Intelligent Radar Signal Recognition and Classification

Ivan Jordanov and Nedyalko Petrov

Abstract This chapter investigates a classification problem for timely and reliable identification of radar signal emitters by implementing and following a neural network (NN) based approach. A large data set of intercepted generic radar signals, containing records of their pulse train characteristics (such as operational frequencies, modulation types, pulse repetition intervals, scanning period, etc.), is used for this research. Due to the nature of the available signals, the data entries consist of a mixture of continuous, discrete and categorical data, with a considerable number of records containing missing values. To solve the classification problem, two separate approaches are investigated, implemented, tested and validated on a number of case studies. In the first approach, a listwise deletion is used to clean the data of samples containing missing values and then feed-forward neural networks are employed for the classification task. In the second one, a multiple imputation (MI) model-based method for dealing with missing data (by producing confidence intervals for unbiased estimates without loss of statistical power, i.e. by using all the available samples) is investigated. Afterwards, a feedforward backpropagation neural network is trained to solve the signal classification problem. Each of the approaches is tested and validated on a number of case studies and the results are evaluated and critically compared. The rest of the chapter is organised as follows: the next section (*Introduction and Background*) presents a review of related literature and relevant background knowledge on the investigated topic. In Sect. 2 (*Data Analysis*), a broader formulation of the problem is provided and a deeper analysis of the available data set is made. Different statistical transformation techniques are discussed and a multiple imputation method for dealing with missing data is introduced in Sect. 3 (*Data Pre-Processing*). Several NN topologies, training parameters, input and output coding, and data transformation techniques for facilitating the learning process are tested and evaluated on a set of case studies in Sect. 4 (*Results and Discussion*). Finally, Sect. 5 (*Conclusion*) summarises the results and provides ideas for further extension of this research.

I. Jordanov (✉) · N. Petrov
School of Computing, University of Portsmouth, Portsmouth PO1 3HE, UK
e-mail: ivan.jordanov@port.ac.uk

N. Petrov
e-mail: nedyalko.petrov@port.ac.uk

1 Introduction and Background

What an irony of fate when Robert Watson-Watt was pulled over in a RADAR (***RA***dar ***D***etection ***A***nd ***R***anging) speed trap during his visit in Canada in the late 1950s. He joked that had he known radar would be used for speed traps, he would never have invented it. Nowadays, this is what most people associate the radar with, but when Watson-Watt invented his primitive radar system in the mid 1930s, it was secretly developed for military purposes. Later, in 1940, it played a vital role in the Battle of Britain, providing early warning of incoming Luftwaffe bombers. During the World War II, USA scientists made the Watson-Watt's radar a lot smaller, more efficient and reliable. This made possible a compact radar unit to be used for warning fighter pilots of enemy aircraft approaching from behind. Also, four of these units were carried on each of the nuclear bombs dropped over Hiroshima and Nagasaki to monitor the bomb distance to the ground, so that detonation could be triggered at a pre-set altitude for maximum destruction. Vigorous development of radar technology after the war led to a wide range of military applications for detecting, locating, tracking, and identifying objects, for surveillance, navigation and weapon guidance purposes for terrestrial, maritime, and airborne systems at small to medium and large distances (from ballistic missile defence systems to fist sized tactical missile seekers) [1].

Later, civilian applications emerged and became wide-spread. This began in air traffic control systems to guide commercial aircrafts in the vicinity of the airports and during their flight and in the sea navigation, used by ships in maritime collision avoidance systems. Nowadays, radars are beginning to serve the same role for the automobile and trucking industries in self-braking systems in cars, crash avoidance and parking assist [2, 3].

Police traffic radar are used for enforcing speed limits; airborne radars are used not only for weather forecast, large-scale weather monitoring, prediction and atmospheric research, but also for environmental monitoring of forestry conditions and land usage, water and ice conditions, pollution control, etc.; space-born (both satellite and space shuttle) serve for space surveillance and planetary observation; in sport they are used for measuring the speed of tennis and baseball serves [1].

A basic block-scheme of a radar system is shown in Fig. 1. Radars are considered to be "active" sensors, as they use their own source of illumination (a transmitter) for locating targets. They transmit energy towards a target and then catch the reflected signal to identify the target. The problem is that (especially for a long range radars) a powerful transmitter and very sensitive receiver are needed
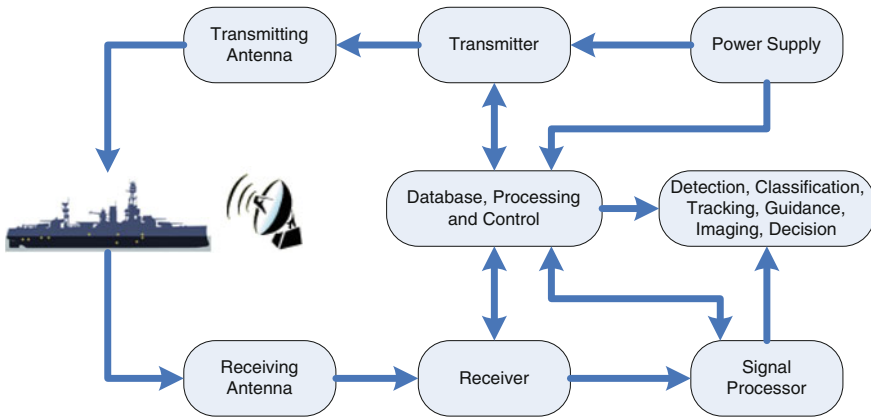
Fig. 1 A block diagram of a basic radar system. Radars operate by transmitting electromagnetic energy toward targets and processing the observed echoes

because the energy spreads out on its way to the target, scatters on reflection and further spreads out on its way back (in general, the decrease of the received signal is proportional to the fourth power of the target distance). The radars range, resolution and sensitivity are generally determined by their transmitter and waveform generator. Although the typical radar systems operate in the microwave region of the electromagnetic spectrum with frequency range of about 200 MHz to about 95 GHz (with corresponding wavelengths of 0.67 m to 3.16 mm), there are also radars that function at frequencies as low as 2 MHz and as high as 300 GHz [4].

The application of the Doppler effect revolutionized the cosmology enabling Doppler spectroscopy to become a powerful tool for finding extrasolar planets and proving the expansion of the universe (the light spectrum of stars (or galaxies) receding from us exhibits *redshift* (increased bandwidth and reduced frequencies), and *blueshift* (higher frequencies and lower bandwidth) if they are moving towards us), but also expanded dramatically the use of radiolocation radars. For the Doppler radars, the reflection from an approaching target electromagnetic wave exhibits higher frequency than the transmitted one and vice versa, a moving away target returns lower frequency wave. The difference between the sent and received frequencies can then be used to estimate the target speed. The problem is that this difference is a very small one, e.g., an incoming target with a 100 km/h increases the received frequency by less than 1e-6, which needs very precise circuits to measure.

A Doppler weather radar with a parabolic antenna situated within a large tiled dome is shown in Fig. 2 [5]. A system with such a radar can measure the distance and lateral speed of falling rain drops, hail particles, or snowflakes, allowing forecasters to predict storms' evolving locations. The presence of debris in the air is

Fig. 2 A Doppler weather radar (*Photo* Brownie Harris/Corbis)

used in similar radar systems to detect tornadoes and define their location, velocity and direction, allowing projections of their movement in real time.

The classical radar imaging uses its antenna to focus a radio frequency beam on a target and capture its reflection to create the image. To work over a long-range it requires powerful transmitters and sensitive receivers because of the way the transmitted energy spreads out on its way to the target and then scatters on reflection. Also, to achieve higher resolution of the image, it needs narrower beams which means that the airborne or space-born platform will need much larger antenna than it could carry. The application of a synthetic aperture technique solves this problem by enabling the use of a smaller antenna through simulating a virtual one with aperture defined by the travel distance of the physical antenna.

The use of the Doppler effect further enhanced the angular resolution in synthetic-aperture radars (SAR) [6] enabling them to acquire surprisingly clear and crisp images [7]. The SAR have been long used on planes and satellites (Fig. 3) for military reconnaissance, mapping ground terrain with intelligence imagery, revealing enemy facilities for enhancing situational awareness and all this in any type of weather, in total darkness and through cloud cover and foliage [8, 69]. They also proved to be very useful in diverse range of civil applications, e.g., in earth-quake damage assessment [9], ice [10] and snow monitoring [11], oceanography, polar ice caps and coastal regions imagery, oil pollution monitoring, solid earth science, hydrology, ecology and planetary science [12, 13].

