Exo-atmospheric infrared objects classification using recurrence-plots-based convolutional neural networks

QIUQUN DENG,* HUANZHANG LU, MOUDA HU, AND BENPONG ZHAO

National Key Laboratory of Science and Technology on ATR, College of Electronic Science, National University of Defense Technology (NUDT), Changsha 410073, China
*Corresponding author: dengqiuqun75@163.com

Received 11 September 2018; revised 11 November 2018; accepted 21 November 2018; posted 21 November 2018 (Doc. ID 344358); published 21 December 2018

Object discrimination plays an important role in infrared (IR) imaging systems. However, at long observing distance, the presence of detector noise and absence of robust features make exo-atmospheric object classification difficult to tackle. In this paper, a recurrence-plots-based convolutional neural network (RP-CNN) is proposed for feature learning and classification. First, it uses recurrence plots (RPs) to transform time sequences of IR radiation into two-dimensional texture images. Then, a CNN model is adopted for classification. Different from previous object classification methods, RP representation has well-defined visual texture patterns, and their graphical nature exposes hidden patterns and structural changes in time sequences of IR signatures. In addition, it can process IR signatures of objects without the limitation of fixed length. Training data are generated from IR irradiation models considering micro-motion dynamics and geometrical shape of exo-atmospheric objects. Results based on time-evolving IR radiation data indicate that our method achieves significant improvement in accuracy and robustness of the exo-atmospheric IR objects classification.

https://doi.org/10.1364/AO.58.000164

1. INTRODUCTION

Discriminating exo-atmospheric objects using IR sensors is a key technology for precise guidance systems and IR surveillance systems [1]. In the case of long observing distance, objects emerge as small dots on the IR image plane, lacking shape and attitude (i.e., orientation) information. Those objects include decoys owning similar IR radiation intensity as the target, which poses a great challenge to discrimination tasks. Fortunately, the time sequence of the IR signature can be used, providing additional information for the classification problem. Thus, the problem of IR objects classification can be converted into classification of the time sequence of IR signatures.

In recent years, object discrimination based on the time sequence of the IR signature has been a hot topic, and different classification methods have been presented [2–4, 6–8]. Most work focuses on feature extraction, which is of vital importance for object classification. Silberman [2] extracted several statistical features of input signatures, e.g., mean and variance, to build a classifier for ballistic targets. The analysis of objects IR signatures showed that different objects may possess different temperatures and cool at varying rates [3]. Wang et al. [4] proposed a probabilistic neural network for exo-atmospheric target discrimination using the temporal evolvement characteristics of temperature and emissivity-area products as input of the neural network. Temperature features can be extracted through the radiation ratio of two different wavelengths. Besides the temperature feature, micro-motion features and geometrical shape also serve as important features. Exo-atmospheric objects are always in rotational or vibrational motion, referred to as micro-motion, until they re-enter the atmosphere [5]. Micro-motion of objects leads to periodic fluctuations of the time sequence of the IR signature [6]. Although the shape of an object cannot be resolved by IR imaging, different shapes of micro-motion objects can induce different time sequences of the IR signature [5]. Wu et al. [7] and Liu et al. [8] estimated micro-motion and shape parameters based on the models they proposed, respectively. However, these approaches are limited by the assumptions of specific parameters; when conditions change, the performance of these handcrafted features obviously decreases. Feature extraction of exo-atmospheric IR objects remains challenging.

Deep learning techniques have recently achieved impressive results in a variety of domains [9–19], which provide a new perspective to the problem mentioned above. Instead of extracting handcrafted features, we can adopt a deep neural network for feature learning from the raw data directly. As one of the deep neural networks, the convolutional neural network (CNN) has been successfully applied to space object
classification [20] and time series classification [21–23]. Inspired by these results, this paper explores CNNs for the exo-atmospheric IR objects classification problem. IR signatures are generated by dynamical systems, which are caused mainly by motion, including orbital motion and micro-motion, and are susceptible to detector noise. The challenge of the application of CNNs is to explore the framework that is appropriate for learning invariant features in a dynamical system and is robust to noise.

