A Novel Method for Pattern Recognition based on Radar Tomographic Images

Muhannad Almutiry[1] and Ahmed Alsheikhy[2]

Electrical Engineering Department, Engineering College, Northern Border University, Arar 73222 Saudi Arabia

Summary
Radar tomographic imaging's foundation is the radar cross-section (RCS) of the pattern and material of the investigative shape. RCS varies when the target's permittivity and conductivity differ in material profile and shape from other objects in the environment. This paper investigates the use of Multi-Layer Perceptron (MLP) for pattern recognition of shapes in the frequency domain. Supervised and Unsupervised Locally Linear Embedding (LLE) algorithms were used as a robust means of addressing the difficulties of nonlinear dimensionality reduction of scattering fields when domains have frequency complexity. The proposed algorithm is capable of recognizing a variety of shape profiles based on the scattering field measured by using a geometrically diverse tomographic radar scanning transmitter/receiver set up. The algorithm is trained in recognition of the patterns of different scattering fields with a scattering field of radar tomographic handwritten digits across the frequency domain, which can be extended to the 3D model. The theoretical basis of this work is validated using the results of simulations.

Key words:
MLP, RF Tomography, Radar Target Recognition, LLE.

.1. Introduction

Research into the employment of neural networks for geometry recognition recently has gained considerable attention for the radar system [1]–[4]. The geometry recognition of radar target represents an inverse problem involving harvesting information regarding the scatterer by analyzing the scattered field response that was discussed in the literature[5]–[8]. The identification of radar targets may occur either through the creation of images with a high enough resolution that human observers can recognize them or by the creation of signatures/target representations that can be automatically recognized by machines [9]–[12]. Tomographic Radar Imaging technology combining angular, spectrum, and polarization degrees of freedom can produce imagery of a target's scattering centers close to optical resolution. Despite this ability, sometimes it can be too expensive or too cumbersome to create the physical aperture required to provide the requisite angular degrees of freedom; at other times their simplicity is not the time to synthesize such apertures by a relative motion between the object being investigated and the radar system (e.g., using SAR or ISAR). This then creates the difficulty of identifying a target when the available information is insufficient for the creation of recognizable imagery. Several authors have recognized neural networks for target recognition in time domain radar signals which suffer from high complexity and noise generated through the process [13], [14]. Using a natural frequency-based neural network as an independent angle was proposed, which has less noise immunity. The work at [15], used principal-component analysis of the resonance of the radar signals to feed the multi-layer perceptron neural network in order to classify the target geometry where the generalized pencil-of-functions method was used to process the data again in the time-domain in order for noise reduction. The method of dimensionality reduced space introduced by [16] has the advantage of increase the classification performance of the high-range resolution profile (HRRP) scattering. The HRRP suffers from a lack of information where the neural network feature extraction will typically just process a dominant scatterer data in the scene.

For data dimension reduction, Locally Linear Embedding (LLE) algorithms are regarded as a robust tool for addressing the difficulties of nonlinear dimensionality reduction [17]. Dimensionality reduction relates to the difficulties of data mapping onto a low-dimensional embedding space when it is on or close to a low-dimensional manifold in high-dimensional data space. Many standard techniques have been developed for linear dimensionality reduction [18]–[21], e.g., Multidimensional Scaling and Principal Component Analysis (PCA) [22]. Nevertheless, such methodologies do not work well when embedding nonlinear manifolds [23]. LLE, an unsupervised learning methodology [24], is being suggested as a practical algorithm to address the difficulties of dimensionality reduction, as it can obtain low dimensional and environmental data while not rejecting embeddings of high dimensional data. It has a number of advantages, including the speed of implementation and its ability to absorb fresh data without the need to run the whole algorithm over again.

We shall demonstrate that collective nonlinear signal processing reduction LLE centered on MLP models in combination with the employment of appropriate target signatures provides the potential for reliable very high-resolution target identification even when information is incomplete. We introduce a neuromorphic means of approaching target identification that can remove the

Manuscript received October 5, 2021
Manuscript revised October 20, 2021
https://doi.org/10.22937/IJCSNS.2021.21.10.1

requirement for sizable expensive apertures required to image remote targets. This research further proposes that nonlinear multidimensional dynamical systems could offer a means of identifying targets using a single operating frequency.

Generally, approaching the radar target identification thought couple methods. One is to form a radar image that will then be examined and identified by human observation or an artificial intelligence system. In this case, system designers investigate the methodologies and concepts that can produce the highest possible resolution images to make human identification as accurate as possible. The twin aims of this type of method are generally economic viability and close to optical resolution. The second way of approaching the problem is to have the target automatically recognized by a machine that can identify appropriate representations of target signatures. This is the optimal method when there is insufficient information regarding the Radar Cross Section (RCS) of different sections of an object so that it can be defined. This method concerns how accurate identification can be made when information is incomplete or thin no matter what the target range, aspect, or location by using a system that can robustly accomplish this that can deal with any faults. For this type of approach, the processing undertaken by the artificial intelligence system to identify images in the alternative approach has to be replicated by machinery. There are a number of reasons for selecting automated recognition, amongst the most important of which are speed and economy. Both of these approaches require amplitude/phase measurement for radar echoes from irregular and complex objects using the functions of polarization, frequency, and orientation by employing the same equipment generally used to make intricate RCS measurements.

This paper has designed a classification system of tomographic measurement domain that represented as the handwritten digit's frequency response (0, 1,……, 9) from the MNIST database employing dimensionality reduction based on supervised LLE and unsupervised LLE classification techniques based on MLP. The frequency response data of the measurement domain contain multiple data such as amplitude and phase for each step frequency in the bandwidth beside the spatial information.

2. Methodology

This paper will investigate the second of the approaches detailed in the previous section and demonstrate the links between the processes and how to target representation. The processes of radar imaging must be understood in order to formulate a methodology to identify targets automatically. The motivation of this research is to improve nonlinear dimensionality reduction through the employment of LLE for radar data. An additional purpose is to reduce the potential costs of any radar diversity imaging system. The intention is to be able to develop an automated recognition using incomplete information, particularly when there is so little target information available that it is practically impossible to formulate a clear image. The focus on neural signal processing is natural, taking into account the associative memory attributes for artificial intelligence systems, how it can complete or supplement any absent information, and the seemingly effortless and swift way in which it can resolve untidy problems frequently found within speech, vision, and cognition processes. Neural processing offers a novel and robust means of processing signals that can easily cope with flawed information.