Fig. 3 JAXA's ALOS-2 Earth-observation radar sat may help the Japanese navy keep track of ship movements in the region. *Photo* JAXA Concept

Another type developed especially to look underground and through walls is the Ground-penetrating radar (GPR), also known as surface-penetrating radar (SPR) [14]. GPR has recently proved to be efficient non-invasive technology with applications in archaeology [15, 16], mining–for both identifying underground rock strata and monitoring instabilities [14, 17], and for optimal irrigation and pollution monitoring [18, 19]. It has been also used for helping police, emergency response and firefighters 'to see' through building walls to locate hostages or help people trapped by fire or under a rubble of a collapsed building [20]. Its ability to see under surface metallic and non-metallic objects makes it useful mapping tool for detection and localisation of underground cables and pipes [21], and buried objects of historical and archaeological importance [22].

The IEEE standard letter nomenclature for the common nominal radar bands is given in Table 1, [23]. The millimetre wave band is sometimes further decomposed into approximate sub-bands of 36–46 GHz (Q band), 46–56 GHz (V band), and 56–100 GHz (W band). The lower frequency bands are usually preferred for longer range surveillance applications due to the low atmospheric attenuation and high available power, and vice versa the higher frequencies tend to be used for shorter range applications and higher resolution, due to the smaller achievable antenna beam widths for a given antenna size, higher attenuation, and lower available power [1]. The radars from the first category (considered a form of radar radiolocation) are capable of covering distances of up to hundreds of kilometres (using high-power transmitters concentrated in a relatively narrow radio bandwidth) and the second group covers radar systems that operate at low power levels, over much smaller distances.

Based on their characteristics, features and application areas, radars can be classified in terms of the following criteria [24]:

- **purpose and function:** surveillance, tracking, guidance, reconnaissance, imaging, data link;
- **frequency band:** radar systems have been operating at frequencies as low as 2 MHz and as high as 300 GHz (see Table 1). Criteria for frequency selection for surveillance radar can be found in [4, 25];
- **waveform:** continuous wave, pulsed wave, digital synthesis;
- **beam scanning:** fixed beam, mechanical scan (rotating, oscillating), mechanical scan in azimuth, electronic scan (phase control, frequency control and mixed in azimuth/elevation), mixed (electronic-mechanical) scan, multi-beam configuration;
- **location:** terrestrial (stable, mobile), marine-borne, air-borne, space-borne;
- **spectrum of collected data:** range (delay time of echo), azimuth (antennae beam pointing, amplitude of echoes), elevation (3D—radar, multifunctional, tracking), height (derived by range and elevation), intensity (echo power), radar cross section (RCS)—(derived by echo intensity and range), radial speed (measurement of differential phase along the time on target due to the Doppler effect—it requires a coherent radar), polarimetry (phase and amplitude of echo in the polarisation channels: horizontally transmitted—HH, horizontally received—HV, VH, VV), RCS profiles along range and azimuth (high resolution along range, imaging radar);
- **configuration:** monostatic (same antenna with co-located transmitter and receiver), bi-static (two antennas), multistatic (one or more spatially dispersed transmitters and receivers). Further detail on variety of radar configurations can be found in [26];
- **signal processing:** coherent (Moving Target Detector/Pulse-Doppler/Super-resolution Signal Processor/Synthetic Aperture Radars (SAR)), non-coherent (integration of envelope signals, moving window, adaptive threshold (Constant False Alarm Rate (CFAR)) and mixed [6];

Table 1 Letter nomenclature for nominal radar frequency bands (IEEE, 2003)

Band	Frequencies	Wavelengths
HF	3–30 MHz	100–10 m
VHF	30–300 MHz	10–1 m
UHF	300 MHz–1 GHz	1–0.3 m
L	1–2 GHz	0.3–0.15 m
S	2–4 GHz	15–7.5 cm
C	4–8 GHz	7.5–3.75 cm
X	8–12 GHz	3.75–2.5 cm
Ku	12–18 GHz	2.5–1.67 cm
K	18–27 GHz	16.7–11.1 mm
Ka	27–36 GHz	11.1–7.5 mm
Q, V, W	36–300 GHz	7.5–1 mm

- **transmitter and receiver technologies:** antenna—reflector plus feed, array (planar, conformal), corporate feed; transmitter—magnetron, klystron, wideband amplifiers (high-power travelling wave tubes (TWT)), solid state; and receiver—analogue and digital technologies, base band, intermediate frequency sampling, low-power TWT;
- **area of application:** large-scale weather forecast and monitoring, air traffic control and guidance (terminal area, en route, collision avoidance, airport apron); police traffic radar used for enforcing speed limits; air defence; anti-theatre ballistic missile defence; vessel traffic surveillance; remote sensing (application to crop evaluation, geodesy, astronomy, defence); environmental monitoring of forestry conditions and land usage; pollution control; geology and archaeology (ground penetrating radar); meteorology (hydrology, rain/hail measurement); study of atmosphere (detection of micro-burst and gust, wind profilers); space-born altimetry for measurement of sea surface height; acquisition and tracking of satellites; monitoring of space debris; marine—navigation and ship collision avoidance; others [5, 12–15].

Radar detection, classification and tracking of targets against a background of clutter and interference are considered as "the general radar problem". For military purposes, the general radar problem includes searching for, interception, localisation, analysis and identification of radiated electromagnetic energy, which is commonly known as radar Electronic Support Measures (ESM). They are considered to be a reliable source of valuable information regarding threat detection, threat avoidance, and, in general, situation awareness for timely deployment of counter-measures [27, 28]. A list of ESM abbreviations is given in Table 2.

A real-time identification of the radar emitter associated with each intercepted pulse train is a very important function of the radar ESM. Typical approaches include sorting incoming radar pulses into individual pulse trains [29], then comparing their characteristics with a library of parametric descriptions, in order to get list of likely radar types. This can be very difficult task as there may be radar modes

Table 2 Commonly adopted ESM abbreviations

Abbreviation	Meaning
EW	Electronic warfare
MOP	Modulation on pulse
PA	Pulse amplitude
PDW	Pulse descriptor word
PPI	Pulse-to-pulse interval
PRI	Pulse repetition interval
PD	Pulse duration
PW	Pulse width
RF	Radio frequency
TOA	Time of arrival
ST	Scanning type
SP	Scan period

for which there is no record in ESM library; overlaps of different radar type parameters; increases in environment density (e.g., Doppler spectrum radars, transmitting hundreds of thousands of pulses per second); agility of radar features, such as radio frequency and scan, pulse repetition interval, etc.; multiplication and dispersion of the modes for military radars; noise and propagation distortion that lead to incomplete or erroneous signals [30].

1.1 Neural Networks in Radar Recognition Systems

There are wide variety of approaches and methods used for radar emitter recognition and identification. For example, [31] investigate a specific emitter identification technique applied to ESM data and by analysing the radar pulses try to extract unique features for each radar, which can be later used for identification. A wavelet transform is employed in [32] for the feature extraction phase in radar signal recognition, as in [33], where they use it before employing probabilistic support vector machines SVMs for the radar emitter recognition task. SVMs are also used in [8, 34] for solving a similar problem. In [35] the authors focus their research on the estimation of a common modulation from a group of intercepted radar pulses and use it as a basis for specific emitter identification. A variety of novel radar emitter recognition algorithms, incorporating clustering and competitive learning, and investigating their advantages over the traditional methods are proposed in [32, 36–42, 70–73].

Among those approaches, a considerable part of the research in the area incorporates NN, due to their parallel architecture, fault tolerance and ability to handle incomplete radar type descriptions and inconsistent and noisy data [43]. NN techniques have previously been applied to several aspects of radar ESM processing [28], including Pulse Descriptor Word (PDW) sorting [44, 45] and radar type recognition [46]. More recently, many new radar recognition systems include NNs as part of a clutter reduction system to improve the information managed by automatic identification systems, such as the detection, positioning, and tracking of surrounding ships [47], or as a key classifier [48–52]. Some examples of NN architectures and topologies used for radar identification recognition and classification based on ESM data include Multilayer Perceptron (MLP) [43], Radial Basis Function (RBF) neural networks as a signal detector [46, 53], a vector neural network [54], and a single parameter dynamic search neural network [50].

In many cases, the NNs are hybridised with other techniques, including fuzzy systems [55], clustering algorithms [29, 56], wavelet packets [32, 57], or Kalman filters [30]. When implementing their "What-and-Where fusion strategy" [30] use an initial clustering algorithm to separate pulses from different emitters according to position-specific parameters of the input pulse stream, and then apply fuzzy ARTMAP (based on Adaptive Resonance Theory (ART) neural network) to classify streams of pulses according to radar type, using their functional parameters. They also complete simulations with a data set that has missing input pattern

components and missing training classes and then incorporate a bank of Kalman filters to demonstrate high-level performance of their system on incomplete, overlapping and complex radar data. In [48] higher order spectral analysis (HOSA) techniques are used to extract information from low probability of intercept (LPI) radar signals to produce 2D signatures, which are then fed to a NN classifier for detecting and identifying the LPI radar signal. The work presented in [49] investigates the potential of NNs (MLPs) when used in Forward Scattering Radar (FSR) applications for target classification. The authors analyse collected radar signal data and extract features, which are then used to train NN for target classification. They also apply K-Nearest Neighbour classifier to compare the results from the two approaches and conclude that the NN solution is superior. In [58] an approach combining rough sets (for data reduction) and NN as a classifier is proposed for radar emitter recognition problem, while [59] combines wavelet packets and neural networks for target classification.

