AF relaying system. Two antenna configurations (2×1) and (2×2) were inspected, considering $N = 512$ samples. The proposed method can maintain better P_{cc} than the others in both configurations, especially at low SNR values. It reaches $P_{cc} = 41\%$ at SNR=-5dB when $N_r = 1$ and 56% when $N_r = 2$, at the same SNR point. When SNR=0dB, the proposed classifier reaches 91% when $N_r = 2$ and 50% when $N_r = 1$ and reaches 100% regardless of the antennas configuration for SNR≥ $15dB$.

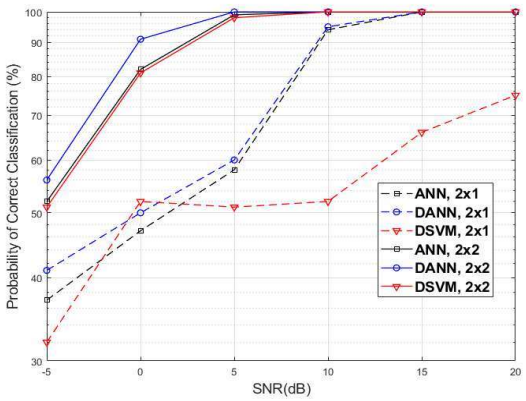

Figure 5: Classification performance per SNR(dB) of cooperative DH-Relaying system with $N = 512$, and different configuration antennas $N_r = 1, 2$.

ANN exhibits exciting behavior exceeding DSVM. However, We observe that for $N_r = 1$, DSVM presents poor performance with P_{cc} around 50% at SNR range [0dB, 10dB] and less than 80% at SNR=20dB. The explanation behind the results obtained by the DSVM classifier can be inferred from Figure 6c. We can see that when $N_r = 2$ DSVM confuses between 8PSK and 16QAM, where 50% of 8PSK modulation was misclassified as 16QAM. Hence, decreasing the number of receive antennas to $N_r = 1$ increases the distortion in DSVM classification and significantly affects its performance in this SNR range.

Based on the results, we conclude that DSVM is not robust against noise when $N_r = 1$.

One can see that the proposed method is robust because decomposing the AMI task into sub-components using the ANN architectures lowers the misclassification rate and defeat the distortion between 16QAM and 8PSK.

To investigate the classification performance of the proposed method, Figure 6 illustrates three confusion matrices generated by the proposed DANN and the other counterpart methods across all four classes for SNR=0dB, $N_r = 2$, and $N = 512$ samples. Most classification errors appear to be found for 8PSK and 16QAM, which can be interpreted using Table 1. We can see that their HOS have the closest values. Consequently, they reflect adjacent instances in the features space, which cause distortion between their instances under low SNR values.

Our results prove that generating enhanced Dendrogram design utilizing AHC is a fundamental source of performance optimization, particularly for 8PSK, where the proposed DANN framework achieves a higher classification performance.

The F-score and Accuracy metrics of the proposed method are compared with DSVM and ANN methods in Figure 7. The performance of the proposed classifier, in terms of the F-score and ACC, is better than applying

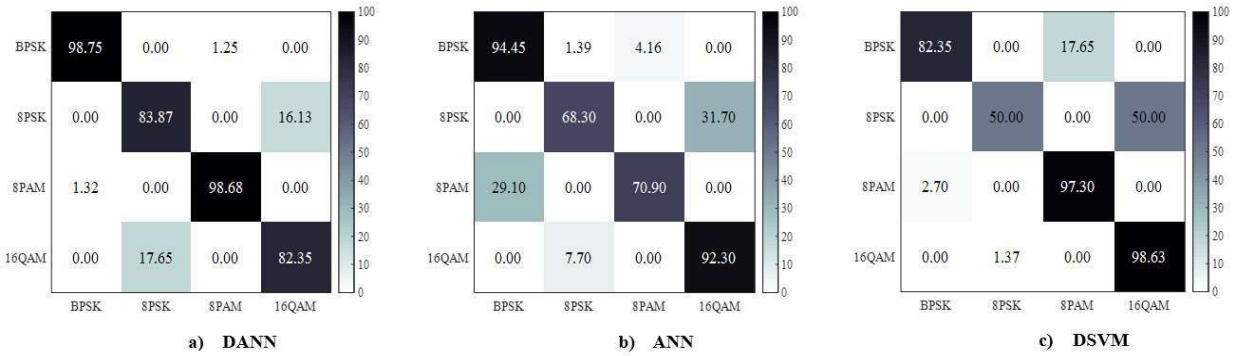

Figure 6: Confusion matrices produced (a) DANN, (b) ANN and (c) DSVM, SNR=0dB, $N = 512$, and $N_r = 2$.