Based on properties of the IR signature, a recurrence-plots-based CNN (RP-CNN) is proposed for exo-atmospheric IR objects classification. The recurrence plot (RP), introduced in the late 1980s [24], is a powerful tool for visualization and analysis of recurrence, which enables us to investigate certain aspects of the m-dimensional phase space trajectory through 2D representation [25–27]. This representation has well-defined visual texture patterns, and their graphical nature exposes hidden patterns and structural changes in data. It can exhibit not only a global characteristic (typology) but also local patterns (texture). Traditionally, measures such as recurrence rate, determinism, mean diagonal line length, and entropy are used to quantify the RP matrix [28,29]. Silva et al. [30] proposed treating plots as gray-level texture images and applied a support vector machine classifier on texture features. In this work, we first transform time sequences of IR signatures into a texture image using RP, and then treat exo-atmospheric IR objects classification as a texture recognition task. Finally, a two-hidden-layers CNN model is applied for classification.

The rest of this paper is organized as follows. We first simulate IR signatures of exo-atmosphere micro-motion objects in Section 2. The IR objects classification framework is presented in Section 3, and experimental results and discussion are presented in Section 4. Concluding remarks are provided in the last section.

2. SIMULATING IR SIGNATURES

Due to time consumption, high expense, and limited number of experiments depending only on measurement, it is inconvenient to analyze the IR signature of exo-atmospheric objects under various complex conditions. Thus, simulation is an alternative method [6], which provides an effective research facility to analyze the IR signature with more flexibility and lower costs. There are several models used for simulating IR signature measurements in literature, and Ref. [8] provides a relative comprehensive consideration for exo-atmospheric IR objects applications. This model is based on geometrical optics and radiation theory and considers a wide range of objects’ physical attributes and dynamic states. The IR signature from a space object varies as a function of the projected area observed by the sensor. Thus, for an object at absolute temperature \( T \), the power received by the sensor in the wave band \( \lambda_1 \sim \lambda_2 \) is

\[
P_T(\lambda_1 \sim \lambda_2) = \frac{\pi D^2}{4} \frac{1}{R^2} \varepsilon \tau A_{\text{proj}} \int_{\lambda_1}^{\lambda_2} M_T(\lambda) d\lambda, \tag{1}\]

where \( D \) is the optical aperture diameter, \( R \) is the distance between the sensor and the object, \( \varepsilon \) is the emissivity of the object, and \( M_T(\lambda) \) is the spectral radiant exitance of blackbody. According to Planck's law, \( M_T(\lambda) \) can be represented as

\[
M_T(\lambda) = \frac{2 \hbar c^2}{\lambda^5} \left[ \exp \left( \frac{hc}{k \lambda T} \right) - 1 \right]^{-1}, \tag{2}\]

where \( c \) denotes the velocity of light in vacuum, \( k \) denotes the Boltzmann entropy constant, and \( b \) denotes the Planck constant. \( A_{\text{proj}} \) is the projection area along the line of sight (LOS) of the detector.

The geometry of the sensor and a coning object is illustrated in Fig. 1. Coning is a rigid body rotation about an axis that intersects with an object body coordinate, which is usually combined with spinning. The sensor coordinate system is \((U, V, W)\), the object body coordinate is \((x, y, z)\), and the reference coordinate \((X, Y, Z)\) is parallel to the sensor coordinate, and its origin locates in the object’s mass center. \( \alpha, \beta \) are the azimuth and elevation angles of the sensor’s LOS in the reference coordinate. We use this model to simulate observations by sampling randomly from a distribution of physical attributes and dynamic states.