The common denominator of all referenced approaches is that they use predominantly supervised NN learning. This means that there is an available data set (or it is on-line collected), on which the NN can be trained and later used to determine the type of the radar emitters detected in the environment. During the training, the NN is presented with labelled samples from the available dataset and the NN weights are adjusted in order to minimise the difference between the NN output and the available target (supervised learning). This difference is expressed by an error function that is minimised by adjusting the NN weights. One of the most popular methods for training is backpropagation (BP), but, as it uses Newton and quasy-Newton deterministic minimisation methods, it could become trapped in a local minimum and in this way to converge to a suboptimal training. Another drawback of the BP algorithm is that it can, sometimes, be slow and unstable. After training, the NN is tested for its ability to generalise, in other words, its ability to correctly classify samples that have not been shown during the learning process.

Among other considerations, the complexity of the training includes selecting the way of showing the samples to the network (i.e. how the training data set is organised and presented to the NN—'batch mode', 'on-line mode', etc.). Another important question is when to stop the training—achieving a zero error function does not always lead to an optimal training. The reality shows that at some point of the learning process, the NN starts to memorise rather than to generalise—this happens when the NN starts to overfit. In order to avoid the overfitting, an additional data subset (called validation subset), is used in parallel with the training set. Initially, the errors on both sets will decrease, but at some point the validation error will start to rise, while the training error will continue to decrease. This point is an indication of overfitting and the training should be stopped, with the current weights assumed to be optimal. This training approach is known as split sample training, where the available dataset is split in training, validation and testing subsets. There are also other training approaches, such as k-fold crossover, or bootstrapping, each with their own specific advantages and drawbacks [43]. One advantage of the k-fold crossover, for example, is that it can be applied when limited number of samples is available for training.

In addition, often before approaching training, the available data set needs to be pre-processed, e.g., [60] use feature vector fusion before feeding the NN classifier. Radar signal processing has specific features that differentiate it from most other signal processing fields. Many modern radars are coherent, meaning that the received signal, once demodulated to baseband, is complex-valued rather than real-valued and as it can be seen from Table 2, many of the collected data is categorical. Another specificity of the radar data sets is that there are usually many missing or incomplete data. Therefore, the problems of representation and statistical pre-processing of the available dataset are very important steps that need to be considered, before starting the actual training. This may also include transformation techniques, such as linear discriminant analysis and principal component analysis, in order to reduce the dimensionality of the problem and dispose of redundant information in the dataset.

1.2 Dealing with Missing Data

According to statistical analysis, the nature of missing data can be classified into three main groups [61–63]: missing completely at random (MCAR), where the probability that an observation is missing is unrelated to its value or to the value of any other variables; missing at random (MAR)—that missingness does not depend on the value of the observed variable, but on the extent of the missingness correlation with other variables that are included in the analysis (in other words, the cause of missingness is considered); and missing not at random (MNAR)—when the data are not MCAR or MAR (missingness still depends on unobserved data). The problem associated with MNAR is that it yields biased parameter estimates, while MCAR and MAR analysis yield unbiased ones (at the same time the main consequence of using MCAR is loss of statistical power), [63].

Dealing with missingness requires an analysis strategy leading to least biased estimates, while not losing statistical power. The problem is these criteria are contradictory and in order to use the information from the partial data in samples with missing data (keeping up the statistical power), and substituting the missing data samples with estimates, inevitably brings bias.

The most popular approaches in dealing with missing data generally fall in three groups: Deletion methods; Single imputation methods; and Model-based methods [62, 64, 65].

Deletion methods include pairwise and listwise deletion. The pairwise deletion (also called "unwise" deletion) keeps as many samples as possible for each analysis (and in this way uses all available information for it), resulting in incomparable analysis, as each is based on different subsets of data, with different sample sizes and different standard errors. The listwise deletion (also known as complete case analysis) is a simple approach, in which all cases with missing data are omitted. The advantages of this technique include comparability across the analyses and it leads to unbiased parameter estimates (assuming the data is MCAR), while its main

disadvantage is that there may be substantial loss of statistical power (because not all information is used in the analysis, especially if a large number of cases is excluded).

The single imputation methods include mean/mode substitution, dummy variable method, and single regression. Mean/mode substitution is an old procedure, currently rejected due to of its intrinsic problems, e.g., it adds no new information (the overall mean stays the same), reduces the variability, and weakens the covariance and correlation estimates (it ignores relationship between variables). The dummy variable technique uses all available information about missing observation, but produces biased estimates. In the regression approach, linear regression is used to predict what the missing value should be (based of the available other variables) and then uses it as an actual value. The advantage of this technique is that it uses information from the observed data, but overestimates the model fit and the correlation estimates, and weakens the variance [62].

Most popular, "modern" model-based approaches, fall into two categories: multiple imputation (MI) and maximum likelihood (ML) methods (often referred to as full-information maximum likelihood), [63]. Their advantage is that they model the missingness and give confidence intervals for estimates, rather than relying on a single imputation. If the assumption for MAR holds, both groups of methods result in unbiased estimates (i.e., tend to "preserve" means, variances, co-variances, correlations and linear regression coefficients) without loss of statistical power.

ML identifies a set of parameter values that produces the highest (log) likelihood and estimates the most likely value that would result in the observed data. It has the advantage that both complete and incomplete cases are used, in other words, it utilises all of the information and produces unbiased parameter estimates (with MCAR/MAR data). The MI approach involves three distinct steps: first, sets of plausible data for the missing observations are created and these sets are filled in separately to create many 'completed' datasets; second, each of these datasets is analysed using standard procedures for complete datasets; and thirdly, the results from previous step are combined and pooled into one estimate for the inference. The aim of the MI process is not just to fill in the missing values with plausible estimates, but also to plug in multiple times these values by preserving important characteristics of the whole dataset. As with most multiple regression prediction models, the danger of overfitting the data is real and can lead to less generalisable results than would have been possible with the original data [66].

The advantage of the MI technique is that it provides more accurate variability by making multiple imputations for each missing value (it considers both variability due to sampling and variability due to imputation) and its disadvantage is that it depends on the correctly specified model. Also, it requires cumbersome coding, but the latter is not an issue due to the existence of easy to use off-shelf software packages. For the purpose of this investigation, a free, open source R statistical software is used.

2   Data Analysis

For the purpose of this research, a data set composed of 29,094 intercepted generic data samples is used. Each of the captured signals is pre-classified by experts in one of 26 categories, in regards to the platform that can carry the radar emitter (aircraft, ship, missile, etc.) and in one of 142 categories, based on the functions it can perform (3D surveillance, weather tracking, air traffic control, etc.).

Each data entry represents a list of 12 recorded pulse train characteristics (signal frequencies, modulation type, pulse repetition intervals, etc. that will be considered as input parameters), a category label (specifying the radar function and being treated as system output) and a data entry identifier (for reference purposes only) (Table 3).

A more comprehensive summary of the data distribution is presented in Table 4, where an overview of the type, range and percentage of missing values for the recorded signal characteristics is given. The collected data consists of both numerical (integer and float) and categorical values, therefore coding of the categorical fields to numerical representations will be required during the data pre-processing stage. Also, due to the large number of missing values for some of the parameters, approaches for handling of missing data will be considered.

3   Data Pre-processing

The pre-processing of the available data is of a great importance for the subsequent machine learning stage and usually can significantly affect the overall success or failure of the application of a given classification algorithm. In this context, the main objective of this stage is to analyse the available data for inconsistencies, outliers and irrelevant entries and to transform it in a form that could facilitate the underlying mathematical apparatus of the machine learning algorithm and lead to an overall improvement of the classifier's performance.