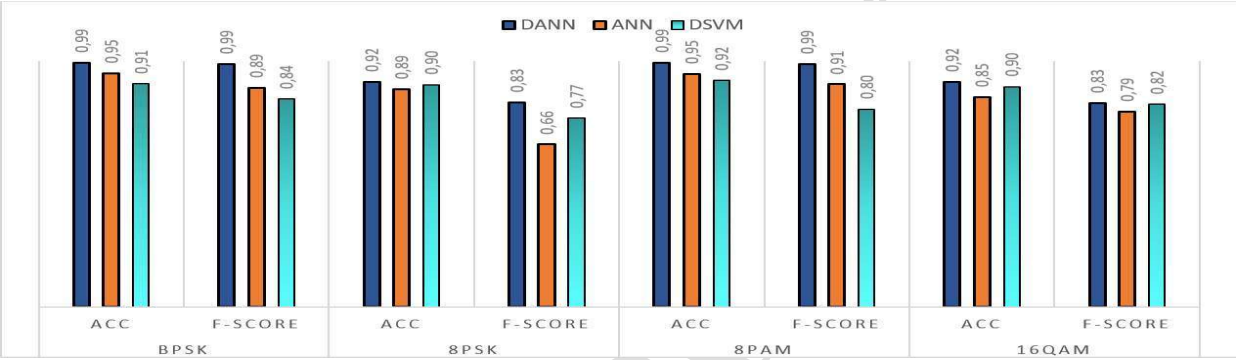

Figure 7: Comparison of the Accuracy and F-score between three methods: ANN, DANN, and DSVM, SNR=0dB, $N = 512$, and $N_r = 2$.

DSVM or standard ANN since they demonstrate lower F-score and ACC performance compared to the DANN. The results confirm the acquired results previously, where we can see lower results for 8PSK and 16QAM modulations for all the classification results. In addition, we can see that the proposed method achieves the highest performance compared to other counterparts.

6.4. Comparison with popular clustering methods for AMI

Table 4 briefly compares our proposed method with some related works.

Despite the underlying assumption that each of the referred papers generates its datasets, we approximate the classifiers mentioned in the referred articles conditions by simulating our proposed model in their appropriate system parameters as presented in Table 4, which can still provide a fair comparison regarding the probability of correct classification. In [11], the authors tackled the issue of AMI in MIMO-STBC systems using low complexity DL and RBFN. However, our proposed method grants a slight improvement in the performance compared to theirs.

Similarly, [7] used DL to recognize the modulation type in MIMO systems based on a convolutional neural network (CNN). They used ZF to recover the received signal; from the results in Table 4, we can see that our proposed method presents a similar performance to theirs.

Other related works [13, 18] generated random bit streams in AWGN channels and modulated them in the referred modulation pool. Then, relying on the modulation constellation, they applied modulation detection. In [13], clustering algorithms were used to recover the constellation and minimum distance classifier to recover the modulation type and lower the execution time, but their method requires a large number of samples and higher SNR values to identify the modulation accurately, whereas in many practical scenarios only a small sample set is available.

However, the authors in [18] addressed the modulation classification by transforming constellations into color images; then, proposed using ResNet-50 and ResNet-V2 to estimate the modulation type and order. Their method achieved minute improvement higher then ours, except at SNR=0dB, our method provides around 12% higher P_{cc}.

A recent work in [24] addressed the AMI in DH transmission system based on AF protocol; however, they did not consider D-STBC. The suggested approach is based on theoretical representations of cross-correlation functions of the received signals. They proved that a family of modulation

Table 4
Comparison the proposed method with popular existing methods.