When using artificial neural networks, machine recognition is reliant on generating target signatures (target feature/attribute representations) that can provide forms of recognition that are "distortion tolerant," i.e., in which the range, orientation, or location of the target (otherwise scale, rotation, and shift-invariant recognition) are not handicapping factors. This paper presents how artificial neural networks can take scanty information regarding complex scattering objects and develop recognition of extremely high resolution.

Radar Tomography looks at targets from numerous angles following the way the transmitters and receivers are spatially distributed. Having a diverse geometric distribution leads to an increase in the amount of information that can be obtained through measurement. The environment that contains the targets requiring imaging is surrounded by multiple (*N*) dipole transmitters and *M* dipole receivers. The location of the transmitter *n* is $\boldsymbol{r_n^t}$ with polarization $\hat{\mathbf{a}}_{\mathbf{n}}^{\mathbf{t}}$; for receiver *m* the location is $\boldsymbol{r_m^r}$ and polarization $\hat{\mathbf{a}}_{\mathbf{m}}^{\mathbf{r}}$. This arrangement is illustrated in Figure 1, which shows the area to be measured and a single pair of transmitters and receivers. Only one transmitter emits a known waveform at a time, with the other transmitters remaining inactive. The target echoes are harvested by the receivers [25].

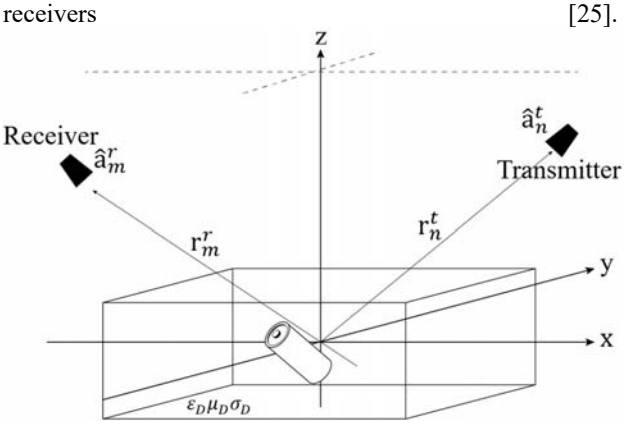

Fig. 1. 3-D model of the target of interest

By employing this geometry, it is possible to introduce elaborations to basic/inverse models in time-harmonic electric fields. It should be noted that time-domain waveforms may be given expression within time-harmonic fields employing stretch processing or simply the Fourier transform.

The assumptions made about the measurement domain is a free space medium having an unknown contrast function $\tau_\delta(r')$ in which r' is a position vector. We can express the scattered field $\mathbf{E}^S(\boldsymbol{r}_n^t, \boldsymbol{r}_m^r)$ for transmitter n and receiver m by employing the Born approximation, as shown below,

$$\mathbf{E}^S(\boldsymbol{r}_n^t, \boldsymbol{r}_m^r) = k_0^2 \iiint_D [\hat{\mathbf{a}}_\mathbf{m}^\mathbf{r} \cdot \bar{\mathbf{G}}(\boldsymbol{r}_m^r, \mathbf{r}')] \cdot [\bar{\mathbf{G}}(\mathbf{r}', \boldsymbol{r}_n^t) \cdot \hat{\mathbf{a}}_\mathbf{n}^\mathbf{t}] \tau_\delta(\mathbf{r}') d\mathbf{r}' \quad (1)$$

For the equation above, k is the wavenumber, and $\bar{\mathbf{G}}(\boldsymbol{r}_m^r, \mathbf{r}')$ is the Green's function. The Born approximation has a linear relationship with the contrast function, and the above equation may be represented as matrix multiplication. In this instance, we may express the relationship as[25]:

$$\mathbf{E}^S(\boldsymbol{r}_n^t, \boldsymbol{r}_m^r) = \mathbf{L}(\tau_\delta(\mathbf{r}')) \quad (2)$$

in which the forward model can be implemented through multiplication by the L matrix. The $\boldsymbol{L}$ operator must undergo inversion for computation of the scattering field. For the general case, we just need to obtain the scattering field from tomographic scanning mode to feed it into the low reduction algorithm.

3. Algorithm Description

Numerous machine learning problems start when raw multidimensional data sets have to be preprocessed, e.g., spectral histograms, speech signals, objects, etc. The aim of preprocessing the scattering field of radar tomographic data is obtaining data representations for the operations that follow that are more descriptive, informative, and useful, such as outlier detection, clustering, visualization, classification, etc. One purpose of data preprocessing is to reduce dimensionality. High dimensional data may contain numerous correlations and redundancies that conceal fundamental relationships; preprocessing aims to remove any redundancies in the processed data. There are two ways in which data dimensionality may be reduced, either supervised, in which data labels are supplied and unsupervised, without data labels. In the majority of practical instances, there will be no prior data knowledge available, as assigning labels to data samples is both expensive and time-heavy.

A. Unsupervised Locally Linear Embedding (LLE)

Let the scattering field vector X be a set of N points a high-dimensional data space R^D as the following

$$X = \{x_1, x_2, \dots x_N\} \quad (3)$$

The scattering field as the frequency response of the handwritten digits is a nonlinear manifold of intrinsic dimensionality d < D (typically d <<D). D is received scattering field dimension at a different tomographic angle, which can be extended into the received scattering field at a set of bandwidth. We create a sufficient frequency response sample from the manifold to find a low reduction of the frequency response sample using LLE by mapping the multidimensional data (D) into one global coordinate system in R^d. Let the frequency response sample has N samples in the embedding space R^D by

$$Y = \{y_1, y_2, \dots y_N\} \quad (4)$$

Each pattern is $D \times 1$ vector. In our case $D = 841$, M is the number of classes and N is the number of patterns

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_D \end{bmatrix} \quad (5)$$

1) Proximity matrix

Identify the K nearest neighbors using Euclidean distance in the D-dimensional space for each data point x_i. For each data point $x_i \epsilon X$, find the set N_i of K nearest neighbors of x_i.

$$Distance = \|X_i - X_j\|^2 \quad (6)$$

The dimension of the proximity matrix is $k \times N$. Where the first row is the indices of all the neighbors of the first pattern, the second row is the indices of all the neighbors of the second pattern and so on.