3.1   Data Cleaning and Imputation

Data cleaning (also known as data cleansing or scrubbing) deals with detecting and removing errors and inconsistencies from data, in order to improve its quality [67]. The most important tasks carried out on this stage would include identification of outliers (entries that are significantly different from the rest and could be a result of an error), resolving of data inconsistencies (values that are not consistent with the specifications or contradict expert knowledge), dealing with missing data (removing the missing values, assigning those values to the attributes' mean, using statistical

Table 3 Sample radar data subset

ID	FN	RFC	RFmin	RFmax	PRC	PRImin	PRImax	PDC	PDmin	PDmax	ST	SPmin	SPmax
84	SS	B	5300	5800	K	–	–	S	–	–	A	5.9	6.1
4354	AT	F	2700	2900	F	1351.3	1428.6	S	–	–	A	9.5	10.5
7488	3D	B	8800	9300	K	100	125	S	13	21	B	1.4	1.6
9632	WT	F	137	139	T	–	–	V	–	–	D	–	–
9839	3D	S	2900	3100	J	–	–	V	99	101	A	9.5	10.5

Missing values (i.e. values that could not have been intercepted or recognised) are denoted by '–'. The rest of the acronyms are defined in Table 4

Table 4 Data description and percentage of missing values

Field	Field description	Type	Categories	Missing (%)
ID	Reference for the line of data	I	–	–
FN	Function performed by the radar ('*3D*'—3D surveillance, '*AT*'—airtraffic control, '*SS*'—surface search, '*WT*'—weather tracker, etc.)	C	142	1.35
RFC	Type of modulation used by the radar to change the frequency of the radar from pulse to pulse ('*A*'—agile, '*F*'—fixed, etc.)	C	12	20.75
RFmin	Min frequency that can be used by the radar	R	–	11.15
RFmax	Max frequency that can be used by the radar	R	–	11.15
PRC	Type of modulation used by the radar to change the Pulse Repetition Interval (PRI) of the radar from pulse to pulse ('*F*'—fixed, etc.)	C	15	15
PRImin	Min PRI that can be used by the radar	R	–	46.70
PRImax	Max PRI that can be used by the radar	R	–	46.70
PDC	Type of modulation used by the radar to change the pulse duration of the radar from pulse to pulse ('*S*'—stable)	C	5	12.92
PDmin	Min pulse duration that can be used by the radar	R	–	46.05
PDmax	Max pulse duration that can be used by the radar	R	–	46.05
ST	Scanning type—method that the radar uses to move the antenna beam ('*A*'—circular, '*B*'—bidirectional, '*W*'—electronically scanned, etc.)	C	28	11.33
SPmin	Min scan period that can be used by the radar	R	–	59.35
SPmax	Max scan period that can be used by the radar	R	–	59.35

In column "Type": I—integer; C—categorical; R—real values

algorithms to predict the missing values) or removing redundant data in different representations.

At this stage of the pre-processing phase, two data sets are prepared. For the purposes of the first two case studies (presented later in this chapter), a data set only containing samples with complete data values is extracted, with the data that could not have been fully intercepted and recognised removed by applying listwise deletion. The second data set (used for the final case study) is received after applying multiple imputation, performed as described below.

3.2 Dealing with Missing Data—Data Imputation

To estimate the values of the missing multivariate data, a sequential imputation algorithm, presented in [68] is used. According to it, if the available data set is denoted with Y and the complete subset with Y_c, the procedure starts from the complete subset to estimate sequentially the missing values of an incomplete observation Y*, by minimizing the covariance of the augmented data matrix Y* = [Y_c, x*]. Subsequently the data sample x* is added to the complete data subset and the algorithm continues with the estimate of next data sample with missing values.

Implementations in R of the original algorithm (available under the function name *"impSeq"*) and two modifications of it (namely *"impSeqRob"* and *"impNorm"*) are considered and tested. As the original algorithm uses the sample mean and covariance matrix, it is vulnerable to the presence of outliers, but this can be enhanced by including robust estimators of location and scatter (which is realised in the *"impSeqRob"* function). However, the outlyingness metric can be computed for a complete dataset only, therefore the sequential imputation of the missing data is done first and then the outlyingness measure is computed and used to define whether the observation is an outlier or not. If the measure does not exceed a predefined threshold, the observation is included in the next stage of the algorithm. In our investigation, however, the use of modified *"impSeqRob"* and *"impNorm"* versions did not produce better results when tested on complete dataset (which may be simply due to the lack of outliers), so the *"impSeq"* function was adopted.

After employing MI on the data samples with missing continuous values, a second dataset of 15656 observations is received, which is more than double the size of the first dataset. Table 5 shows the inputted values produced by the MI algorithm for the sample subset, presented previously in Table 3.

3.3 Data Coding and Transformation

This stage of the pre-processing aims to transform the data into a form that is appropriate for feeding to the selected classifier and would facilitate faster and more accurate machine learning.

In particular, a transformation known as coding is applied to convert the categorical values presented in the data set into numerical ones. Three of the most broadly applied coding techniques are investigated and evaluated—continuous, binary and introduction of dummy variables.

For the first type of coding, each of the categorical values is substituted by a natural number, e.g., the 12 categories for the RFC input are encoded with 12 ordinal numbers, the 15 PRC categories—with 15 ordinal numbers, etc. A sample of data subset coded with continuous values is given in Table 6.

Table 5 Sample radar data subset with imputed values for the missing data entries

ID	FN	RFC	RFmin	RFmax	PRC	PRImin	PRImax	PDC	PDmin	PDmax	ST	SPmin	SPmax
84	SS	B	5300	5800	K	963.2	5625	S	5.8	17	A	5.9	6.1
4354	AT	F	2700	2900	F	1351	1428	S	4	6.3	A	9.5	10.5
7488	3D	B	8800	9300	K	100	125	S	13	21	B	1.4	1.6
9632	WT	F	137	139	T	622.6	31,312	V	61.1	93.1	D	12	47.8
9839	3D	S	2900	3100	J	2058	48,128	V	99	101	A	9.5	10.5

Table 6 Sample subset with imputed radar data and natural number coding for the 'RFC', 'PRC', 'PDC', and 'ST' signal characteristics

ID	RFC	RFmin	RFmax	PRC	PRImin	PRImax	PDC	PDmin	PDmax	ST	SPmin	SPmax
84	2	5300	5800	7	963.2	5625	1	5.8	17	1	5.9	6.1
4354	4	2700	2900	4	1351	1428	1	4	6.3	1	9.5	10.5
7488	2	8800	9300	7	100	125	1	13	21	2	1.4	1.6
9632	4	137	139	11	622.6	31,312	2	61.1	93.1	4	12	47.8
9839	9	2900	3100	6	2058	48,128	2	99	101	1	9.5	10.5

Binary coding, wherein each non-numerical value is substituted by log_2N (where N is the number of categories taken by that variable) new binary variables (i.e. taking value of either 0 or 1), is illustrated in Table 7 for 32 categories.

Finally, the non-numerical attributes are coded using dummy variables. In particular, every N levels of a categorical variable are represented by introducing N dummy variables. An example of dummy coding for 32 categorical levels is shown in Table 8.

Taking into account the large number of categories presented for the categorical attributes in the input data set (Table 4), continuous and binary codings are considered for transforming the input variables. On the other hand, binary and dummy variable codings are chosen for representing the output parameters.

Finally, in order to balance the impact of the different input parameters on the training algorithm, data scaling is used. Correspondingly, each of the conducted experiments in this chapter is evaluated using 3 forms of the input data set: the original data (with no scaling); normalised data (scaled attribute values within [0, 1] interval); and standardised data (i.e. scaling the attribute values to a zero mean and unit variance). A sample binary coded and standardised data subset is given in Table 9.

Table 7 Example of binary coding for 32-level categorical variable

Original category		Encoded variables				
Index	Label	B1	B2	B3	B4	B5
1	'2D'	0	0	0	0	0
2	'3D'	0	0	0	0	1
3	'AA'	0	0	0	1	0
...						
16	'CS'	0	1	1	1	1
...						
32	'ME'	1	1	1	1	1

Table 8 Example of dummy coding for 32-level categorical variable

Original category		Encoded variables								
Index	Label	D1	D2	D3	D4	D5	...	D16	...	D32
1	'2D'	1	0	0	0	0	...	0	...	0
2	'3D'	0	1	0	0	0	...	0	...	0
3	'AA'	0	0	1	0	0	...	0	...	0
...										
16	'CS'	0	0	0	0	0	...	1	...	0
...										
32	'ME'	0	0	0	0	0	...	0	...	1

Table 9 Sample subset with imputed radar data and binary coding

ID	RFC	RFmin	RFmax	PRC	PRImin	PRImax	PDC	PDmin	PDmax	ST	SPmin	SPmax
84	0001	5300	5800	0110	963	5625	0	5.8	17	00000	5.9	6.1
4354	0011	2700	2900	0011	1351	1428	0	4	6.3	00000	9.5	10.5
7488	0001	8800	9300	0110	100	125	0	13	21	00001	1.4	1.6
9632	0011	137	139	1010	622	31,312	1	61.1	93.1	00011	12	47.8
9839	1000	2900	3100	0101	2058	48,128	1	99	101	00000	9.5	10.5

3.4 System Training

The investigated neural network topologies include one hidden layer, with fully connected neurons in the adjacent layers and batch-mode training. For a given experiment with P learning samples, the error function is presented as:

$$E_P = \frac{1}{2} \sum_{p=1}^{P} \sum_{i=1}^{L} (x_i^p - t_i^p)^2, \tag{1}$$

where for each sample $p = 1, ..., P$ and each neuron of the output layer $i = 1, ..., L$, a pair (x_i, t_i) of NN output and target values, respectively, is defined.

