

A deep neural network based method for magnetic anomaly detection

Yizhen Wang¹  | Qi Han¹  | Guanyi Zhao¹ | Minghui Li² | Dechen Zhan¹ | Qiong Li¹

¹ School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

² School of International Studies, Harbin Institute of Technology, Harbin, China

Correspondence

Qi Han, School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China.

Email: qi.han@hit.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Number: 61771168

Abstract

Magnetic anomaly detection (MAD) is a technique to find ferromagnets hiding in strong and complicated magnetic background. In many practical cases, the targets are very far from the detection sensor, which leads to low signal-to-noise ratio (SNR) and high detection difficulty. Most of the current methods determine the existence of target by some approaches based on signal analysis, such as the orthogonal basis function (OBF) and the minimum entropy (ME). However, although these methods consume low resources, the detection performances are not satisfactory enough. In recent years, due to the increase of computer capability, complex methods become applicable in MAD. In this study, a deep neural network (DNN) is adopted to detect the magnetic anomalies. The DNN has shown its better ability to represent natural data in many applications. A feature automatically learned by a DNN from data in the raw form is more effective for detecting target signals and suppressing irrelevant variations. Herein, a convolutional network with residual structure to implement the feature extraction is designed and an MAD method based on it is proposed. Through the semi-real tests, the proposed method exhibits a strong capability to extract features and shows excellent performances on detection.

1 | INTRODUCTION

Magnetic anomaly detection (MAD) is a technique to find invisible ferromagnetic targets through measuring the magnetic field. This technique originates from the 1940s [1], and attracts increasing attention in recent years [2]. Nowadays, this technique has been widely used in the field of military, humanitarian aid, and industries [3].

In some applications, such as detecting urban pipelines [4, 5] and underground unexploded ordnances (UXOs) [6, 7], benefiting from the close detection distances, targets can cause strong magnetic anomalies. However, in some other applications, magnetometers are very far from targets, which leads to very low signal-to-noise ratios (SNRs) and make it very hard to determine the existences of targets. Most of the traditional MAD methods are mainly based on the signal processing technology, because they have low complexity and can satisfy the real-time requirement. One of the commonest methods is a matched filtering method based on the orthogonal basis functions (OBFs) [8, 9]. This method regards the target signal as a dipole and

decomposes it into orthogonal bases, then uses the bases as kernels to filter the collected signal. For this method, there is a limitation that the target model should be previously known, hence, some methods facing the background noise rather than the target model are presented, such as the minimum entropy (ME) detector [10, 11] and the stochastic resonance (SR) detector [12]. These methods have strong accessibility and can detect any type of anomaly signal. In general, both kinds of methods follow a common strategy, that is, designing a feature to represent the collected signal and using a threshold to determine whether there is a target. However, for signals with low SNRs, these features can hardly take effects. Moreover, a single feature which is designed manually is often insufficient to represent the magnetic field.

To improve the detection capability, deep neural network (DNN) has been adopted. In contrast to the traditional methods, a detector based on DNN automatically learns the features with multiple layers, which is important for detecting target signals and suppressing irrelevant variations [13]. In [14], Liu et al. propose a MAD method based on the fully connected neural

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *IET Science, Measurement & Technology* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

network classifier (called MLP method in this paper) with the inputs of OBF coefficients and signal variance. This method performs better than OBF method under colored Gaussian background noise, but it still relies on features extracted by traditional methods. In [15], the authors propose a MAD detector based on convolutional neural network (called CNN method in this paper). This method firstly convert the 1D signal into 2D using the Wigner–Ville time-frequency transform [16], and then adopt a network with 2 convolutional layers to be a detector. It yields excellent detection results. However, due to the complexity of the processing procedures, it consumes too much time. In [17], Xu et al. propose a fully convolutional network model (called FCN method in this paper) with 1D convolutional kernels. The whole network framework includes a classifier and a regressor to detect and denoise magnetic anomaly, respectively. This method achieves better results than OBF method on semi-real data.

In this paper, we design an end-to-end DNN-based detector directly on 1D signal, which has a better detection ability than other methods. The proposed detector is mainly formed by 1D convolutional layers and contains residual structures. We have tested this detector on the semi-real datasets, and compared it with three other DNN-based detectors and three traditional detectors. The training dataset and the test dataset use different background data. The detection results show that the trained DNN-based detector has higher accuracy and detection rate (DR) than other detectors, and the time cost is low enough to be applied in a real-time MAD system.

2 | DETECTION METHOD

2.1 | Magnetic target model

The source of the ferromagnetic target's magnetic field is usually very complex, so it is difficult to model it accurately [18–20]. However, when the distance from the target to the magnetometer is long enough, a magnetic target can usually be regarded as a magnetic dipole [21]. The magnetic target model is shown in Figure 1, the target is located at the origin O , and the X , Y , and Z axes point to the magnetic east, the magnetic north, and the upward direction, respectively. The detection platform equipped with the magnetometer moves along the path at a speed v . \mathbf{R}_0 is the closest distance from the target to the platform trajectory, and is called the closest proximity approach (CPA). E is the intersection point of \mathbf{R}_0 and the platform track.

The magnetic field \mathbf{B} generated by a magnetic target can be expressed as

$$\mathbf{B} = \frac{\mu_0}{4\pi} \left[\frac{3(\mathbf{M} \cdot \mathbf{R})\mathbf{R}}{|\mathbf{R}|^5} - \frac{\mathbf{M}}{|\mathbf{R}|^3} \right], \quad (1)$$

where μ_0 is the permeability of vacuum, \mathbf{M} is the target magnetic moment, \mathbf{R} is the distance from the magnetometer to the target.