Method	Modulation pool	Symbols Length	System model	Channel type	P_{cc} (%) at -5dB	0dB	5dB	10dB
DNN [11]	{BPSK, QPSK, 8PSK, 16QAM}	500	STBC-MIMO system	Quasi-static	×	98	100	100
RBFN [11]				Rayleigh	×	96	100	100
Proposed work 'DANN'			DH-AF relaying system	fading	89	99	100	100
ResNet-50 [18]	{2ASK, 4ASK, QPSK, 8PSK, 8QAM, 16QAM, 32QAM, 64QAM}	1000	Randomly generated bit streams	AWGN	32.5	48.25	78.7	99
Proposed work 'DANN'			DH-AF relaying system	Quasi-static Rayleigh fading	32	60	77	98
Cross-Correlation [24]	{BPSK, QPSK, 8PSK, 16QAM}	2048	DH-AF cooperative system (2×1)	Rayleigh fading	×	×	×	67.95
Proposed work 'DANN'			DH-AF relaying system (2×1)	Quasi-static Rayleigh fading	65	96	100	100
CNN/ZF-AMC [7]	{BPSK, QPSK, 8PSK, 16QAM}	128	MIMO system (2×4)	Rayleigh fading	83	100	100	100
Proposed work 'DANN'			DH-AF relaying system (2×4)	Quasi-static Rayleigh fading	83	99	100	100
Clustering [13]	{8PSK, 16QAM}	2048	Randomly generated bit streams	AWGN	0	0	0	89
Proposed work 'DANN'			DH-AF relaying system	Quasi-static Rayleigh fading	65	90	100	100
CCNN [25]	{BPSK, QPSK, 8PSK, 16QAM, 32QAM, 64QAM}	500	Randomly generated bit streams	AWGN	×	78.4	94	99
Proposed work 'DANN'			DH-AF relaying system	Quasi-static Rayleigh fading	35	63	83	92

"×" indicates that P_{cc} at the relative SNR point was note reviewed in the related work

types induces spikes for certain cross-correlation functions while others do not. The P_{cc} was derived through a tree-based method utilizing the false-alarm criterion for spike detection, yet this method requires a large number of samples and high SNR values for accurate detection.

In [25], the authors proposed an AMI based on combining clustering with neural networks. AWGN corrupts the generated signals with specified SNR range. The results show that their method presents higher P_{cc} compared to ours, but one must keep in mind that there are different causes for these results that must be acknowledged. The first reason is the difference in the channel type. We used Rayleigh multi-paths fading channel, whereas in their work, they used AWGN as is the case with [13, 18]. Secondly, the difference in the dataset, we used a DH-AF relaying system while they used randomly generated bit streams.

Figure 8, presents a comparison of training time among different classifiers. One can see that DL is the most computationally expensive classifier (68 sec for training) compared to other classifiers because of its inherent architecture complexity involving deeper layers resulting in additional operations. Therefore, our proposed classifier provides a more practical alternative with lower training time (5.05 sec) and higher performance.

This study strives to investigate the possibility of optimizing modulation classification performance by integrating the advantages of ANN classification (Storing information on the entire network, having a distributed memory, parallel processing capability, etc.) with induced Dendrogram methods.

According to the comparison presented in this work, DSVM presented the lowest training time because it is the simplest model, but it demonstrated the lowest robustness, particularly for one receive antenna, while ANN presented stable results regardless of the antennas configuration and

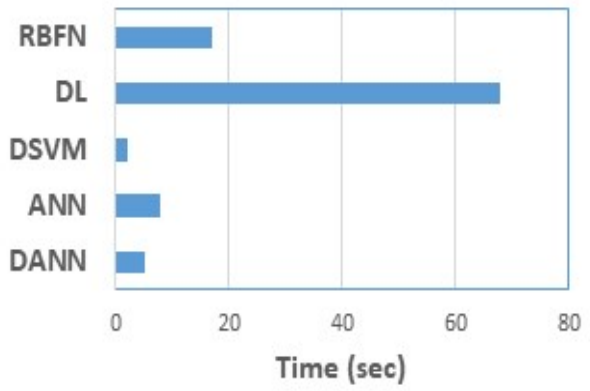

Figure 8: Comparison of the consumed training for proposed framework with the other models.

sample size used for the evaluation. DANN provided higher accuracy and robustness compared to DSVM, ANN, and other classification models in the literature as shown in Table 4, and displayed the lowest training time compared to other models with similar architectures (DL, RBFN, and ANN).