4 Results and Discussion

A number of experiments are designed, implemented, executed and evaluated to test and validate the performance of the proposed intelligent system for identification and classification of radar signals. Two separate approaches are considered and the related results are grouped and presented in the following two case studies. MATLAB® and its *Statistics*, *Neural Networks* and *Global Optimisation* toolboxes are used for coding and running of all the experiments.

4.1 Case Study 1—Listwise Deletion and Feedforward Neural Networks

For the purposes of the first case study, samples that contain incomplete data (i.e. data that was not fully intercepted or recorded) are removed from the considered data set, resulting in a subset of 7693 complete data samples of radar signal values.

Subsequently, depending on the experiment to be performed, the samples are sorted by experts in several groups of major interest according to their application. In two classes for the first two experiments ("*Civil*" and "*Military*"), and in 11 classes for the purpose of the final one (4 from the "*Civil*" and 7 from the "*Military*" application areas).

A randomly selected, no missing data sample subset (after listwise deletion) is presented in Table 10. Its first column (the ID attribute) is retained for referencing purposes only and it is not used during the classifier's training.

Next, a coding transformation (as described in Sect. 3.2) is applied to convert the categorical values in the data set to numerical ones. Taking into account the large number of categories in the inputs (Table 4), continuous and binary codings are

Table 10 Sample radar data subset with no missing values, received after listwise deletion

ID	FN	RFC	RFmin	RFmax	PRC	PRImin	PRImax	PDC	PDmin	PDmax	ST	SPmin	SPmax
983	AT	F	15,700	17,700	F	100	142.9	S	0.03	0.05	A	0.9	1.1
1286	SS	A	5500	5800	K	909.1	1111.1	S	0.6	0.8	A	1.9	2.1
4846	SS	F	172	180	F	2439	2564.1	S	1.6	1.8	G	28	32
12,097	3D	D	5250	5850	F	2703	2777.8	S	3	3.3	A	5.8	6.2
28,059	WT	F	5300	5700	F	1127	1132.5	S	0.75	0.85	C	12	60

considered for transforming the input variables. On the other hand, binary and dummy variable representations are used for transforming the output parameters.

In order to balance the impact of the different input parameters on the training algorithm, data scaling is applied. Respectively, each of the experiments conducted for the purposes of this case study is evaluated using three forms of the input data set—the data itself (with no scaling), after normalisation (i.e., scaling the attribute values to fall within a specific range, for example [0 1]), and after standardisation (i.e. scaling the attribute values to a zero mean and unit variance). A sample binary encoded and normalised data subset is given in Table 11.

The investigated NN topologies include one hidden layer, with fully connected neurons in the adjacent layers and batch-mode training. For a given experiment with P learning samples, the error function is given with Eq. 1. Supervised NN learning with Levenberg-Marquardt algorithm and tangent sigmoid transfer function is used. A split-sample technique using randomly selected 70 % of the available data for training, 15 % for validation and 15 % for testing, and mean squared error (MSE) is adopted for evaluating the learning performance. The stopping criteria is set to 500 training epochs, gradient reaching less than 1.0e-06 or if 6 consequent validation checks fail, whichever occurs first.

For the purposes of the first experiment, the categorical attributes of the input data are coded with consecutive integers. In this way a total of 12 input variables are received (Table 6). Two neural network topologies are examined—12-10-1 (12 neurons in the input, 10 neurons in the hidden and 1 neuron in the output layers) and 12-10-2, where the output parameter is coded as one binary neuron taking values *0* (*"Civil"*) and *1* (*"Military"*) for the first topology and 2 binary neurons, taking values *10* (*"Civil"*) and *01* (*"Military"*) for the second topology (Fig. 4). The performance of each of the topologies is investigated, evaluated and compared after training with the original, normalised and standardised data. The results are summarised in Table 12 and Fig. 5.

The second experiment investigates two additional NN topologies: 22-22-1 and 22-22-2, where the output parameter is again coded by one binary neuron (*0* for *"Civil"* and *1* for *"Military"*) for the first topology and by two binary neurons for the second one (*10* for *"Civil"* and *01* for *"Military"*). Again, the performance of each of the topologies is investigated, evaluated and compared using the original data, after normalisation and after standardisation. The results are summarised in Table 13.

Similarly to the first experiment, sample confusion matrices are presented in Fig. 6 for a 22-22-2 NN classifier trained with standardised input data. A very high accuracy of 84.3 % on the testing data set is achieved after 114 epochs and activation of the validation check stopping criteria (unsatisfactory performance on the validation data set in six successive iterations).

The final experiment in this case study investigates a broader output space of 11 classes (4 from the *"Civil"* and 7 from the *"Military"* domain) and evaluates a 22-22-11 NN classifier with unscaled, normalised and standardised training data using dummy variable coded outputs. Summary of the obtained results is presented in Table 14 and a sample confusion matrix for the investigated classifier with

Table 11 Sample radar data subset with no missing values (listwise deletion), after binary encoding and normalisation

| ID	RFC enc.				RFmin	RFmax	PRC enc.				PRImin	PRImax	PDC enc.	PDmin	PDmax	ST enc.				SPmin	SPmax
983	0	0	1	1	0.228	0.249	0	0	1	1	0.0006	0.0003	0	0.000010	0.000014	0	0	0	0	0.00025	0.00030
1286	0	0	0	0	0.080	0.082	0	1	1	0	0.0056	0.0022	0	0.000303	0.000363	0	0	0	0	0.00054	0.00058
4846	0	0	1	1	0.002	0.003	0	0	1	1	0.0151	0.0051	0	0.000815	0.000828	0	1	1	0	0.00789	0.00877
12,097	0	0	0	0	0.076	0.082	0	0	1	1	0.0167	0.0055	0	0.001533	0.001526	0	0	0	0	0.00163	0.00170
28,059	0	0	1	1	0.077	0.080	0	0	1	1	0.0070	0.0023	0	0.000379	0.000386	0	0	1	0	0.00338	0.01644
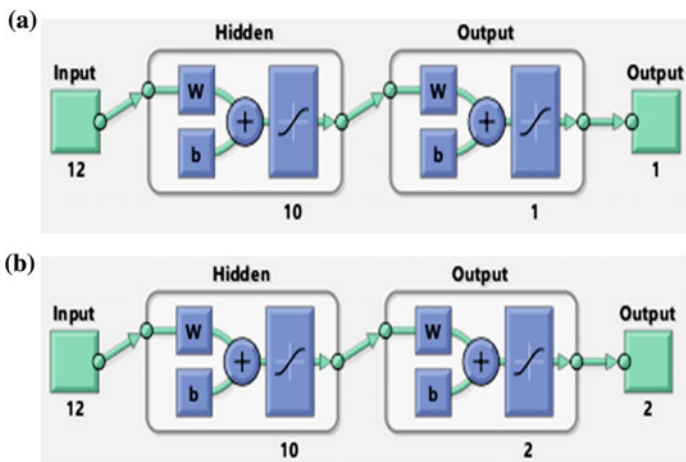
Fig. 4 Investigated NN topologies for case study 1: 12 neurons in the input layer; 10 in the hidden; and 1 (**a**), or 2 (**b**) neurons in the output layer

Table 12 Classification performance (over the testing set) for continuous input coding and 12-10-N topologies with no data scaling, after normalisation and after standardisation

NN topology	Inputs scaling	Classification accuracy (%)
12-10-1	No scaling	78.12
	Normalisation	80.82
	Standardisation	80.76
12-10-2	No scaling	80.14
	Normalisation	81.60
	Standardisation	82.18

Table 13 Classification performance (over the testing set) for binary input coding and 22-22-N topologies with no data scaling, after normalisation and after standardisation

NN topology	Inputs scaling	Classification accuracy (%)
22-22-1	No scaling	81.90
	Normalisation	83.34
	Standardisation	83.01
22-22-2	No scaling	81.77
	Normalisation	83.90
	Standardisation	84.30

Fig. 5 Classification results for 12-10-2 NN classifier with normalised input data and a validation stop after 118 epochs. The values in *green* specify the correctly classified samples for each class (10—Civil, 01—Military)

standardised input training data is given in Fig. 7, where a good recognition rate of 67.49 % can be observed.