In practical applications, the target's magnetic field is much smaller than the Earth's magnetic field \mathbf{B}_e [22], that is, $|\mathbf{B}| \ll |\mathbf{B}_e|$.

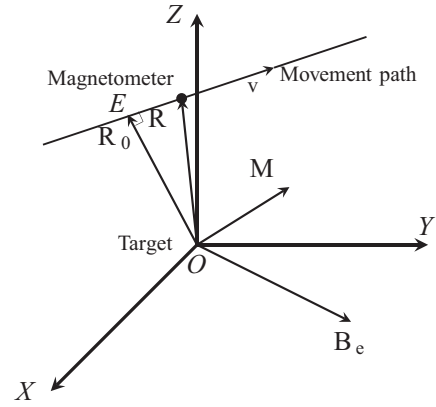


FIGURE 1 Configuration of magnetic target detection by a moving magnetometer.

\mathbf{B}_e . Consequently, a target signal S measured by a scalar magnetometer can be regarded as a projection of \mathbf{B} along the direction of the geomagnetic field \mathbf{B}_e [22]:

$$S = \frac{\mathbf{B} \cdot \mathbf{B}_e}{|\mathbf{B}_e|}, \quad (2)$$

$$= \frac{\mu_0}{4\pi |\mathbf{B}_e|} \left[\frac{3(\mathbf{M} \cdot \mathbf{R})(\mathbf{R} \cdot \mathbf{B}_e)}{|\mathbf{R}|^5} - \frac{\mathbf{M} \cdot \mathbf{B}_e}{|\mathbf{R}|^3} \right]. \quad (3)$$

Suppose η , ϕ , and ψ_0 are the angles between \mathbf{M} and \mathbf{R} , \mathbf{R} and \mathbf{B}_e , \mathbf{M} and \mathbf{B}_e , respectively, Equation (3) can be rewritten as

$$S = \frac{M\mu_0}{4\pi} \left[\frac{3 \cos \eta \cos \phi - \cos \psi_0}{|\mathbf{R}|^3} \right]. \quad (4)$$

The distance vector \mathbf{R} is the vectorial sum of the distance from E to the measuring point and the CPA, which can be expressed as

$$\mathbf{R} = \mathbf{R}_0 + \mathbf{v}(t - t_0). \quad (5)$$

From Equations (4) and (5), it can be known that the target signal type is controlled by \mathbf{R} , \mathbf{v} , η , ϕ , and ψ_0 . The three angles can be calculated from the direction of \mathbf{M} , \mathbf{R} , and \mathbf{B}_e , and \mathbf{R} is related to \mathbf{R}_0 and \mathbf{v} .

2.2 | Data representation

The total-field signal used in MAD can be represented as $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$, where x_i ($i = 1, 2, \dots, n$) is a sampling point. The sampling rate is written as F_s . Traditionally, an MAD method transforms \mathbf{x} into a new series $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$, namely the feature (e.g. the feature adopted by ME detector is the information entropy). The feature \mathbf{y} often has a higher SNR than \mathbf{x} and makes the target more noticeable. Afterwards, a threshold T_b , often confirmed on the basis of the priori knowledge or the experience, is used to classify all the y_i ($i = 1, 2, \dots, n$) into two categories. We use v_i to label whether the corresponding y_i

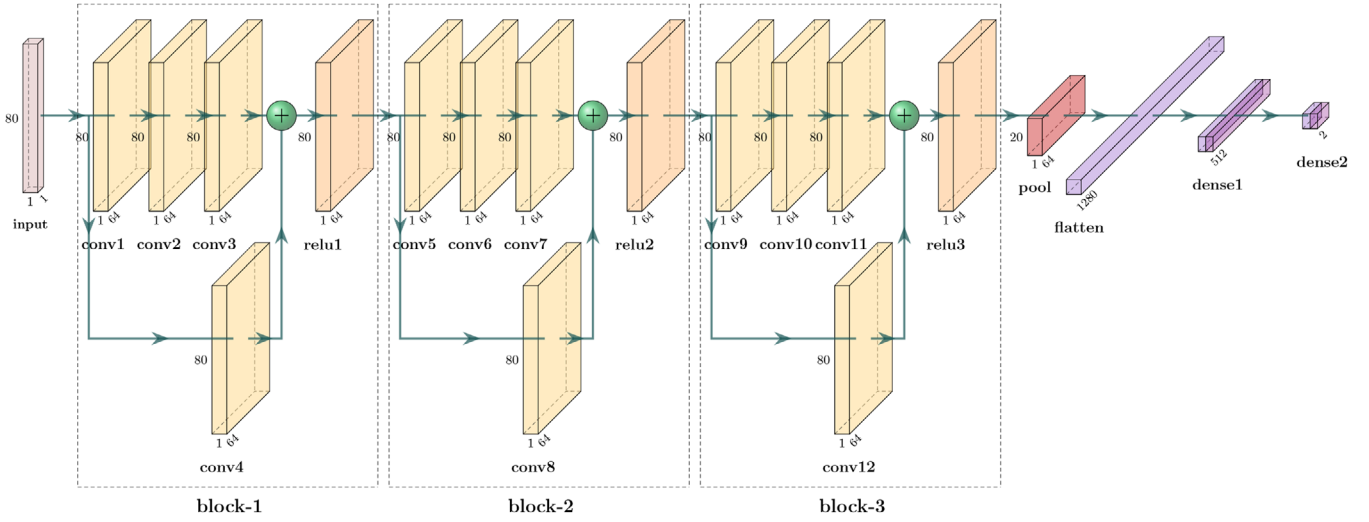


FIGURE 2 The architecture of MADRN (MAD residual network)

reaches T_b :

$$v_i = \begin{cases} 1, & y_i \geq (\text{or } \leq) T_b \\ 0, & y_i < (\text{or } >) T_b \end{cases}, \quad (i = 1, 2, \dots, n). \quad (6)$$

However, this point-to-point representation is not very direct, because discovering an entire target needs a signal segment rather than only a point. Those anomalous points which reach the threshold accidentally will be labeled as targets by the detector and cause the increase of false alarm rate (FAR). Therefore, we design a sequence-to-point type to represent the collected signals.