ANN has a higher training time (7.95 sec) than DANN, while DANN reduces the training time by 36.47%.

Generally, improving classification accuracy leads to an increase in training time due to the use of deeper neural networks. However, despite the overall deep architecture of DANN compared to traditional ANN, training time is reduced due to the modularization of the multi-class classification task into smaller classification tasks by inducing a Dendrogram, where traditional ANN is trained on a larger dataset.

7. Conclusion

This paper proposed a new modulation classification method, namely Dendrogram-based Artificial Neural Networks (DANN) by merging ANN with Dendrogram for DH-AF relaying system using DSTBC scheme operating in multipath fading channels.

We compare the performance of the proposed method with other existing methods.

The simulation results have shown that the proposed classification method outperformed the existing methods and presented good performance even when using a single antenna at each operating node in the DH-AF relaying system, responding to the demands in terms of low profile for power and complexity.

The DANN approach is promising in terms of reducing overall training time and improving accuracy as well. Instead of using deeper architectures to extract the subtle differences between schemes. This work creates deeper ANN architecture by separately trained ANN classifiers, whose hierarchy is determined based on the correlation using generated Dendrogram. The overall neural network learns distinct features at higher levels and subsequently learns features with higher similarities at lower levels. In our future work, we will consider including OFDM in DH-AF relaying system; in addition, we will consider the effect of Carrier Frequency Offset (CFO) and channel estimation errors on the classification accuracy.

Acknowledgments

This work was supported by Directorate General for Scientific Research and Technological Development (DGSRTD), Algeria.

References

[1] H. Tayakout, I. Dayoub, K. Ghanem, H. Bousbia-Salah, Automatic modulation classification for d-stbc cooperative relaying networks, IEEE Wireless Communications Letters 7 (5) (2018) 780–783. doi:http://dx.doi.org/10.1109/LWC.2018.2824813.

[2] T. Morehouse, N. Rahimi, M. Shao, R. Zhou, Baseband modulation classification using incremental learning, in: 2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS), IEEE, 2020, pp. 225–228. doi:http://dx.doi.org/10.1109/MWSCAS48704.2020.9184646.

[3] J. Rodriguez, Fundamentals of 5G mobile networks, John Wiley & Sons, 2015. doi:http://dx.doi.org/10.1002/9781118867464.

[4] V. D. Orlic, M. L. Dukic, Automatic modulation classification: Sixth-order cumulant features as a solution for real-world challenges, in: 2012 20th Telecommunications Forum (TELFOR), IEEE, 2012, pp. 392–399. doi:https://doi.org/10.1109/TELFOR.2012.6419230.

[5] B. Jdid, W. H. Lim, I. Dayoub, K. Hassan, M. R. B. M. Juhari, Robust automatic modulation recognition through joint contribution of hand-crafted and contextual features, IEEE Access 9 (2021) 104530–104546. doi:http://dx.doi.org/10.1109/ACCESS.2021.3099222.

[6] H. Zhang, Y. Wang, L. Xu, T. A. Gulliver, C. Cao, Automatic modulation classification using a deep multi-stream neural network, IEEE Access 8 (2020) 43888–43897. doi:http://dx.doi.org/10.1109/ACCESS.2020.2971698.

[7] Y. Wang, J. Gui, Y. Yin, J. Wang, J. Sun, G. Gui, H. Gacanin, H. Sari, F. Adachi, Automatic modulation classification for mimo systems via deep learning and zero-forcing equalization, IEEE Transactions on Vehicular Technology 69 (5) (2020) 5688–5692. doi:https://doi.org/10.1109/TVT.2020.2981995.

[8] A. K. Ali, E. Erçelebi, Automatic modulation classification using different neural network and pca combinations, Expert Systems with Applications 178 (2021) 114931. doi:http://dx.doi.org/10.1016/j.eswa.2021.114931.

[9] B. Zhang, S. Qi, P. Monkam, C. Li, F. Yang, Y.-D. Yao, W. Qian, Ensemble learners of multiple deep cnns for pulmonary nodules classification using ct images, IEEE Access 7 (2019) 110358–110371. doi:http://dx.doi.org/10.1109/ACCESS.2019.2933670.