Although a straightforward comparison with radar classification studies reported by other authors might be misleading, due to the different data sets, model parameters and training methods used, the achieved results appeared to be strongly competitive when compared to the ones reported in [30, 32, 48, 49, 60]. Furthermore, additional improvement is expected, if further statistical pre-processing techniques, missing data handling routines, NN topologies or training algorithm parameters are investigated (as shown in the next two case studies).

Fig. 6 Classification results for 22-22-2 NN classifier with normalised input data and a validation stop after 114 epochs. The values in *green* specify the correctly classified samples for each class (*10—"Civil", 01—"Military"*)

Table 14 Classification performance (over the testing set) for binary input coding and 22-22-11 topology with no data scaling, after normalisation and after standardisation

NN topology	Inputs scaling	Classification accuracy (%)
22-22-11	No scaling	61.94
	Normalisation	66.70
	Standardisation	67.49

Test Confusion Matrix

Output Class	M1	M2	M3	M4	M5	M6	M7	C1	C2	C3	C4	
10000000000	3	0	0	0	0	0	0	0	0	0	0	100%
01000000000	0	5	0	0	0	0	0	0	0	0	0	100%
00100000000	0	1	31	0	1	2	0	2	1	0	0	81.6%
00010000000	3	5	6	237	39	22	6	32	1	5	21	62.9%
00001000000	0	0	0	20	63	18	0	3	0	0	2	59.4%
00000100000	2	0	2	4	4	16	0	0	0	0	0	57.1%
00000010000	0	0	0	0	0	0	0	0	0	0	0	0.0%
00000001000	0	1	3	33	4	1	0	305	2	1	4	86.2%
00000000100	1	1	1	3	3	1	0	4	23	1	2	57.5%
00000000010	0	0	1	6	4	4	1	1	2	51	3	69.9%
00000000001	1	0	1	27	8	0	3	0	1	1	86	67.2%
	30.0%	38.5%	68.9%	71.8%	50.0%	25.0%	0.0%	87.9%	76.7%	86.4%	72.9%	67.49%

Target Class

Fig. 7 Classification results for 22-22-11 NN classifier with standardised data on 7 military (*M1*—*"Multi-function"*, *M2*—*"Battlefield"*, *M3*—*"Aircraft"*, *M4*—*"Search"*, *M5*—*"Air Defense"*, *M6*—*"Weapon"* and *M7*—*"Information"*) and 4 civil classes (*C1*—*"Maritime"*, *C2*—*"Airborne Navigation"*, *C3*—*"Meteorological"* and *C4*—*"Air Traffic Control"*)

4.2 Case Study 2—Multiple Imputation and Feedforward Neural Networks

The second case study follows the same sequence of experiments and NN topologies, as introduced in the first study, however, this time an extended dataset, received after multiple imputation of the missing data values (as described in Sect. 3) is used.

For the purposes of the first experiment in this study, the categorical attributes of the input data are coded with consecutive integers. Two NN topologies are examined—12-10-1 and 12-10-2, where the output parameter is coded as one binary neuron taking values 0 (*"Civil"*) and 1 (*"Military"*) for the first topology and 2 neurons, taking binary values 10 (*"Civil"*) and 01 (*"Military"*) for the second one.

The performance of each of the topologies is investigated, evaluated and compared using training with the original data (no pre-processing), and after normalisation and standardisation. The results are summarised in Table 15 showing up to 5 % accuracy improvement for the case introducing imputation.

Sample confusion matrices for a 12-10-2 NN classifier trained with normalised input data and a validation stop activated after 106 epochs are given in Fig. 8. They demonstrate improved accuracy rates (especially for the *"Military"* class) when compared to the case studies using listwise deletion to cope with the incomplete data samples (Fig. 5).

The second experiment in this study investigates two additional NN topologies —22-22-1 and 22-22-2, where the output is again coded by one binary neuron (*0* for *"Civil"* and *1* for *"Military"*) for the first topology and by two binary neurons for the second one (*10* for *"Civil"* and *01* for *"Military"*).

The NN performance for each of the topologies is investigated, evaluated and compared using the original, normalised and standardised data for both the cases— with and without imputed values. The performance results are summarised in Table 16, again showing improved NN performances for the cases with imputed data.

The final experiment investigates a broader output space of 11 classes (4 *"Civil"* and 7 *"Military"*) and evaluates 22-22-11 NN classifiers with the original, normalised and standardised training data, and with dummy variable coded outputs. Summary of the obtained results when training on data subsets with and without imputation is presented in Table 17.

Sample confusion matrices for the imputed 22-22-11 NN case, trained with standardised input data and a validation stop activated after 98 epochs are presented in Fig. 9. Although the results seem slightly inferior to the listwise deletion case (Fig. 7), they give higher statistical confidence because of the increased number of samples.

It can also be seen from Fig. 9 that although the accuracy of the NN classifier is relatively the same (compared to the NN trained after listwise deletion (Fig. 7)), the

Table 15 Classification performance (over the testing set) for continuous input coding and 12-10-N topologies with no data scaling, after normalisation and after standardisation

Topology	Input data	% Accuracy	
		No imputation	With imputation
12-10-1	No scaling	78.1	83.3
	Normalised	80.8	84.5
	Standardised	80.8	85.2
12-10-2	No scaling	80.1	82.1
	Normalised	81.6	83.6
	Standardised	82.1	84.5

Comparison between NN training with data received after listwise deletion and after multiple imputation

Fig. 8 Classification results for imputed data case for 12-10-2 NN classifier with normalised input data and a validation stop after 106 epochs. The values in green specify the correctly classified samples for each class (*10—"Civil"*, *01—"Military"*)

number of hits is largely increased and with a better distribution. This is especially evident for the *'M7'* class, for which there were no hits in the case without imputation. The best accuracy is again achieved for the *'M4'* and *'C1'* classes, but the more important achievement as a result of the imputation is the uniform

Table 16 Classification performance (over the testing set) for binary input coding and 22-22-N topologies with no data scaling, after normalisation and after standardisation

Topology	Input data	% Accuracy	
		No imputation	With imputation
22-22-1	No scaling	81.9	85.6
	Normalised	83.3	87.3
	Standardised	83.1	87.2
22-22-2	No scaling	81.8	84.8
	Normalised	83.9	85.0
	Standardised	84.3	86.8

Comparison between NN training with data received after listwise deletion and after multiple imputation

Table 17 Classification performance (over the testing set) for binary input coding and 22-22-11 topology with no data scaling, after normalisation and after standardisation

Topology	Input data	% Accuracy	
		No imputation	With imputation
22-22-11	No scaling	61.9	66.1
	Normalised	66.7	66.4
	Standardised	67.5	66.7

Comparison between NN training with data received after listwise deletion and after multiple imputation

distribution of correctly classified samples. As illustrated in Fig. 7, the class accuracy variance for the classification with no missing data is very high, from 0 to 87.9 %, whereas in the case using imputed data (Fig. 9), it is between 22.6 and 87.4 %. In other words, while keeping the best accuracy almost the same, the
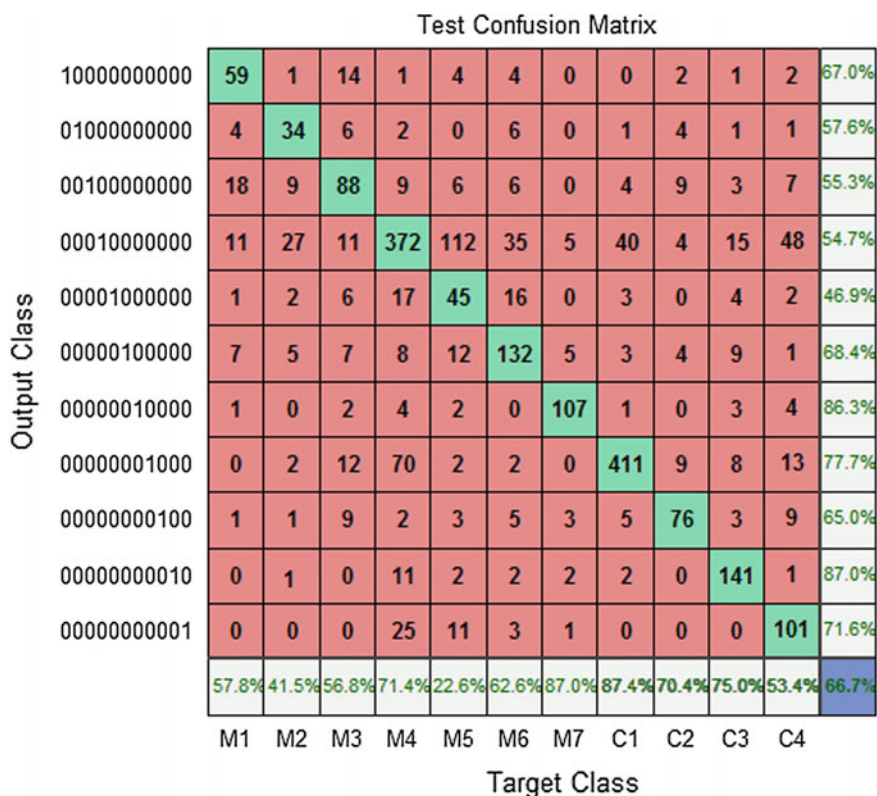
Fig. 9 Classification results for inputed data and 22-22-11 NN classifier with standardised data on 7 military (*M1*—"*Multi-function*", *M2*—"*Battlefield*", *M3*—"*Aircraft*", *M4*—"*Search*", *M5*—"*Air Defense*", *M6*—"*Weapon*" and *M7*—"*Information*") and 4 civil classes (*C1*—"*Maritime*", *C2*—"*Airborne Navigation*", *C3*—"*Meteorological*" and *C4*—"*Air Traffic Control*")