Generally, the shape model of exo-atmospheric objects detected by the IR detector include irregular fragments and symmetries, such as flat-base cone, ball-base cone, cone-cylinder, etc. In this work, our classification considers mainly four typical categories: flat-base cone, ball-base cone, cone-cylinder, and arc-shaped debris. The simulation parameters setting is displayed in Table 1. The detection distance is computed based on the ellipse trajectory theory. We set the start position of geographical coordinates of objects as \((130°E, 50°N, 150 \text{ km})\) and the end position as \((80°E, 40°N, 150 \text{ km})\). The peak distance of the trajectory from the ground is 469.3 km. The detector observes the objects with speed of 6 km/s from the start position (95°E, 35°N, 300 km). The data are gathered from 300 s to 330 s. All objects are assumed to have the same micro-motion parameters, initial temperature, and emissivity.

3. EXO-ATMOSPHERIC IR OBJECTS CLASSIFICATION FRAMEWORK

Given a time sequence of the IR signature, our goal is to learn a set of discriminative features. In fact, our IR signature classification problem can be viewed as a special time series classification problem. There are complicated factors inducing the fluctuation of time sequence of the IR signature in both short range and long range. Moreover, time sequence of the IR signature is often distorted by detector noise. It remains a
Table 1. Simulation Parameters of Four Classes and IR Detector

<table>
<thead>
<tr>
<th>3D models</th>
<th>( r = 0.3 \pm 0.05 \text{ m} )</th>
<th>( r = 0.3 \pm 0.05 \text{ m} )</th>
<th>( r = 0.3 \pm 0.05 \text{ m} )</th>
<th>( r = 0.3 \pm 0.05 \text{ m} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b = 1.0 \pm 0.25 \text{ m} )</td>
<td>( b = 1.0 \pm 0.25 \text{ m} )</td>
<td>( b_1 = 0.4 \pm 0.15 \text{ m} )</td>
<td>( b_2 = 0.6 \pm 0.10 \text{ m} )</td>
<td></td>
</tr>
<tr>
<td>( b = 0.5 \pm 0.20 \text{ m} )</td>
<td>( q = 0.6 \pm 0.1\pi )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Micro-motion parameters

| \( a_i = 0.0\pi \) | \( a_i = 0.0\pi \) | \( a_i = 0.0\pi \) | \( a_i = 0.0\pi \) |
| \( \beta_i = 0.2\pi \) | \( \beta_i = 0.2\pi \) | \( \beta_i = 0.2\pi \) | \( \beta_i = 0.2\pi \) |
| \( \theta = 0.25\pi \pm 0.25\pi \) | \( \theta = 0.25\pi \pm 0.25\pi \) | \( \theta = 0.25\pi \pm 0.25\pi \) | \( \theta = 0.25\pi \pm 0.25\pi \) |
| \( \omega_i = 0.3\pi \pm 0.05\pi \) | \( \omega_i = 0.3\pi \pm 0.05\pi \) | \( \omega_i = 0.3\pi \pm 0.05\pi \) | \( \omega_i = 0.3\pi \pm 0.05\pi \) |
| \( \varphi_0 = 1.0\pi \pm 1.0\pi \) | \( \varphi_0 = 1.0\pi \pm 1.0\pi \) | \( \varphi_0 = 1.0\pi \pm 1.0\pi \) | \( \varphi_0 = 1.0\pi \pm 1.0\pi \) |

Initial temperature

| 320 K | 320 K | 320 K | 320 K |

Emissivity

| 0.5 | 0.5 | 0.5 | 0.5 |

IR detector parameters

| Wave band: 8 ~ 12 μm; observation time: \( T = 30 \text{ s} \); sample frequency: \( f = 20 \text{ Hz} \); |