[10] P. Zhang, J. Zhang, K. P. Peppas, D. W. K. Ng, B. Ai, Dual-hop relaying communications over fisher-snedecor f-fading channels, IEEE Transactions on Communications 68 (5) (2020) 2695–2710. doi:https://doi.org/10.1109/TCOMM.2020.2973263.

[11] M. H. Shah, X. Dang, Low-complexity deep learning and rbfn architectures for modulation classification of space-time block-code (stbc)-mimo systems, Digital Signal Processing 99 (2020) 102656.

[12] M. H. Shah, X. Dang, An effective approach for low-complexity maximum likelihood based automatic modulation classification of stbc-mimo systems, Frontiers of Information Technology Electronic Engineering 21 (3) (2020) 465–475.

[13] J. P. Mouton, M. Ferreira, A. S. Helberg, A comparison of clustering algorithms for automatic modulation classification, Expert Systems with Applications 151 (2020) 113317. doi:http://dx.doi.org/10.1016/j.eswa.2020.113317.

[14] M. Saber, A. El Rharras, R. Saadane, A. Chehri, N. Hakem, H. Kharraz, Spectrum sensing for smart embedded devices in cognitive networks using machine learning algorithms, Procedia Computer Science 176 (2020) 2404–2413.

[15] H. YİĞİT, A. Kavak, A learning approach in link adaptation for mimo-ofdm systems, Turkish Journal of Electrical Engineering and Computer Sciences 21 (5) (2013) 1465–1478. doi:http://dx.doi.org/10.3906/elk-1110-24.

[16] J. Xu, C. Luo, G. Parr, Y. Luo, A spatiotemporal multi-channel learning framework for automatic modulation recognition, IEEE Wireless Communications Letters 9 (10) (2020) 1629–1632.

[17] N. Jafar, A. Paeiz, A. Farzaneh, Automatic modulation classification using modulation fingerprint extraction, Journal of Systems Engineering and Electronics 32 (4) (2021) 799–810.

[18] Y. Kumar, M. Sheoran, G. Jajoo, S. K. Yadav, Automatic modulation classification based on constellation density using deep learning, IEEE Communications Letters 24 (6) (2020) 1275–1278.

[19] B. Dehri, M. Besseghier, A. B. Djebbar, I. Dayoub, Blind digital modulation classification for stbc-ofdm system in presence of cfo and channels estimation errors, IET Communications 13 (17) (2019) 2827–2833. doi:http://dx.doi.org/10.1049/iet-com.2019.0362.

[20] K. Hassan, I. Dayoub, W. Hamouda, C. N. Nzeza, M. Berbineau, Blind digital modulation identification for spatially-correlated mimo systems, IEEE Transactions on Wireless Communications 11 (2) (2011) 683–693. doi:https://doi.org/10.1109/TWC.2011.122211.110236.

[21] O. A. Dobre, Y. Bar-Ness, W. Su, Higher-order cyclic cumulants for high order modulation classification, in: IEEE Military Communications Conference, 2003. MILCOM 2003., Vol. 1, IEEE, 2003, pp. 112–117. doi:http://dx.doi.org/10.1109/MILCOM.2003.1290087.

[22] S. Huang, Y. Yao, Z. Wei, Z. Feng, P. Zhang, Automatic modulation classification of overlapped sources using multiple cumulants, IEEE Transactions on Vehicular Technology 66 (7) (2016) 6089–6101. doi:http://dx.doi.org/10.1109/TVT.2016.2636324.

[23] T. Lajnef, S. Chaibi, P. Ruby, P.-E. Aguera, J.-B. Eichenlaub, M. Samet, A. Kachouri, K. Jerbi, Learning machines and sleeping brains: automatic sleep stage classification using decision-tree multi-class support vector machines, Journal of neuroscience methods 250 (2015) 94–105. doi:https://doi.org/10.1016/j.jneumeth.2015.01.022.

[24] M. Marey, H. Mostafa, Modulation awareness method for dual-hop cooperative transmissions over frequency-selective channels, Sensors 22 (14) (2022) 5441.

[25] A.-S. Liu, Z. Qi, Automatic modulation classification based on the combination of clustering and neural network, The Journal of China Universities of Posts and Telecommunications 18 (4) (2011) 13–38. doi:http://dx.doi.org/10.1016/S1005-8885(10)60077-5.