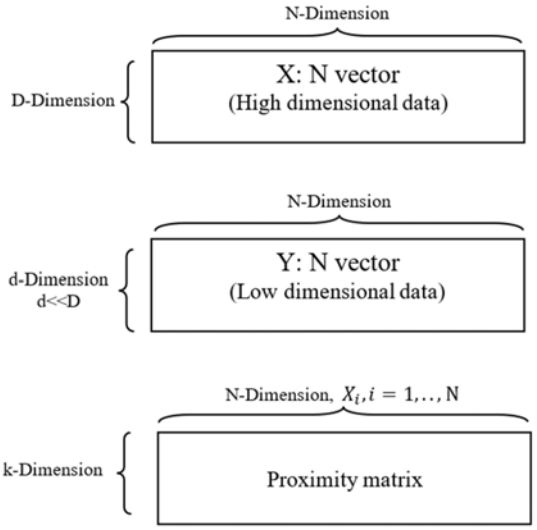

Fig. 2. The dimensionality reduction of unsupervised LLE.

2) Compute the weights W_{ij}

Define each data point X_i as a weighted combination of all its neighbors.

$$\widetilde{X}_i = \sum_{j=1}^{N} W_{ij}\, X_j \tag{7}$$

$W_{ij} = 0$ if X_i and X_j are NOT neighbors.

Where W_{ij} is a square matrix with $N \times N$ dimensions.

$$W_{ij} = \begin{bmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{N1} & \cdots & w_{NN} \end{bmatrix} \tag{8}$$

The error can be calculated using the following equation

$$E(W_i) = \left\| X_i - \sum_{j=1}^{N} W_{ij}\, X_j \right\|^2 \tag{9}$$

The total error

$$E(W) = \sum_{i=1}^{N} \left\| X_i - \sum_{j=1}^{N} W_{ij}\, X_j \right\|^2 \tag{10}$$

This error should be minimized to find the best reconstruction weights W_{ij}. Subject to $\sum_{j=1}^{N} W_{ij} = 1$ to have rotation, scaling translation invariance (sum to one condition).

By using the Raleigh-Ritz theorem, we can solve the system of k^2 equations with variables k^2

$$E(W_i) = \left\| X_i - \sum_{j=1}^{N} W_{ij}\, X_j \right\|^2 = \left\| \sum_{j=1}^{N} W_{ij}\, (X_i - X_j) \right\|^2 \tag{11}$$

$$E(W_i) = \sum_{j=1}^{N} \sum_{k=1}^{N} W_{ij}\, W_{ik} (X_i - X_j)\,(X_i - X_k) \tag{12}$$

$$\sum_{k=1}^{N} W_{ik}\, (X_i - X_j)\,(X_i - X_k) = 1, \quad for\ all\ i\ and\ j \tag{13}$$

System of k^2 linear equations with k^2 variables.

$$g_{jk} = (X_i - X_j)(X_i - X_k) \tag{14}$$

$$\sum_{k=1}^{N} W_{ik}\ g_{jk} = 1 \qquad for\ all\ i, j = 1,2, \ldots N \tag{15}$$

3) Compute the d dimensional embeddings

Compute the d dimensional embeddings best reconstructed based on the weight's matrix W_{ij} obtained. This corresponds to minimizing the following function:

$$\Phi(Y_i) = \left\| Y_i - \sum_{j=1}^{N} W_{ij}\, Y_j \right\|^2 \tag{16}$$

We create another set of data Y_i using the same weight matrix W_{ij}.

Y_i is the low dimensional representation of X_i and it is a d-dimension vector, and X_i is the D-dimension vector. The total error can be calculated through the following:

$$\Phi(Y) = \sum_{i=1}^{N} \left\| Y_i - \sum_{j=1}^{N} W_{ij}\, Y_j \right\|^2 \tag{17}$$

(17) is an ill-posed problem, in order to create the well-posed problem [26]:

$$\sum_{i=1}^{N} y_i = 0 \tag{18}$$

$$\frac{1}{N} \sum_{i=1}^{N} y_i\, y_i^T = I \tag{19}$$

y_i is reconstructed input vector with d-dimension, and I is an identity matrix with $d \times d$ dimension. Furthermore, the weight matrix W_{ij} is updating the M matrix:

$$M = (I - W)^T (I - W) \tag{20}$$

Where I, W, and M are three matrices with $N \times N$ dimension. We apply the Eigenvalue analysis on matrix M to choose d Eigenvectors corresponding to the lowest d non-zero Eigenvalues. Usually, the value of d is equal to the number of classes minus one.

B. For testing:

1) We have the test pattern $X_{test}, where\ t$he X_{test} is D-dimension

$$X_{test} = \begin{bmatrix} X_1 \\ \vdots \\ X_D \end{bmatrix} \quad (21)$$

2) Find the k-nearest neighbors.
3) Solve the following system of linear equations

$$\sum_{k=1}^{N} W_{test,k}\,(X_{test} - X_j)\,(X_{test} - X_k) = 1 \quad for\ j = 1,2,\dots N \quad (22)$$

$$\sum_{k=1}^{N} W_{test,k}\ g_{jk} = 1 \quad (23)$$

X_j and X_k are the data points in the training set which are k neighbors of X_{test}.

$W_{test,j} = 0$ if X_{test} and X_j are NOT neighbors.

and $\sum_{j=1}^{N} W_{test,j} = 1$ sum to one.
By solving the system of equations, we get the matrix $W_{test,j}, j = 1,2,\dots,N$

4) Compute Y_{test}

$$Y_{test} = \sum_{j=1}^{N} W_{test,j}\, Y_j \quad (24)$$

Y_{test} and Y_j is d-dimension. Y_j's are the reduced dimensional representation of X_j (k neighbors of X_{test}).

C. *Supervised Locally Linear Embedding (SLLE)*

The main disadvantage of the LLE is complexity. In the SLLE, we use class information instead of using all the data point as one class (in LLE).

1) Proximity matrix

Identify the K nearest neighbors using Euclidean distance in the D-dimensional space for each data point within the class (the neighborhood of a data point only from those points that belong to the same class)

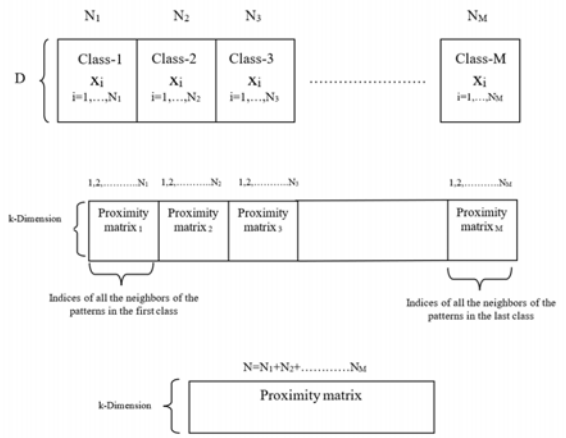

Fig. 3. Dimensionality reduction of supervised LLE.