A. Encoding IR Signatures as Texture Images

The encoding of the input IR signature as a texture image is inspired by Marwan et al. [28]. The IR imaging system for detecting exo-atmospheric motion objects is a periodical dynamical system. As the RP is a suitable tool for analyzing recurrences of states of a dynamical system and measuring recurrences of a trajectory \( \tilde{x}_i \in R^m \) in phase space, \( m \) is the dimension of the phase space trajectory. We adopt RP for the encoding of input. It can be formulated as

\[
R_{ij} = \theta(e - \| \tilde{x}_i - \tilde{x}_j \|), \quad i, j = 1, \ldots, N, \quad (3)
\]

where \( N \) is the number of measured points \( \tilde{x}_i \), \( e \) is a threshold distance, \( \theta(\cdot) \) is the Heaviside function, and \( \| \cdot \| \) is a norm. There is also another key parameter involved in generating the RP matrix: embedding time delay \( \tau \). The \( R \)-matrix contains textures that are single dots and diagonal lines, as well as vertical and horizontal lines, and typology information is characterized as homogeneous, periodic, drift, and disrupted. For instance, fading to the upper left and lower right corners means the process contains a trend or drift; or vertical and horizontal lines/clusters show that some states do not change or change slowly for some time, and this can be interpreted as laminar states [29].

In this paper, we understand \( R \)-matrix as a function of the spatial variation of IR irradiation intensity. Figure 3 shows the \( R \)-matrix calculation for four exo-atmospheric objects' IR signatures (using the phase space dimension \( m = 3 \) and embedding time delay \( \tau = 10 \)). The first row is the IR radiation signature plot. The second row is the \( 3D \) phase space trajectory constructed from the 1D IR signature. The third row is the scaled colors image visualization of \( R \)-matrix, which is calculated based on the closeness of the states in phase space. The last row is the binary plot of \( R \)-matrix with threshold parameter \( \theta = 0.15 \).

The size of \( R \)-matrix is always changed with input data length and RP parameters, e.g., phase space dimension and embedding time delay. However, CNN requires a fixed input size. Thus, down-sampling is applied to generate a fixed input size. In other words, it enables RP-CNN to process input data of any length. It also can achieve a balance between performance and computation of RP-CNN via image size selection.

B. CNN for IR Objects Image Classification

In order to achieve high classification performance, an appropriate architecture and learning algorithm should be chosen. In this work, we propose a two-stage CNN model that contains two convolutional layers, two max pooling layers, and one fully connected layer. The architecture and parameters of the network are selected via extensive experiments. Details of layouts of each network are described in Table 2.

The convolution operation between an input feature map \( x \) and a convolutional kernel \( W \) is defined by

\[
b(x) = f(x \ast W + b), \quad (4)
\]

where \( \ast \) denotes the convolution operator, and \( f \) is the activation function for each layer that adds nonlinearity to the

![Fig. 2. RP-CNN framework.](image-url)
feature vector. We use ReLu for layer activation, which is defined by

\[ f(x) = \max(0, x) \tag{5} \]

Following the convolutional layer, a max-pooling layer is applied to reduce feature maps’ size as well as the number of following layers’ parameters to reduce redundancy and improve computation efficiency. After convolution and max pooling, a fully connected layer is introduced for the later feature extraction block. ReLu is used for layer activation, and dropout is applied to avoid overfitting. A SoftMax function is used to restrict outputs in the ranges (0,1), which is defined by

\[ y_j(h(x)) = \frac{\exp(h_j)}{\sum_i \exp(h_i)} \tag{6} \]

Cross entropy is used as loss function and can be expressed as

\[ L(\theta) = -\frac{1}{n} \sum_{i=1}^{n} \log P(y^i|x^i; \theta) \tag{7} \]

where vectors \( \theta \) are the parameters of the network, \( \{(x^i, y^i), i = 1, 2, \ldots, n\} \) are the set of labeled training set, and \( y^i \) corresponds to the true label of the sample \( x^i \). Then the RP-CNN is trained by adaptive moment estimation (Adam) \[30\] by minimizing the cross-entropy loss between outputs and labeled data. The parameters updating rule is

\[ \theta = \theta - \eta \nabla L(\theta) \]