Let \mathbf{X}_i represent a sequence of sampling points as

$$\mathbf{X}_i = [x_{i1}, x_{i2}, \dots, x_{im}]^T \in \mathbb{R}^m (i = 1, 2, \dots, N), \quad (7)$$

where m is the length, and N is the amount of \mathbf{X}_i . For each \mathbf{X}_i , we define a corresponding binary pair $\mathbf{V}_i \in \{0, 1\}^2$ as

$$\mathbf{V}_i = \begin{cases} [1, 0]^T, & \text{if there is no target in sample } \mathbf{X}_i \\ [0, 1]^T, & \text{otherwise} \end{cases}, \quad (8)$$

where $i = 1, 2, \dots, N$. If we regard \mathbf{X}_i as a sample and \mathbf{V}_i as a label, the signal \mathbf{x} is represented by the type of $\mathbf{D} = \{(\mathbf{X}_i, \mathbf{V}_i), i = 1, 2, \dots, N\}$, where \mathbf{D} is regarded as dataset in this paper. This representation is useful to locate the target and reduce the FAR.

2.3 | Feature extraction

In this work, we use convolutional layers to extract features and use fully connected layers to implement classification. For

TABLE 1 The detailed network settings

Layers	Parameter settings
conv1(5, 9)	Kernel number: 64; kernel size: 8×1 ; strides: 1; padding: "same"
conv2(6, 10)	Kernel number: 64; kernel size: 8×1 ; strides: 1; padding: "same"
conv3(7, 11)	Kernel number: 64; kernel size: 8×1 ; strides: 1; padding: "same"
conv4(8, 12)	Kernel number: 64; kernel size: 1×1 ; strides: 1; padding: "same"
pool	Maxpooling; pool size: 4×1 ; strides: 4; padding: "same"
dense1	Activate function: "relu"
dense2	Activate function: "softmax"

convolutional layers, we adopt residual structure, which is proposed in [23] and facilitates training deeper network [24].

The entire architecture of the network applied in this work is shown in Figure 2, where the structures of residual blocks (res-blocks) are shown in dashed boxes. There are 3 res-blocks in this network. In each res-block, there are 3 stacked convolutional layers with 8×1 convolutional kernels and an additional convolutional layer with 1×1 convolutional kernels. The activation function is the relu function. For convenience, we call this network as MADRN (MAD residual network). The detailed parameter settings for convolution layers, pooling layer and full connection layers in MADRN are shown in Table 1, where we take the convolution layers in res-block1 as an example, the convolution layers in res-block2 and res-block3 are annotated in brackets.

To confirm the parameters mentioned in the previous paragraph, we did a additional test to compare the performances of networks with different structure parameters. The results are shown in Table 2, where n represents the number of res-blocks, m represents the number of kernels in each convolutional

TABLE 2 The areas under the curves (AUCs) and time consumptions (TCs) of networks with different parameters

Parameters	AUC ^{train}	AUC ^{test}	TC (ms)
$n = 2, m = 64$	0.9997	0.9965	~1.81
$n = 3, m = 64$	1	0.9992	~1.85
$n = 4, m = 64$	0.9999	0.9987	~2.46
$n = 3, m = 128$	1	0.9988	~2.96

layer. In this test, the criteria are the area under the receiver operating characteristic (ROC) curve (AUC) and the time consumption (TC) of detecting a sample. The datasets will be described in the next section. From this table we can find that networks with 3 res-blocks have better classification capability, and increasing the number of kernels from 64 to 128 is adverse to the classification performance of the test data and is time-consuming.

A well-trained MADRN has a strong ability to represent the samples. We display the outputs of the three blocks in our trained network in Figure 3. Those three charts in the first row show the original samples. They are the strong target sample, the weak target sample, and the pure background sample, respectively. Corresponding to these samples, those three charts in the second row display the outputs of the first res-block (block-1). The outputs of the second and third res-blocks (block-2 and block-3) are shown in the third and fourth rows of Figure 3.