Wherer m[th] class A matrix should be shifted by $\sum_{i=1}^{m-1} N_i$

2) The same procedure in the LLE to compute the weighted matrix W_{ij}.

W_{ij} is $N \times N$ dimension.

3) The same procedure in the LLE to compute another set of data Y_i

Y_i is the low dimensional representation of X_i as d-dimension. (d<<D). For the testing, the part is the same as LLE.

4. Classification using a Multi-Layer Perceptron

Train a Multi-layer perceptron network to classify the low dimensional data to their respective
classes (the input data is Y_i with d-dimension, and the output is ten classes), as shown in Fig. *4*.

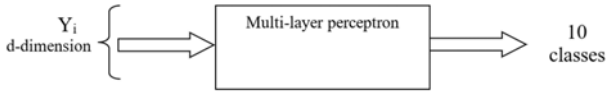

Fig. 4. Multi-layer Perceptron Block diagram processing low dimension tomographic radar data.

For those working with high-dimensional data, dimensionality reduction is a significant means of allowing us to discover the essential relationships between data points and the primary modes that vary in the data. In dimensionality reduction procedures, high dimension data is input, with an output of a map of the data on a low-dimensional manifold. Neighborhoods should be preserved

in dimensionality reduction, i.e., if points are adjacent in high dimensional space, they should still be adjacent in the calculated low dimensional space. Also, dimensionality reduction methods ought to provide for nonlinear low-dimensional manifolds. The essential assumption of LLE is that even when we consider the whole of a manifold embedded in high dimensional space as being nonlinear, we may still assume it to be locally linear if all its data points and their neighbors are situated on or close to a patch of the manifold that is locally linear. All data points may be represented through the weighted linear combination of the closest neighbors. For this research, local linear embedding is employed for the reduction of the high dimensional data set into a suitable low dimensional data set. Firstly, we employ supervised LLE, and then unsupervised LLE.

5. Simulation Results

For the Tomographic radar imaging response, we calculated the scattering field of the measurement domain, which is, in our case, the handwritten digits to provide more diversity in shape looking. The handwritten images, as shown in Fig. *6*, is divided into 28×28 pixels as experimental measurement domain. Each pixel is $0.0037 \times 0.0037\ m$. In order to apply the Forward problem to obtain the theoretical scattering field of the measurement domain, we calculated the Green's function of the in the measurement domain. We design the simulation to have the dimension of the 2D measurement domain in $0.105 \times 0.105\ m$. We operate the simulation at a single frequency was sat up to be $8\ GHz$ ($\lambda = 0.0375\ m$) to function the scanning system. We used 28 transmitters and 28 receivers that placed equally spaced circularly. The transmitters are located $2\ m$ away from the center of the measurement domain, while the receivers are located $0.4\ m$ away from the center as shown in Fig. *5*. The scattering field matrix is 29×29, which leads to D a vector of 841 complex elements that represent the scattering field in the frequency domain for only a single frequency was obtain by the forward problem. The scattering field power density of the measurement domain is shown in Fig. *7* where we can see the varying due to the change of the handwritten digit shape. The phase of the scattering field and the Eigen-function of the measurement domain is shown in Fig. *8* and Fig. *9*, respectively.

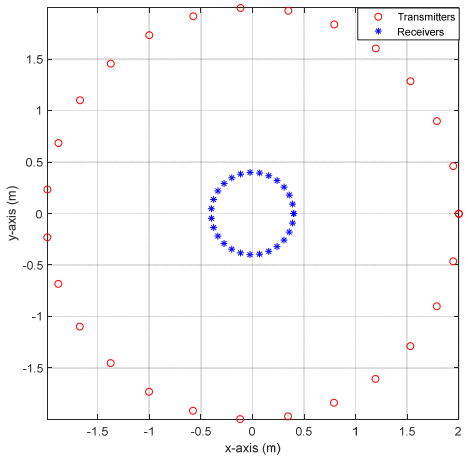

Fig. 5. The location of the transmitters and the receivers.

The training and testing sets are selected and rearranged before the inputs are applied (X_i, i=1, 2… 841). For the training, 1000 scattering field of handwritten digit images is chosen (100 images per digit). In this process, the first 100 images called with the digit "0", the second 100 images accord with the digit "1", and so on. In addition, 100 images are employed for testing (10 images for each digit), and the testing image is rearranged so that the initial 10 images are for the digit "0" and the second 10 images for the digit "1". The input data sets are then normalized. For this research, we use D=841 (input pattern dimension), N=1000 (number of data point (training set number)), M=10 (number of classes), d=M-1=9 (low dimension sets dimension (Y_i)), and k=18 (numbers of neighbors).

A. Unsupervised Locally Linear Embedding (LLE)

The initial stage in using the LLE algorithm is to discover the neighborhood for all data points X_i, i=1,2,....N. This may be achieved either through the identification of a set number of nearest neighbors K for each data point in relation to the Euclidean distance or by selecting every point inside a fixed-radius sphere. For this research, the neighborhood is identified using Euclidean distance. Thus for every data point, we have N-1 Euclidean distance. A proximity matrix is then constructed employing every neighbor for every data input. The proximity matrix measures $k \times N$ (18×1000). We will employ this matrix in the second stage.

Once we have found K nearest neighbors for every data point and defined each data point Xi as being a weighted combination of every neighbor, the next stage is to find an assigned weight for each pair of neighboring points. The

weights comprise a weight matrix W_{ij}, i,j=1,2,…,N, with every element (W_{ij}) characterizing the degree of closeness for a pair of specific points (x_i and x_j). Naturally, the weighted matrix is a square matrix having dimensions ($N \times N$)(in this instance 1000×1000). This means that apart from the elements X_j, neighbors of X_i, the majority of elements in the matrix are zeros (non-active). For every row of the weight matrix, there is a k number of $W_{ij} \neq 0$ and k-N number of W_{ij}=0. Besides, at this stage, we have the condition (sum to one) referred to in the second stage that requires satisfaction for the provision of rotation, scaling, and translation invariance. We can obtain the weighted matrix through the resolution of a system of linear equations.

The final stage is the computation of the low dimensional embedding Y_i. Employing the weighted matrix W_{ij}, the computation of the embedded coordinates Y_i is achieved through the minimization of the function $\Phi(Y)$ (referred to in step 3). At this stage, two conditions need to be satisfied ($\sum_{i=1}^{N} y_i = 0$ and $\frac{1}{N} \sum_{i=1}^{N} y_i y_i^T = I$). We must then compute the matrix M employing the formula for stage 3. The M matrix represents a square matrix having dimensions $N \times N$ (1000×1000 in this instance). We can now apply the eigenvalue analysis to this matrix. d=9 eigenvectors are selected that correspond to the lowest non zero eigenvalues. Having completed this step, we arrive at another set of data Y_i (with low dimension). The Y dimension is $d \times N$ (9×1000 in this instance).