minimum accuracy is improved by more than 22 %. This should be attributed to the greater number of available training and testing samples as a result of the imputation, which increases the statistical power of the dataset and subsequently improves the classification performance of the NN.

5 Conclusion

Reliable and real-time identification of radar signals is of crucial importance for timely threat detection, threat avoidance, general situation awareness and timely deployment of counter-measures. In this context, this chapter investigates the potential application of NN-based approaches for timely and trustworthy identification of radar types, associated with intercepted pulse trains.

A number of experiments are designed, implemented, executed and evaluated for testing and validating the performance of the proposed intelligent systems for solving the investigated classification tasks. The different experiments study a variety of NN topologies, data transformation techniques and missing data handling approaches.

The simulations are divided in two broad case studies, each of which conducts several sub-experiments. In the first one, all the signals are pre-classified by experts into between 2 and 11 classes, depending on the experiment, and then a listwise deletion is used to clean the data from incomplete samples. As a result, very competitive classification accuracy of about 81, 84 and 67 % is received for the different recognition tasks.

In the second one, a study applying a multiple imputation model-based approach for dealing with the large number of missing data (contained in the available radar signals data set) is investigated. The experiments conducted for the purposes of the first case study are repeated, but this time using the imputed data set for training of the classifiers. An improved accuracy of up to 87.3 % is achieved. The results are compared and critically analysed, showing overall improved accuracy when the NN are trained on the larger subset with imputed values.

Although a straightforward comparison to radar classification studies, reported by other authors might be misleading, due to the different data sets, model parameters, data transformations, training and optimisation methods used, the achieved results are strongly competitive to the ones reported in [30, 42, 48, 49, 52, 60].

Potential areas for further extension of this research include investigation of additional statistical transformation techniques, such as Principal Component Analysis (PCA), Non-Linear Principal Component Analysis (NLPCA), and Linear Discriminant Analysis, for decreasing the dimensionality of the problem and increasing the separability between the classes. In terms of classifiers, we presented supervised learning and classification, but unsupervised learning techniques (such as self-organising maps (SOM)) can also be considered, as well as varying other

training parameters and exploring additional NN topologies. Finally, additional classes can be introduced, in order to achieve more specific classification of the intercepted radar data.

References

1. Richards, M.A.: Fundamentals of Radar Signal Processing. Tata McGraw-Hill Education, New Delhi (2005)
2. Dickmann, J., Appenrodt N., Brenk, C.: Making Bertha see. IEEE Spectrum, 41–45 (2014)
3. Mure-Dubois, J., Vincent, F., Bonacci, D.: Sonar and radar SAR processing for parking lot detection. In: Proceedings of the IEEE International Radar Symposium (IRS), pp. 471–476 (2011)
4. Skolnik, M.I.: Introduction to Radar Systems. McGraw Hill, Boston (2001)
5. Lazarus, M.: Radar everywhere. IEEE Spectr. **52**(2), 52–59 (2015)
6. Cumming, I.G., Wong, F.H.: Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation. Artech House, Norwood (2005)
7. Krieger, G., Younis, M., Gebert, N., Huber, S., Bordoni, F., Patyuchenko, A., Moreira, A.: Advanced concepts for high-resolution wide-swath SAR imaging. In: 8th European Conference on Synthetic Aperture Radar (EUSAR), pp. 1–4. VDE (2010)
8. Zhai, S., Jiang, T.: A new sense-through-foliage target recognition method based on hybrid differential evolution and self-adaptive particle swarm optimization-based support vector machine. Neurocomput. **149**(1), 573–584 (2015)
9. Brunner, D., Lemoine, G., Bruzzone, L.: Earthquake damage assessment of buildings using VHR optical and SAR imagery. IEEE Trans. Geosci. Remote Sens. **48**(5), 2403–2420 (2010)
10. Arcone, S.A., Spikes, V.B., Hamilton, G.S.: Stratigraphic variation within polar firn caused by differential accumulation and ice flow: interpretation of a 400 MHz short-pulse radar profile from West Antarctica. J. Glaciol. **51**(174), 407–422 (2005)
11. Storvold, R., Malnes, E., Larsen, Y., Høgda, K., Hamran, S., Mueller, K., Langley, K.: SAR remote sensing of snow parameters in Norwegian areas—Current status and future perspective. J. Electromagn. Waves Appl. **20**(13), 1751–1759 (2006)
12. Evans, D.L., Alpers, W., Cazenave, A., Elachi, C., Farr, T., Glackin, D., Holt, B., Jones, L., Liu, W.T., McCandless, W.: Seasat—a 25-year legacy of success. Remote Sens. Environ. **94** (3), 384–404 (2005)
13. Secades, C., O'Connor, B., Brown, C., Walpole, M.: Earth observation for biodiversity monitoring: a review of current approaches and future opportunities for tracking progress towards the Aichi Biodiversity Targets, CBD Technical Series 72, (2014)
14. Persico, R.: Introduction to Ground Penetrating Radar: Inverse Scattering and Data Processing. Wiley, New York (2014)
15. Goodman, D., Piro, S.: GPR Remote Sensing in Archaeology. Springer, New York (2013)
16. Sambuelli, L., Bohm, G., Capizzi, P., Cardarelli, E., Cosentino, P.: Comparison between GPR measurements and ultrasonic tomography with different inversion algorithms: an application to the base of an ancient Egyptian sculpture. J. Geophys. Eng. **8**(3), 106–116 (2011)
17. Francke, J.: Applications of GPR in mineral resource evaluations. In: Proceedings of the 13th IEEE International Conference on Ground Penetrating Radar (GPR), pp. 1–5 (2010)
18. Lambot, S., Slob, E.C., van den Bosch, I., Stockbroeckx, B., Vanclooster, M.: Modeling of ground-penetrating radar for accurate characterization of subsurface electric properties. IEEE Trans. Geosci. Remote Sens. **42**(11), 2555–2568 (2004)
19. Soldovieri, F., Prisco, G., Persico, R.: A strategy for the determination of the dielectric permittivity of a lossy soil exploiting GPR surface measurements and a cooperative target. J. Appl. Geophys. **67**(4), 288–295 (2009)