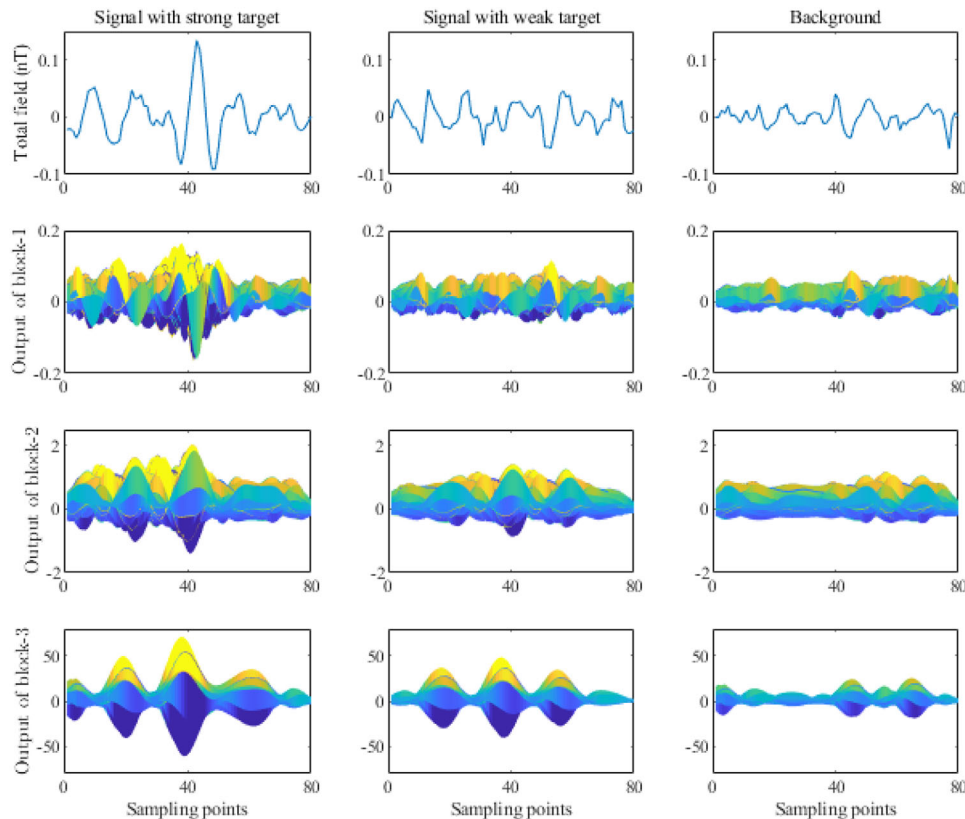
For better illustration, we unify the axis-scales of the charts in the same row. From this figure, we can find that the strength of the weak target is roughly equal to the background; yet the res-blocks' outputs of these two types of signals are different. As the res-blocks' depth in the network increases, the difference among the extracted feature becomes more obvious. It indicates that the automatically extracted features have the ability to represent the target signals and the potentiality to detect them.

2.4 | Training and decision

Based on the features extracted by the res-blocks, some fully connected (fc) layers are designed for classification. The fc layers have 1280, 512, and 2 neurons, respectively. Between the res-blocks and the fc layers, a max-pooling layer [25, 26] is used to reduce the number of neurons and avoid overfitting.

We utilize the Adam optimizer [27] and the cross-entropy loss function to train the network. All the former layers are activated by the relu function, except the last layer followed by the softmax activation function.

The trained network can be used to decide whether there is a target in the inputted sample. For each sample, the final output of the network is a pair of real numbers, written as $\mathbf{P}_i = [p_i, q_i]^T$, where $0 \leq p_i, q_i \leq 1$, and $p_i + q_i = 1$. Each number represents the probability that the sample \mathbf{X}_i belongs to the corresponding category. It is a common practice to map \mathbf{P}_i into $\{0, 1\}^2$ through rounding. According to (3), if the rounded \mathbf{P}_i

**FIGURE 3** Samples and their features extracted by the trained network

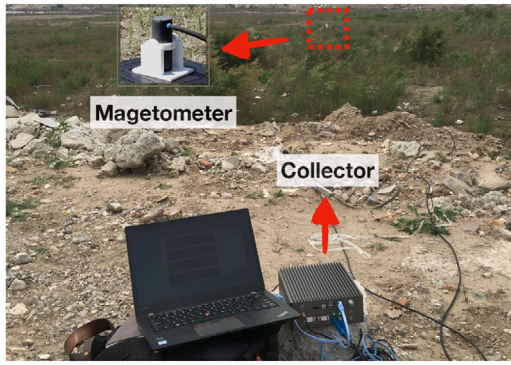


FIGURE 4 The ambient magnetic field collection system

equals to $[1, 0]^T$, there is no target in \mathbf{X}_i , and if the rounded \mathbf{P}_i equals to $[0, 1]^T$, there is a target in \mathbf{X}_i .

3 | TESTS

3.1 | Datasets

Semi-real data is often used in MAD experiments, which usually consists of the real ambient magnetic field and the simulated target magnetic field [10–12, 15]. In our work, the influence of the detection platform magnetic field was also taken into account.

3.1.1 | Background magnetic field

The background magnetic field is the sum of the platform magnetic field, the geomagnetic field, and the diurnal variation of the Earth's magnetic field.

The platform magnetic field is generated according to the simulation method in [28]. In our work, we assume that the detection platform is an aircraft and use the X-Plane flight simulator to get its motion and position information. We establish the distribution model of magnetic dipoles, soft iron, and conductive coils on the aircraft geometry. Based on this model and aircraft motion information, we generate the permanent, inductive, and eddy magnetic fields of the aircraft.

The geomagnetic field is obtained using aircraft position information and the Enhanced Magnetic Model 2017 (EMM2017) [29]. The EMM2017 is a degree and order 790 internal field model with high resolution, which mainly can describe the characteristics of the long-wavelength part of the Earth's internal magnetic field.

The diurnal variation of the Earth's magnetic field is simulated by subtracting the DC term from the ambient magnetic field. The ambient magnetic field signals are real data measured by an optically pumped cesium vapor magnetometer and collected by a homemade collector, shown in Figure 4. The sensitivity of the cesium vapor magnetometer is lower than $0.002\text{nT}/\sqrt{\text{Hz}}@1\text{Hz}$; its principle can be seen in [30–32]. The range of measured magnetic field is about 55001 nT to 55031 nT , and the sampling rate is 10 Hz .