For the element of testing, we first need to find the k nearest neighbors and then resolve the systemic linear equation for computation of $W_{test,j}$ with two condition constraints $W_{test,j}$ =0 if X_{test} and X_j are NOT neighbors, and $\sum_{j=1}^{N} W_{test,j} = 1$ sum to one), then calculate the Y_{test} as illustrated in the testing section for the algorithm.

B. Supervised Locally Linear Embedding (SLLE)

About the supervision of LLE, this approach employs class information. Rather than treating every data point as a single class and revealing the k nearest neighbors, we search for each class's k nearest neighbors. We then obtain the proximity matrix cab through a concatenation of every proximity matrix in every class. The same procedure as was used for the unsupervised LLE algorithm is employed to find the weighted matrix and to compute the low dimensional embedding Y_i.

Generally, looking at the last tables and bar charts, learning conversion characteristics and testing performance for MLP are dependent on numerous elements. Reducing dimensionality can make the system less complex. Low dimensional data sets must retain the crucial information that was held in high dimensional data sets. The advantages of dimensionality reduction are frequently employed to obtain compact data representations before the application of classifiers. In this instance, the chief aim is the obtaining of low dimensional data representations with good class separability. In this paper, we employed the LLE approach for compacting all the available data into low dimensional data sets rather than high dimensional ones. Dimensional reduction employing LLE goes through three central stages: discover the k nearest neighbors for every point X_i in the original space, undertake computation of the weighted matrix W_{ij} in order to achieve the optimal reconstruction of each point X_i from its neighbors in accordance with the requirement to keep rotation, scaling, and translation invariant, and finally the d-dimensional embedding is calculated using the best reconstruction with the weighted W_{ij}. Following these three stages, the new data sets are obtained, these being a low dimensional representation of the original high dimensional data sets. To test these, we map test data sets onto an identical low dimensional space, which is the unsupervised LLE. It is called unsupervised because there is no information available regarding the data, e.g., labels, target output, etc., and so data processing must be undertaken using a methodology that does not require label information to be inputted. With supervised LLE, the k nearest neighbors are found using previous class knowledge. With the MLP in LLE and SLLE, the conversion characteristics move towards zero error, as illustrated in Fig. *10* and Fig. *11*. It is conjectured that the most likely cause of this is that rather than working with high dimensional data that generally carries redundant information (a.k.a. noise), we are working with low dimension data that comprises almost exclusively relevant information. Naturally, extraneous or redundant information adds to the system's complexity. This is why training of the MLP is undertaken with minor fluctuations. To achieve good data separation, the embedded space's dimensionality, d, should be lower than the number of classes, M, by one. As we have ten classes, they must be set at nine. The parameter k has no significant impact on SLLE's ability to achieve good data separation. By good separation, we mean that every point from the identical class in the high dimensional space can be mapped onto a single point in the embedded space. We are aware that LLE has also been named a piecewise approach, meaning that the complete nonlinear manifold is divided into sections of the linear hyperplane. Should the k value be increased, the number of pieces increases as well, and more complexity is introduced to the system. Alternatively, there is a linear hyperplane.

A balance has to be struck between system performance and the number of nearest neighborhoods. Fig. *10* and Fig. *11* show the testing performed for the handwritten digit recognition system employing LLE and SLLE. Looking at these results, it is clear that for either case, the system is performing acceptably. With SLLE, the system performance was slightly superior to that with LLE. For LLE, test results varied between 50% and 70%, while with SLLE, the variation was between 73% and 90%. It is the nature of the original data sets that prevent the percentages

reaching more than 90%. The system performance could be improved by segmenting the patterns and cropping the handwritten digits (handwritten digital dictation) at the center of the patterns. Additionally, there is some similarity between the shapes of the patterns, e.g., numbers 3 and 8 have similarities, as do numbers 9 and 6.

From simulation results, it is obvious in table 1 and table 2 that SLLE give much better results than LLE because in SLLE the comparison between the test image and the other training images will be from the same family; thus the nearest neighbors is much more accurate than if the comparison is made with different digits i.e. the error probability is much less in the case of SLLE than LLE. Moreover, the number of iteration and time needed for SLLE is much less than the LLE case. This, due to the same reason, i.e., calculating the nearest distance from the same family of the testing image, takes less time and gives more accurate results. Furthermore, the Error percentage for each did digit is varying with changing the settings of the input images; this is due to the shape of the image itself, for specific images the number's picture has a big difference if we compared it to the other numbers of the same family, so the results give higher error and vice versa. Also, the Error percentage for each digit and the total error are changed for each execution, this because of varying the initial weighting matrix.

TABLE I
DISPLAYS ERROR PERCENTAGE FOR EACH DIGIT FOR BOTH LLE & SLLE

Digits	LEE ERROR %	SLLE ERROR %
0	12.9032	3.2258
1	29.7297	5.4054
2	57.1429	28.5714
3	56	32
4	62.963	37.037
5	64.2857	25
6	48.2759	17.2414
7	37.931	17.2414
8	70.9677	32.2581
9	51.3514	27.027

TABLE 2
THE TOTAL ERROR, NUMBER OF ITERATIONS, AND REQUIRED TIME OF LLE AND SLLE

	LEE	SLLE
Total Error %	48.3444	21.8543
No of iteration	610	5.4054
Required time (sec)	235.811385	252.760857

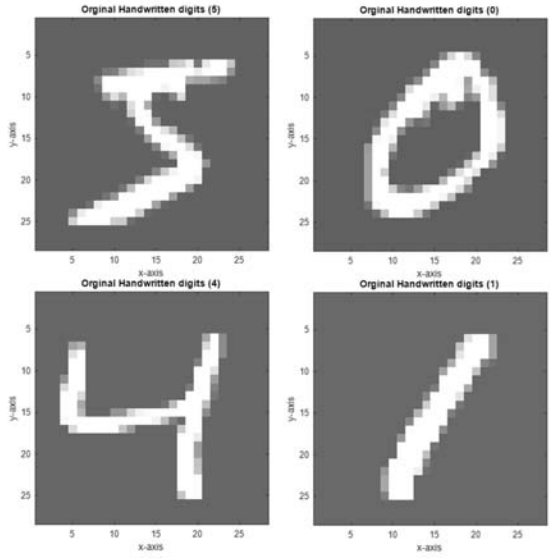

Fig. 6. The handwritten digit images examples from the MNIST data set.