20. Peabody, J.E., Charvat, G.L., Goodwin, J., Tobias, M.: Through-wall imaging radar. Lincoln Lab. J. **19**(1), 62–72 (2012)
21. Pettinelli, E., Di Matteo, A., Mattei, E., Crocco, L., Soldovieri, F., Redman, J.D., Annan, A.P.: GPR response from buried pipes: Measurement on field site and tomographic reconstructions. IEEE Trans. Geosci. Remote Sens. **47**(8), 2639–2645 (2009)
22. Slob, E., Sato, M., Olhoeft, G.: Surface and borehole ground-penetrating-radar developments. Geophysics **75**(5), 75A103–175A120 (2010)
23. IEEE Standard Letter Designations for Radar-Frequency Bands. Standard 521–2002 (2003). doi:10.1109/IEEESTD.2003.94224
24. Capraro, G.T., Farina, A., Griffiths, H., Wicks, M.C.: Knowledge-based radar signal and data processing: a tutorial review. IEEE Sig. Process. Mag. **23**(1), 18–29 (2006)
25. Galejs, R.J.: Volume surveillance radar frequency selection. In: Proceedings of the IEEE International Radar Conference, pp. 187–192 (2000)
26. Johnsen, T., Olsen, K.E.: Bi-and multistatic radar. DTIC Document (2006)
27. Schleher, D.C.: Electronic warfare in the information age. Artech House, Boston, London (1999)
28. Sciortino, J.C.: Autonomous ESM systems. Naval Eng. J. **109**(6), 73–84 (1997)
29. Wang, Z., Zhang, D., Bi, D., Wang, S.: Multiple-parameter radar signal sorting using support vector clustering and similitude entropy index. Circ. Syst. Sig. Process. **33**(6), 1985–1996 (2014)
30. Granger, E., Rubin, M.A., Grossberg, S., Lavoie, P.: A what-and-where fusion neural network for recognition and tracking of multiple radar emitters. Neural Netw. **14**(3), 325–344 (2001)
31. D'Agostino, S., Foglia, G., Pistoia, D.: Specific emitter identification: Analysis on real radar signal data. In: Proceedings of the IEEE European Radar Conference (EuRAD), pp. 242–245 (2009)
32. Feng, B., Lin, Y.: Radar signal recognition based on manifold learning method. Int. J. Control Autom. **7**(12), 399–440 (2014)
33. Li, J., Ying, Y.: Radar signal recognition algorithm based on entropy theory. In: Proceedings of the 2nd IEEE International Conference on Systems and Informatics (ICSAI 2014), pp. 718–723 (2014)
34. Zhang, G., Hu, L., Jin, W.: A novel approach for radar emitter signal recognition. In: Proceedings of the IEEE Asia-Pacific Conference on Circuits and Systems, pp. 817–820 (2004)
35. Lunden, J., Koivunen, V.: Automatic radar waveform recognition. IEEE J. Sel. Top. Sig. Process. **1**(1), 124–136 (2007)
36. Hassan, H.: A new algorithm for radar emitter recognition. In: Proceedings of the 3rd IEEE International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 1097–1101 (2003)
37. Liu, H.-J., Liu, Z., Jiang, W.-L., Zhou, Y.-Y.: Approach based on combination of vector neural networks for emitter identification. IET Sig. Proc. **4**(2), 137–148 (2010)
38. Ting, C., Wei, G., Bing, S.: A new radar emitter recognition method based on pulse sample figure. In: Proceedings of the 8th IEEE International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), pp. 1902–1905 (2011)
39. Wang, L., Ji, H., Shi, Y.: Feature extraction and optimization of representative-slice in ambiguity function for moving radar emitter recognition. In: Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 2246–2249 (2010)
40. Zeng, Y., Li, O.: A new algorithm for signal emitter recognition. In: Proceedings of the IEEE International Conference on Image Analysis and Signal Processing (IASP), pp. 446–449 (2010)
41. Pang, J., Lin, Y., Xu, X.: The improved radial source recognition algorithm based on fractal theory and neural network theory. Int. J. Hybrid Inf. Technol. **7**(2), 397–402 (2014)

42. Li, J., Ying, Y.: Radar signal recognition algorithm based on entropy theory. In: Proceedings of the 2nd IEEE International Conference on Systems and Informatics (ICSAI), pp. 718–723 (2014)

43. Ripley, B.D.: Pattern Recognition and Neural Networks (2008)

44. Kamgar-Parsi, B., Kamgar-Parsi, B., Sciortino Jr., J.C.: Automatic Data Sorting Using Neural Network Techniques. DTIC Document (1996)

45. Pape, D.R., Anderson, J.A., Carter, J.A., Wasilousky, P.A.: Advanced signal waveform classifier. In: Proceedings of the 8th Optical Technology for Microwave Applications, pp. 162–169 (1997)

46. Ince, T.: Polarimetric SAR image classification using a radial basis function neural network. PIERS **41**(4), 636–646 (2010)

47. Vicen-Bueno, R., Carrasco-Álvarez, R., Rosa-Zurera, M., Nieto-Borge, J.C., Jarabo-Amores, M.-P.: Artificial neural network-based clutter reduction systems for ship size estimation in maritime radars. EURASIP J. Adv. Sig. Process. **9**(1), 1–15 (2010)

48. Anjaneyulu, L., Sarma, N., Murthy, N.: Identification of LPI radar signals by higher order spectra and neural network techniques. Int. J. Inf. Commun. Technol. **2**(1), 142–155 (2009)

49. Ibrahim, N., Abdullah, R.R., Saripan, M.: Artificial neural network approach in radar target classification. J. Comput. Sci. **5**(1), 23 (2009)

50. Yin, Z., Yang, W., Yang, Z., Zuo, L., Gao, H.: A study on radar emitter recognition based on SPDS neural network. Inf. Technol. J. **10**(4), 883–888 (2011)

51. Zhang, Z.-C., Guan, X., He, Y.: Study on radar emitter recognition signal based on rough sets and RBF neural network. In: Proceedings of the IEEE International Conference on Machine Learning and Cybernetics, pp. 1225–1230 (2009)

52. Peipei, D., Hui, L.: The radar target recognition research based on improved neural network algorithm. In: Proceedings of the Fifth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA), pp. 1074–1077 (2014)

53. Chen, S., Cowan, C.F., Grant, P.M.: Orthogonal least squares learning algorithm for radial basis function networks. IEEE Trans. Neural Netw. **2**(2), 302–309 (1991)

54. Shieh, C.-S., Lin, C.-T.: A vector neural network for emitter identification. IEEE Trans. Antennas Propag. **50**(8), 1120–1127 (2002)

55. Lin, C.-M., Chen, Y.-M., Hsueh, C.-S.: A self-organizing interval type-2 fuzzy neural network for radar emitter identification. Int. J. Fuzzy Syst. **16**(1), 20 (2014)

56. Liu, J., Lee, J.P., Li, L., Luo, Z.-Q., Wong, K.M.: Online clustering algorithms for radar emitter classification. IEEE Trans. Pattern Anal. Mach. Intell. **27**(8), 1185–1196 (2005)

57. Li, L., Ji, H., Wang, L.: Specific radar emitter recognition based on wavelet packet transform and probabilistic SVM. In: Proceedings of the IEEE International Conference on Information and Automation (ICIA), pp. 1308–1313 (2009)

58. Ting, C., Jingqing, L., Bing, S.: Research on rough set-neural network and its application in radar signal recognition. In: Proceedings of the 8th International Conference on Electronic Measurement and Instruments (ICEMI), pp. 764–768 (2007)

59. Azimi-Sadjadi, M.R., Yao, D., Huang, Q., Dobeck, G.J.: Underwater target classification using wavelet packets and neural networks. IEEE Trans. Neural Netw. **11**(3), 784–794 (2000)

60. Lee, S.-J., Choi, I.-S., Cho, B., Rothwell, E.J., Temme, A.K.: Performance enhancement of target recognition using feature vector fusion of monostatic and bistatic radar. Prog. Electromagnet. Res. **144**, 291–302 (2014)

61. Baraldi, A.N., Enders, C.K.: An introduction to modern missing data analyses. J. Sch. Psychol. **48**(1), 5–37 (2010)

62. Enders, C.K.: Applied Missing Data Analysis. Guilford Press, New York (2010)

63. Graham, J.W.: Missing data analysis: making it work in the real world. Annu. Rev. Psychol. **60**(1), 549–576 (2009)

64. Horton, N.J., Lipsitz, S.R.: Multiple imputation in practice: comparison of software packages for regression models with missing variables. Am. Stat. **55**(3), 244–254 (2001)

65. Little, R.J., Rubin, D.B.: Statistical Analysis with Missing Data. Wiley, New York (2002)

66. Osborne, J.W., Overbay, A.: Best practices in data cleaning. Sage, Thousand Oaks (2012)

67. Rahm, E., Do, H.H.: Data cleaning: Problems and current approaches. IEEE Data Eng. Bull. **23**(4), 3–13 (2000)
68. Verboven, S., Branden, K.V., Goos, P.: Sequential imputation for missing values. Comput. Biol. Chem. **31**(5), 320–327 (2007)
69. Morring Jr., F., Perrett, B.: L-band SAR Satellite May Help JAXA's New Military Job, Aviation Week & Space Technology (2014)
70. O'Reilly, D., Bowring, N., Harmer, S.: Signal processing techniques for concealed weapon detection by use of neural networks. In: IEEE 27th Convention of Electrical & Electronics Engineers in Israel (IEEEI), pp. 1–4 (2012)
71. Lee, J.-H., Choi, I.-S., Kim, H.-T.: Natural frequency-based neural network approach to radar target recognition. IEEE Trans. Sig. Process. **51**(12), 3191–3197 (2003)
72. Khairnar, D., Merchant, S., Desai, U.: Radar signal detection in non-gaussian noise using RBF neural network. J. Comput. **3**(1), 32–39 (2008)
73. Briones, J., Flores, B., Cruz-Cano, R.: Multi-mode radar target detection and recognition using neural networks. Int. J. Adv. Rob. Syst. **9**(177), 1–9 (2012)