TABLE 3 Parameters for target signal simulation

Parameters	Values
Local geomagnetic direction	$(0.86603, 0, -0.5)$
Speed of magnetometer	100 m/s
Target moment	$400000\text{ A} \cdot \text{m}^2$
Target moment direction	$[0^\circ, 360^\circ]$
Magnetometer moving direction	$[0^\circ, 360^\circ]$
CPA	$[450\text{ m}, 650\text{ m}]$

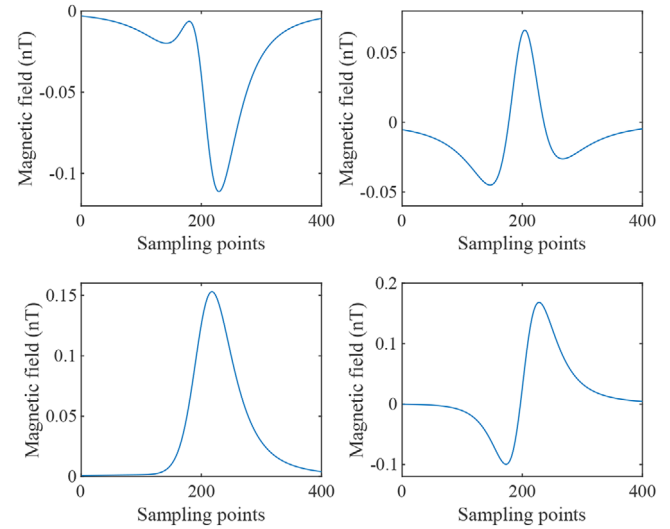


FIGURE 5 Some target signals with different shapes and strengths

3.1.2 | Target magnetic field

The target magnetic field is generated based on Section 2.1. According to Equations (4) and (5), the shape of simulated target signal is related to the magnetometer moving direction, the target moment direction and the local geomagnetic direction, while the strength of the simulated target signal is related to platform speed and the CPA.

In our work, we fixed the local geomagnetic direction and the platform speed, and randomly selected the other three factors in certain ranges. The values of target simulation parameters are listed in Table 3. Using different combinations of parameters, we produce dipole signals with different shapes and strengths. The strength ranges from about 0.05 nT to 0.6 nT . Figure 5 shows some typical target signals with different shapes and strengths by changing parameters of the target model.

3.1.3 | Dataset preparation

Based on the background magnetic field and the target magnetic field, we build a dataset for the MADRN network. We generate 13,000 background magnetic fields, of which 10,000 are used to prepare the training dataset and the remaining 3000 are used to prepare the test dataset. From all samples, we randomly selected

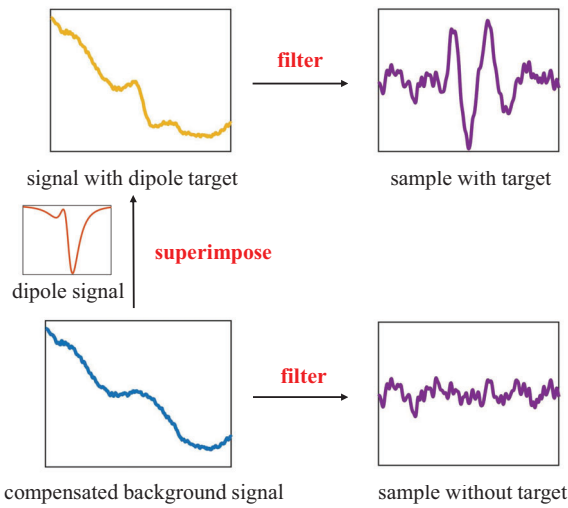


FIGURE 6 The dataset producing procedures

some of them to superimpose the target magnetic field. As mentioned in [33, 34], a technique called compensation is often used in MAD to improve detection performance. For each sample, we use the compensation method proposed in [28] to reduce the aircraft interference and then use a band-pass filter to remove the DC term and trend term. Figure 6 shows the procedures of producing a sample with or without the target. The first row of Figure 3 shows different samples.

3.1.4 | Training with the dataset

With the sampling rate of 10 Hz, there are 400 sampling points in each sample. To reduce the amount of calculation, each sample has been down-sampled to 2 Hz before being inputted into the network. Without this process, the network should be scaled up largely to yield the roughly equal ability, which is time-consuming.

During the training process, 30% of training dataset are selected to be a validation set. The model with the lowest loss on the validation set is saved to be the trained model. The network is trained for 2000 epochs, and has been already well converged after about 1000 epochs, shown in Figure 7.

3.2 | Comparisons

In this subsection, we compare the MADRN detector with six methods mentioned above, including MLP detector, FCN detector, 2D CNN detector, SR detector, OBF detector, and ME detector. We test all the seven methods on the same dataset that contains targets with different types.