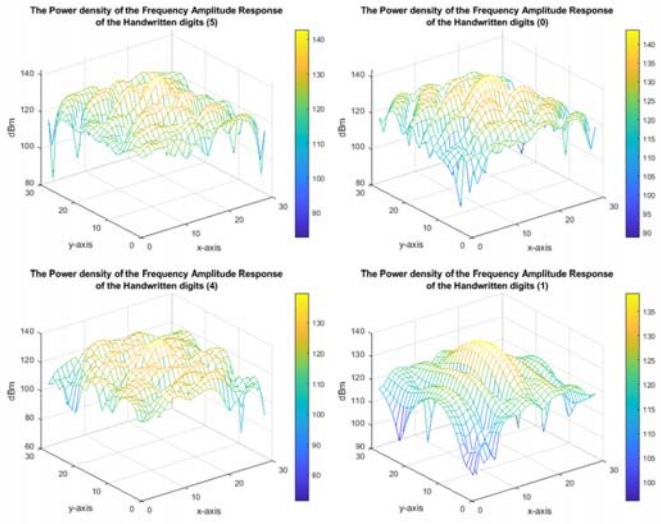

Fig. 7. The Power density of the measurement domain for various handwritten digits.

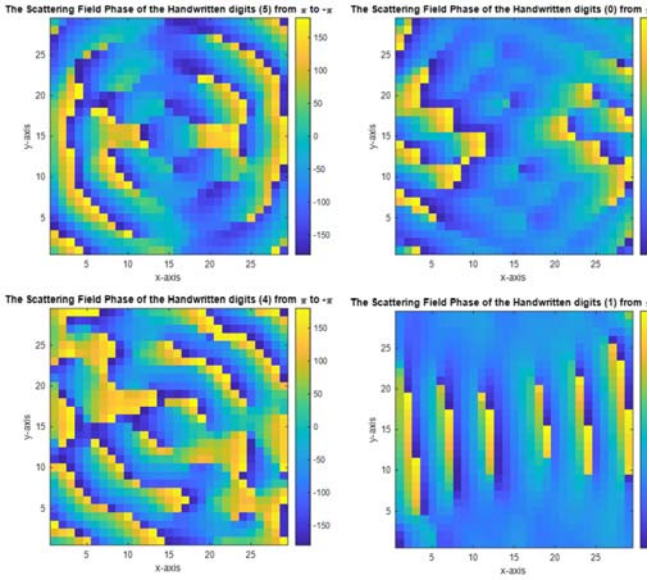

Fig. 8. The phase of the scattering field of the measurement domain for various handwritten digits.

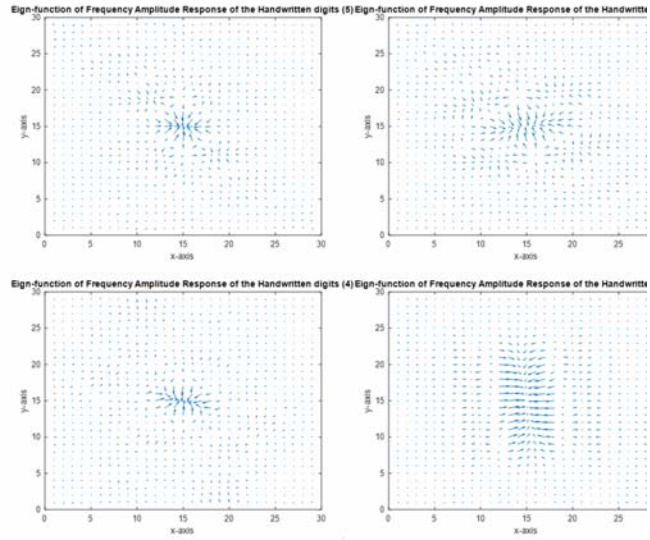

Fig. 9. The Eigen-function of the the measurement domain for various handwritten digits.

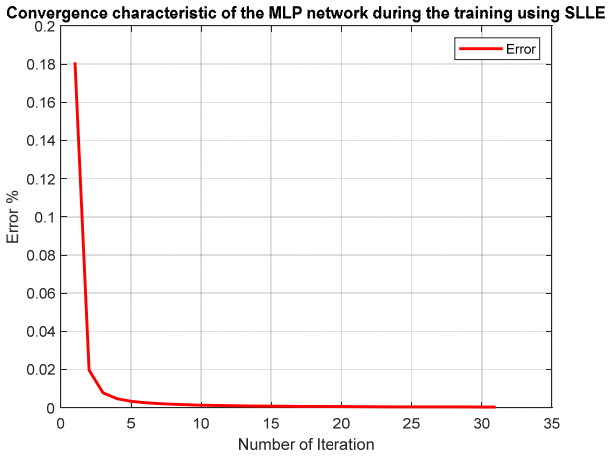

Fig. 10. Convergence characteristic of the MLP network during the training using SLLE, with a large database (10,000 images for training from 0 to 10,000.

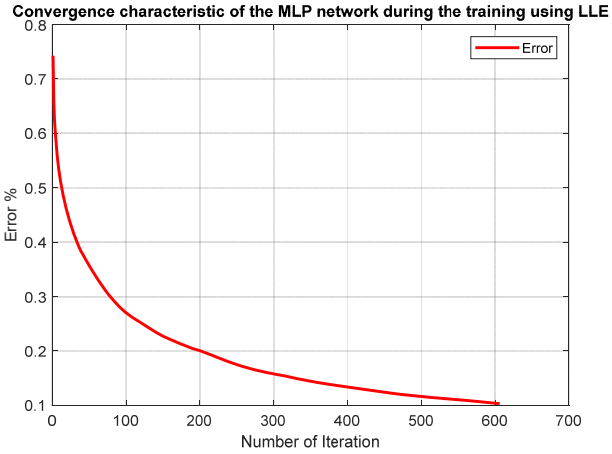

Fig. 11. Convergence characteristic of the MLP network during the training using LLE, with a large database (10,000 images for training from 0 to 10,000.