3.2.1 | ACCs, DRs, and FARs

In Table 4, we compare the detection results of the seven methods in the test dataset. As usual, the thresholds of the four

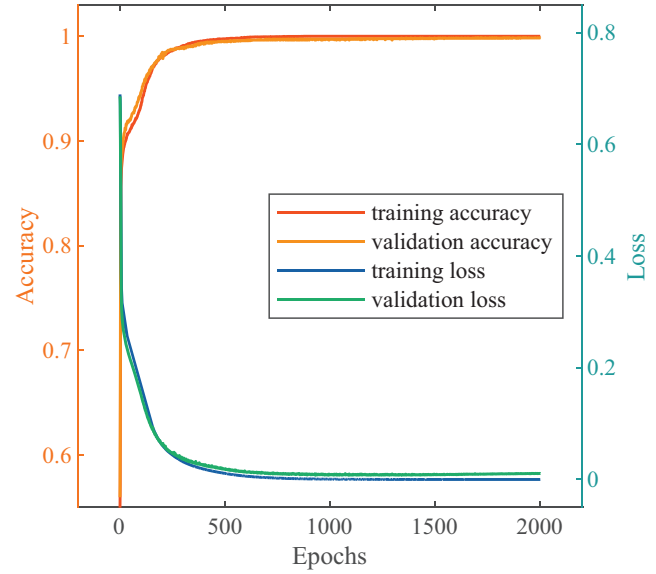


FIGURE 7 The accuracies and losses in the training procedure

TABLE 4 The comparison results

Method	ACC (%)	DR (%)	FAR (%)	DR (%)	TC (ms)
				@10%FAR	
MADRN	98.80	98.28	0.67	99.93	~1.45
CNN	94.17	94.59	6.09	97.03	~1.45
FCN	90.04	89.05	8.29	47.90	~1.43
MLP	76.87	70.45	18.72	57.57	~0.43
SR	65.50	70.84	39.96	11.74	~0.04
ME	68.57	62.14	24.87	7.19	~0.01
OBF	73.83	65.24	17.39	16.75	~0.04

DNN-based detectors are set as 0.5 (namely, rounding the network outputs). Besides, to exhibit the best performances of the traditional methods, their thresholds are adjusted to make them achieve the highest accuracies (ACCs). At this situation, MADRN method provides an ACC of 98.80%, a DR of 98.28%, and an FAR of 0.67%, better than other methods. The 2D CNN detector, FCN detector, and MLP detector also provide better results than the traditional methods; however, the improvement of MLP detector is not very obvious. To avoid the interferences of setting thresholds, we compare the DRs at the same FAR of 10%. The MADRN method still gives a higher DR. Herein, the ACC, DR, and FAR are defined as [35, 36]

$$\text{ACC} = \frac{S_D + S_N}{S} \quad \text{DR} = \frac{S_D}{S_T} \quad \text{FAR} = \frac{S_F}{S}, \quad (9)$$

where S is the total number of samples, S_D is the number of detected target samples, S_N is the number of samples without target and alarm, S_T is the number of total real target samples, S_F is the number of false alarm samples.

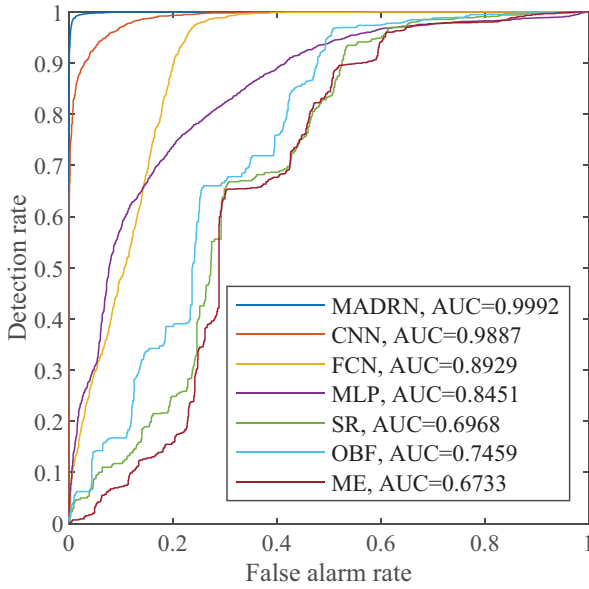


FIGURE 8 The receiver operating characteristic (ROC) curves and the areas under the curves (AUCs) of the seven detection methods

TABLE 5 The time complexity of seven detectors

Detector	Time complexity
MADRN	$4.502 \times 10^7 \text{ FLOPs}$
CNN	$1.173 \times 10^9 \text{ FLOPs}$
FCN	$2.317 \times 10^7 \text{ FLOPs}$
MLP	$O(n) + 986 \text{ FLOPs}$
SR	$O(n)$
ME	$O(nm)$
OBF	$O(nm)$

3.2.2 | ROC curves and AUCs

Figure 8 shows the ROC curves of the seven methods and gives their respective AUC values.

We can find that MADRN detector and 2D CNN detector are much better than other five methods, and the results of FCN detector and MLP detector are next. The AUC of MADRN detector is extremely close to 1, and the AUC of 2D CNN is 0.9887. The AUCs of FCN detector and MLP detector are both larger than 0.8. Among the three traditional methods, the OBF method yields the best result, the performance of the SR method is better than the ME method. This result denotes that the MADRN method has a better ability to correctly classify the samples.

3.2.3 | TCs

According to [37, 38], the time complexity of seven detectors are listed in Table 5, where n is the window length.

The tested TCs of detecting a sample are listed in Table 4. Although the TC of MADRN is longer than the traditional

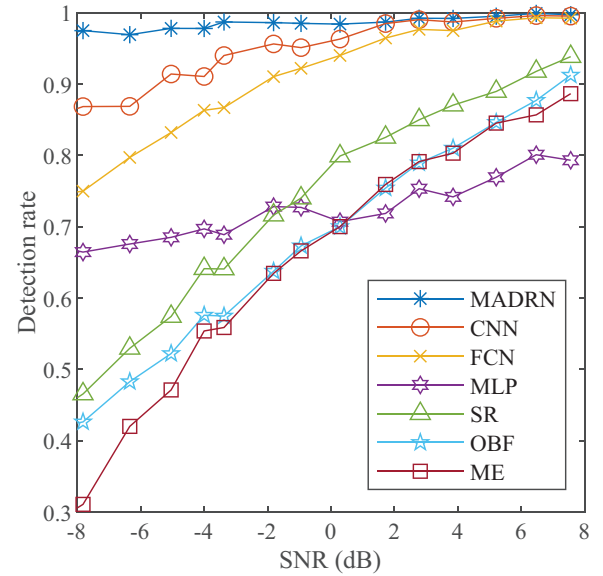


FIGURE 9 The detection rates (DRs) of these detectors on datasets with different signal-to-noise rates (SNRs)

methods, it is much shorter than the typical sampling interval of 100 ms (10 Hz sampling rate). This result denotes that the proposed MADRN method can be implemented in real-time systems. The TCs are calculated using a computer with the CPU of Intel Core i7-8700K and a GPU of NVIDIA GeForce 1080Ti.

3.2.4 | DRs with different SNRs

For further comparing these seven detectors, we test their DRs on different SNRs, which are calculated by the following equation [12]:

$$\text{SNR} = \frac{1}{M} \sum_{i=1}^M 10 \log_{10} \frac{\sum_{j=1}^N s_i(j)^2}{\sum_{j=1}^N n_i(j)^2}, \quad (10)$$

where s_i is a sampling point of the target signal, n_i is a sampling point of the corresponding background noise, N is the length of these sequences, and M is the amount of the samples in a dataset. The datasets with specific SNRs are prepared by controlling the CPA value.

Figure 9 shows the results. The MADRN detector always yields higher DRs than other methods on datasets with different SNRs. The 2D CNN detector and FCN detector also achieve good results. Among the three traditional methods, the OBF detector has higher DRs than the other two, in most cases. DRs of ME detector are lower. The DRs of the MLP detector change little in the test datasets with different SNRs, and always between 0.65 and 0.75. It maybe because the features extracted by MLP detector with only 4 input neurons are not representative enough. In brief, the results in Figure 9 indicates that the MADRN detector is robust enough to resist the noise.

3.2.5 | Summary

From the above comparisons, it can be seen that MADRN, 2D CNN, and FCN detectors are obviously better than other methods. Among them, MADRN detector performs best in detection accuracy and robustness, but it has a bit higher computational complexity than FCN detector. The detection capability of 2D-CNN is second only to MADRN, but the computational complexity of this method is much higher than others because the signal needs to be converted into 2D images first. The FCN detector is also an end-to-end network based on 1D convolution. However, the DR of FCN detector decreases significantly when the SNR is low. The MLP detector performs worse than other DNN-based methods, mainly because of its weaker capability to extract features, which might be caused by the dependency on traditional methods and the simple network structure. The other three traditional methods yield similar results. Their detection performances are significantly weaker than the DNN-based methods, but the time consumptions are very low.

4 | DISCUSSION

In this paper, aiming at detecting a faraway target with a moving magnetometer (onboard usually), we propose a MAD method based on a 1D convolutional network with residual structures. As a DNN-based method, this detector can automatically extract the features after being well trained. Under the same conditions, we compare our method with three other DNN-based detectors and three traditional detectors. The test results show that the DR of our algorithm is 98.28%, the FAR is 0.67, and the AUC is 0.9992. These results are very outstanding, especially in the cases with low SNRs. Besides, our proposed method detects a sample in only about 1.45 ms, satisfies the real-time requirement.

In our opinion, besides the high DR, the MADRN detector also has some unverified highlights: (1) the application of the MADRN is not limited by the target type or the noise strength only if the training dataset contains enough samples to represent these situations; (2) a trained MADRN can be easily transferred to another similar task to reduce the training cost, because the convolutional kernels are not related to the sample size.

In the future, we plan to continue this research with the following two items: (1) detecting the target shape, orientation and other specific properties; (2) researching the unsupervised detection.

ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (Grant No. 61771168).

CONFLICT OF INTEREST

The authors have declared no conflict of interest.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Yizhen Wang  <https://orcid.org/0000-0003-0312-3439>