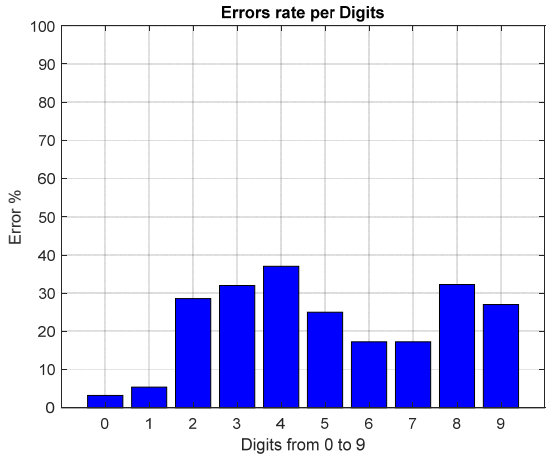

Fig. 12. Bar chart of LLE testing performance for each digit.

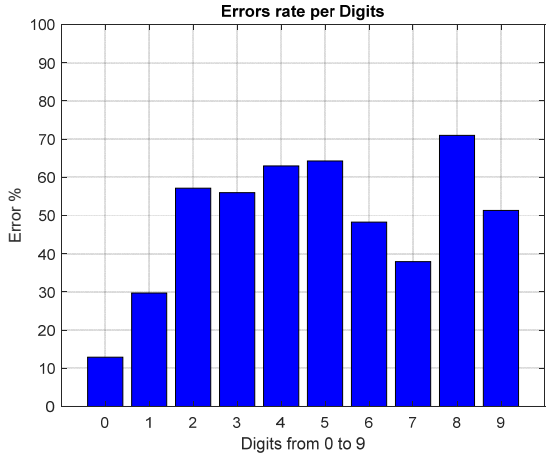

Fig. 13. Bar chart of LLE testing performance for each digit.

6. Further Discussion

In the majority of real-world applications, we do not have prior knowledge regarding data, due to the cost and time burden of applying labels to data samples. In addition, human operatives could label the same data sample differently, which could create false discovery of data sample relationships, which would then impact upon the outcomes of any further calculations. Dimensionality reduction is undertaken through the division of a whole data point (nonlinear hyperplane) into sections of the linear hyperplane. Three stages are undertaken for the production of new data sets, having low dimensionality, weighted matrix, k nearest neighbors, and embedding into a low dimensional space. The MLP is undertaken to employ a low dimensional data set with both LLE and SLLE. For both algorithms, the training process's convergence characteristics were plotted. To undertake to test, behind dimensional data set was mapped onto an identical low dimensional space that was employed during training. Tables and bar charts are used to illustrate how the system performed with LLE and SLLE. The chief drawbacks of employing LLE and SLLE is their complexity and the time they take. SLLE has slightly superior performance to LLE. An advantage of both SLLE and LLE is that they only require a few parameters, i.e., k, the number of nearest neighbors, and d, the dimensionality of the embedded space. System performance is only slightly affected by the number of nearest neighbors, but this is significantly affected by the number of eigenvectors selected.

Additionally, system performance was directly impacted by the data set's nature. How these elements are balanced, and how complex the system becomes, must be decided with reference to the task at hand. In order to achieve optimal training/testing performance, a multiplicity of factors must come into consideration.

3D tomographic imaging radars can provide shape estimations for 3D distribution within a scattering field. The resolution of this type of system is dependent on the spectral and angular windows employed to acquire data and also on polarization diversity. Employing polarization, angular, and spectral degrees of freedom in this type of imaging system increases the amount of information available for object-scattered wavefields. This allows for broadband polarization-selective array apertures to gain a greater amount of information regarding scattering objects than they would in monochrome (i.e., on one frequency) or with a single polarization, as these results demonstrate. There is a useful balance that can be struck between angular and spectral degrees of freedom. As angular degrees of freedom have associations with the quantity of elements/stations in an array, replacing them with more economic spectral degrees of freedom, which have associations with the quantity and frequency points employed in data acquisition, may reduce costs and therefore significantly increase the economic viability of the system. Although microwave diversity imaging systems are attractive, there can be times when either synthetic or physical baselines needed for the realization of the wide-angle windows required for achieving high resolution are either unavailable or insufficient to create an identifiable image. There may be important implications of this research for automated object identification systems, either microwave or other types, and it is worthwhile searching for the next phase of automated neuromorphic radar recognition systems that can accomplish target identification with the minimal examination. Much of the outcomes of this research also have applicability to the robotic recognition and machine vision fields. Transferring the technology into these areas does introduce greater complexity as the objects to be identified will not be discovered in complete isolation, as aerospace targets are.

7. Conclusion

Numerous difficulties with machine learning starts when we preprocess raw multidimensional data sets, such as the scattering field of the target in frequency response. One reason for data preprocessing is to reduce dimensionality. As high dimensional data often carries many redundancies and correlations that leave essential relationships hidden, the aim of the operation is the elimination of redundancy in the processed data. In this paper, a pattern classification system was created for classification of tomographic radar image frequency domains for handwritten digits (0, 1,…., 9) in the MNIST database employing LLE-based dimensionality reduction and MLP-based classification methods. The database employed for training/testing was a normalized complex frequency domain vector for handwritten digits. The system was trained using 1000 images and tested using 100. For data mining, dimensionality reduction is crucial. The aim is to obtain a low-dimensional representation of high-dimensional data without losing any of the data's essential characteristics. Generally, dimensionality reduction is employed as a preprocessing stage prior to data analysis. The low dimensional representation of data is a description of the data's true structure because the reconstruction weights capture and preserve information regarding the data's local neighborhoods. These are invariants in terms of scaling, rotation, and translation. Supervised (with data labels) and unsupervised (without data labels), LLE algorithms are used to reduce dimensionality in the input patterns.

Acknowledgments

This work was supported by the Deanship of Scientific Research with Northern Border University, Arar, Saudi Arabia, funded under Grant ENG-2017-1-7-F-7014.

References

[1] V. G. Sigillito, S. P. Wing, L. V. Hutton, and K. B. Baker, “Classification of radar returns from the ionosphere using neural networks,” *Johns Hopkins APL Tech. Dig. (Applied Phys. Lab.*, Vol. 10, No. 3, pp. 262–266, 1989.

[2] Y. Hong, K. L. Hsu, S. Sorooshian, and X. Gao, “Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system,” *J. Appl. Meteorol.*, 2004.

[3] Y. Hara, R. G. Atkins, S. H. Yueh, R. T. Shin, and J. A. Kong, “Application of neural networks to radar image classification,” *IEEE Trans. Geosci. Remote Sens.*, Vol. 32, No. 1, pp. 100–109, 1994.

[4] S. Haykin and C. Deng, “Classification of radar clutter using neural networks,” *IEEE Trans. Neural Networks*, Vol. 2, No. 6, pp. 589–600, 1991.