Qi Han  <https://orcid.org/0000-0001-9432-3131>

REFERENCES

1. Fromm, W.E.: The magnetic airborne detector. In: *Advances in Electronics and Electron Physics*, pp. 257–299. Elsevier, New York (1952)
2. Gao, X., Yan, S., Li, B.: A novel method of localization for moving objects with an alternating magnetic field. *Sensors* 17(4), 923 (2017)
3. Dames, P.M., Schwager, M., Rus, D., et al.: Active magnetic anomaly detection using multiple micro aerial vehicles. *IEEE Robot. Automation Lett.* 1(1), 153–160 (2015)
4. Zhao, D., Guo, Z., Du, J., et al.: Geometric modeling of underground ferromagnetic pipelines for magnetic dipole reconstruction-based magnetic anomaly detection. *Petroleum* 6(2), 189–197 (2020)
5. Sheinker, A., Moldwin, M.B.: Magnetic anomaly detection (MAD) of ferromagnetic pipelines using principal component analysis (PCA). *Meas. Sci. Technol.* 27(4), 045104 (2016)
6. Gang, Y., Yingtang, Z., Hongbo, F., et al.: Detection, localization and classification of multiple dipole-like magnetic sources using magnetic gradient tensor data. *J. Appl. Geophys.* 128, 131–139 (2016)
7. Butler, D.K., Simms, J.E., Furey, J.S., et al.: Review of magnetic modeling for uxo and applications to small items and close distances. *J. Environ. Eng. Geophys.* 17(2), 53–73 (2012)
8. Ginzburg, B., Frumkis, L., Kaplan, B.Z.: Processing of magnetic scalar gradiometer signals using orthonormalized functions. *Sens. Actuators A* 102(1–2), 67–75 (2002)
9. Ginzburg, B., Frumkis, L., Kaplan, B.Z.: An efficient method for processing scalar magnetic gradiometer signals. *Sens. Actuators A* 114(1), 73–79 (2004)
10. Sheinker, A., Salomonski, N., Ginzburg, B., et al.: Magnetic anomaly detection using entropy filter. *Meas. Sci. Technol.* 19(4), 045205 (2008)
11. Tang, Y., Liu, Z., Pan, M., et al.: Detection of magnetic anomaly signal based on information entropy of differential signal. *IEEE Geosci. Remote Sens. Lett.* 15(4), 512–516 (2018)
12. Wan, C., Pan, M., Zhang, Q., et al.: Magnetic anomaly detection based on stochastic resonance. *Sens. Actuators A* 278, 11–17 (2018)
13. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* 521(7553), 436 (2015)
14. Liu, S., Chen, Z., Pan, M., et al.: Magnetic anomaly detection based on full connected neural network. *IEEE Access* 7, 182198–182206 (2019)
15. Zhang, K., Hu, M., Du, C., et al.: Detection of magnetic dipole target signals by using convolution neural network. In: *2018 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC)*, pp. 1–3. IEEE, Piscataway (2018)
16. Cohen, L.: Time-frequency distributions-a review. *Proc. IEEE* 77(7), 941–981 (1989)
17. Xu, X., Huang, L., Liu, X., et al.: Deepmad: Deep learning for magnetic anomaly detection and denoising. *IEEE Access* 8, 121257–121266 (2020)
18. Holmes, J.J.: Exploitation of a ship's magnetic field signatures. *Synth. Lect. Comput. Electromagn.* 1(1), 1–78 (2006)
19. Holmes, J.J.: Modeling a ship's ferromagnetic signatures. *Synth. Lect. Comput. Electromagn.* 2(1), 1–75 (2007)
20. McFee, J.E., Das, Y., Ellingson, R.O.: Locating and identifying compact ferrous objects. *IEEE Trans. Geosci. Remote Sens.* 28(2), 182–193 (1990)
21. Sheinker, A., Ginzburg, B.: Magnetic anomaly detection using high-order crossing method. *IEEE Trans. Geosci. Remote Sens.* 50(4), 1095–1103 (2012)
22. Balaji, B., Nelson, J.B.: Parameter estimation and tracking of a magnetic dipole. *Proc. SPIE* 9091, 90910A (2014)
23. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778. IEEE, Piscataway (2016)
24. Wei, Y., Yuan, Q., Shen, H., et al.: Boosting the accuracy of multispectral image pansharpening by learning a deep residual network. *IEEE Geosci. Remote Sens. Lett.* 14(10), 1795–1799 (2017)

25. LeCun, Y., Bottou, L., Bengio, Y., et al.: Gradient-based learning applied to document recognition. *Proc. IEEE* 86(11), 2278–2324 (1998)
26. Boureau, Y.L., Bach, F., LeCun, Y., et al.: Learning mid-level features for recognition. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2559–2566. IEEE, Piscataway (2010)
27. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. *arXiv preprint, arXiv: 1412.6980* (2014)
28. Dou, Z., Han, Q., Niu, X., et al.: An aeromagnetic compensation coefficient-estimating method robust to geomagnetic gradient. *IEEE Geosci. Remote Sens. Lett.* 13(5), 611–615 (2016)
29. The Enhanced Magnetic Model. <https://www.ngdc.noaa.gov/geomag/EMM/>. Accessed 9 April 2021
30. Ness, N.F.: Magnetometers for space research. *Space Sci. Rev.* 11(4), 459–554 (1970)
31. Yabuzaki, T., Ogawa, T.: Frequency shifts of self-oscillating magnetometer with cesium vapor. *J. Appl. Phys.* 45(3), 1342–1355 (1974)
32. Hardwick, C.D.: Non-oriented cesium sensors for airborne magnetometry and gradiometry. *Geophysics* 49(11), 2024–2031 (1984)
33. Hood, P.: History of aeromagnetic surveying in Canada. *The Leading Edge* 26(11), 1384–1392 (2007)
34. Noriega, G., Marszalkowski, A.: Adaptive techniques and other recent developments in aeromagnetic compensation. *First Break* 35(9), 31–38 (2017)
35. Chang, S., Du, B., Zhang, L.: Baso: A background-anomaly component projection and separation optimized filter for anomaly detection in hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* 56(7), 3747–3761 (2018)
36. Du, B., Zhang, L.: A discriminative metric learning based anomaly detection method. *IEEE Trans. Geosci. Remote Sens.* 52(11), 6844–6857 (2014)
37. He, K., Sun, J.: Convolutional neural networks at constrained time cost. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5353–5360. IEEE, Piscataway (2015)
38. Molchanov, P., Tyree, S., Karras, T., et al.: Pruning convolutional neural networks for resource efficient inference. *arXiv preprint, arXiv:1611.06440* (2016)

How to cite this article: Wang, Y., Han, Q., Zhao, G., Li, M., Zhan, D., Li, Q.: A deep neural network based method for magnetic anomaly detection. *IET Sci. Meas. Technol.* 16, 50–58 (2022).
<https://doi.org/10.1049/smt2.12084>