[5] J. A. Anderson, M. T. Gately, P. A. Penz, and D. R. Collins, “Radar signal categorization using a neural network,” *Proc. IEEE*, Vol. 78, No. 10, pp. 1646–1657, 1990.

[6] M. Gong, J. Zhao, J. Liu, Q. Miao, and L. Jiao, “Change Detection in Synthetic Aperture Radar Images Based on Deep Neural Networks,” *IEEE Trans. Neural Networks Learn. Syst.*, Vol. 27, No. 1, pp. 125–138, 2016.

[7] S. Skaria, A. Al-Hourani, M. Lech, and R. J. Evans, “Hand-Gesture Recognition Using Two-Antenna Doppler Radar With Deep Convolutional Neural Networks,” *IEEE Sens. J.*, Vol. 19, No. 8, pp. 3041–3048, 2019.

[8] I. Jouny, F. D. Garber, and S. C. Ahalt, “Classification of radar targets using synthetic neural networks,” *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 29, No. 2, pp. 336–344, 1993.

[9] M.-C. Lin and Y.-W. Kiang, “Target discrimination using multiple-frequency amplitude returns,” *IEEE Trans. Antennas Propag.*, Vol. 38, No. 11, pp. 1885–1889, 1990.

[10] H. Lin and A. A. Ksienski, “Optimum Frequencies for Aircraft Classification,” *IEEE Trans. Aerosp. Electron. Syst.*, Vol. AES-17, No. 5, pp. 656–665, 1981.

[11] J. s. Chen and E. K. Walton, “Comparison of Two Target Classification Techniques,” *IEEE Trans. Aerosp. Electron. Syst.*, Vol. AES-22, No. 1, pp. 15–22, 1986.

[12] E. Rothwell, D. Nyquist, K.-M. Chen, and B. Drachman, “Radar target discrimination using the extinction-pulse technique,” *IEEE Trans. Antennas Propag.*, Vol. 33, No. 9, pp. 929–937, 1985.

[13] S. Chakrabarti, N. Bindal, and K. Theagharajan, “Robust radar target classifier using artificial neural networks,” *IEEE Trans. Neural Networks*, Vol. 6, No. 3, pp. 760–766, 1995.

[14] R. Soleti, L. Cantini, F. Berizzi, A. Capria, and D. Calugi, “Neural Network for polarimetric radar target classification,” in *2006 14th European Signal Processing Conference*, 2006, pp. 1–5.

[15] J. A. Garzon-Guerrero, D. P. Ruiz, and M. C. Carrion, “Classification of Geometrical Targets Using Natural Resonances and Principal Components Analysis,” *IEEE Trans. Antennas Propag.*, Vol. 61, No. 9, pp. 4881–4884, 2013.

[16] H. Zhang, D. Ding, Z. Fan, and R. Chen, “Adaptive Neighborhood-Preserving Discriminant Projection Method for HRRP-Based Radar Target Recognition,” *IEEE Antennas Wirel. Propag. Lett.*, Vol. 14, pp. 650–653, 2015.

[17] J. Chen and Y. Liu, “Locally linear embedding: a survey,” *Artif. Intell. Rev.*, Vol. 36, No. 1, pp. 29–48, 2011.

[18] D. DeMers and G. W. Cottrell, “Non-Linear Dimensionality Reduction,” in *Advances in Neural Information Processing Systems 5, [NIPS Conference]*, 1992, pp. 580–587.

[19] H. H. Bock, “On the Interface between Cluster Analysis, Principal Component Analysis, and Multidimensional Scaling BT - Multivariate Statistical Modeling and Data Analysis: Proceedings of the Advanced Symposium on Multivariate Modeling and Data Analysis May 15–16, 19,” H. Bozdogan and A. K. Gupta, Eds. Dordrecht: Springer Netherlands, 1987, pp. 17–34.

[20] J. P. Cunningham and Z. Ghahramani, “Linear Dimensionality Reduction: Survey, Insights, and Generalizations,” *J. Mach. Learn. Res.*, Vol. 16, No. 89, pp. 2859–2900, 2015, [Online]. Available: http://jmlr.org/papers/v16/cunningham15a.html.

[21] R. P. W. Duin and M. Loog, “Linear dimensionality reduction via a heteroscedastic extension of LDA: the Chernoff criterion,” *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 26, No. 6, pp. 732–739, 2004.

[22] H. Abdi, L. J. Williams, D. Valentin, and M. Bennani-Dosse, “STATIS and DISTATIS: optimum multitable principal component analysis and three way metric multidimensional scaling,” *WIREs Comput. Stat.* , Vol. 4, No. 2, pp. 124–167, Mar. 2012.

[23] S. T. Roweis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” *Science (80-.).*, Vol. 290, No. 5500, pp. 2323–2326, 2000.

[24] L. K. Saul and S. T. Roweis, “Think Globally, Fit Locally: Unsupervised Learning of Low Dimensional Manifolds,” *J. Mach. Learn. Res.*, Vol. 4, No. null, pp. 119–155, Dec. 2003.

[25] M. Almutiry, L. Lo Monte, and M. C. Wicks, "Extraction of Weak Scatterer Features Based on Multipath Exploitation in Radar Imagery," *Int. J. Antennas Propag.*, Vol. 2017, No. Article ID 5847872, p. 13, 2017.
[26] P. S. and P. Wang, *Pattern Recognition, Machine Intelligence and Biometrics*. CCISNortheastern UniversityBostonUSA: Springer, Berlin, Heidelberg, 2011.

Muhannad Almutiry received a B.Sc. degree in Electrical Engineering from Umm Al Qura University, Makkah, Saudi Arabia in 2007. He received a M.Sc. degree in Electrical Engineering and a Ph.D. in Electrical Engineering, from the University of Dayton, Dayton, Ohio in 2010 and 2016, respectively. He has been an Associate Professor with the Department of Electrical Engineering, Northern Border University, Saudi Arabia, since 2016. From 2013 to 2016, he was a Research Assistant with the Mumma Radar Laboratory, University of Dayton, Ohio, USA. His main research interests include radar imaging, subsurface sensing, UAV detection, Sonar and intelligent adaptive radar network.

Ahmed Alsheikhy, received the B.S. and M.S. degrees in Electrical and Computer Engineering from King Abdulaziz University 2004 and 2010, respectively. In 2016, he received the Ph.D. degree in Computer Science and Engineering from University of Connecticut, USA. During 2004-2010, he stayed in Technical Vocational Training Corporation (TVTC), Saudi Arabia to train military soldiers about computer networks and technical support. His research interests are on Performance Evaluation, Artificial Intelligence, Satellite and Radar communication systems and Internet of Things.