Table 2. Architectures of the CNN

<table>
<thead>
<tr>
<th>Type</th>
<th>Patch Size</th>
<th>Feature Maps</th>
<th>Strides</th>
<th>Activation</th>
<th>Learning Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conv-1</td>
<td>5 x 5</td>
<td>32</td>
<td>1</td>
<td>ReLu</td>
<td>Adam</td>
</tr>
<tr>
<td>Maxpool-1</td>
<td>2 x 2</td>
<td>32</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conv-2</td>
<td>5 x 5</td>
<td>32</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maxpool-2</td>
<td>2 x 2</td>
<td>32</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fully connected</td>
<td>200</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SoftMax</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. Application of RP on four classes of IR objects.
\[\theta_t = \theta_{t-1} - \frac{\alpha}{\sqrt{\hat{v}_t} + \varepsilon} \hat{m}_t,\]  
(8)

\[\hat{m}_t = m_t / (1 - \beta_1^t),\]

\[\hat{v}_t = v_t / (1 - \beta_2^t),\]  
(9)

where \(\alpha\) is the learning rate, \(t\) is the iteration step, and \(\hat{m}_t, \hat{v}_t\) are the first moment estimation and second moment estimation, respectively. Parameters \(\beta_1, \beta_2\) are exponential decay rate, and the default value of \(\varepsilon\) is \(10^{-8}\) generally. Adam is an algorithm for first-order gradient-based optimization of stochastic objective functions, which are well suited for problems with noisy or sparse gradients. Empirical results show that the Adam algorithm performs well in practice and has great advantages compared to other stochastic optimization algorithms.

4. EXPERIMENTS

In this section, we carry out experiments to compare the performance of multiple approaches and explore the influence of some factors on accuracy. The classification framework is validated using simulated data generated from physics-based models as mentioned above.

A. Experiments Description

To evaluate the proposed method, we need data with different lengths and noise levels. First, we truncate the simulated 30 s objects IR signature data into three parts: the first 10 s, the first 20 s, and 30 s, respectively. We make this division according to the micro-motion periodicities of simulated signatures, which are set in the range of 6 s to 8 s, as displayed in Table 1. Then, we add white Gaussian noise to the data with different signal-to-noise ratios (SNR = 5 dB, 10 dB, 15 dB, 20 dB, 25 dB, 30 dB). The SNR is defined as the ratio of signal power \(P_s\) to the noise power \(P_n\) and is computed by the formula \(\text{SNR} = 10 \log_{10}(P_s/P_n)\). 3200 training data are randomly generated, including 800 testing sets and 600 validating sets. Figure 4 shows the labeled training data where the four class objects are shown using different colors. We can see that they have similar IR irradiance intensity and waveforms, especially class a, class b, and class c. A clear display is given in Fig. 5, which is a set of shuffled training data, showing that the IR signature has periodical fluctuation characteristics. It should be noted that we use only the first 20 s data for displaying. Our method is implemented in Python with the TensorFlow wrapper and run on a PC with 2.7 GHz CPU and 8 GB 1600 MHz DDR3 memory.

B. Comparison with Traditional Methods

We evaluate the performance of the proposed framework with two classical baseline methods: long short-term memory (LSTM) and standard CNN. LSTM is a canonical recurrent network that has superior performance in sequence modeling tasks [31–33]. We conduct two groups of experiments. The parameter settings of RP are determined according to the experiments conducted in the next subsection, and the CNN parameters are chosen by lots of training, displayed in Table 2. We use the Adam optimizer with learning rate 0.001. Cross validation is applied to select the optimal model parameters. The classification results are the average of 20 runs under a range of noise level when data lengths are 10 s and 20 s, respectively, as shown in Figs. 6 and 7.
The results illustrate that classification accuracies of the three methods increase with the increment of SNR. This is because the higher noise level can distort the signature, which directly causes the decrease in classification accuracy. From comparisons of the three methods, we can see that RP-CNN outperforms the other two networks in both conditions of different noise levels and in the scenes of different input lengths. In addition, the proposed framework still shows prominent performance even at a lower noise level and with limited data length, while the LSTM classifier does not perform well in our data. Comparisons of CNN and RP-CNN demonstrate the effectiveness of RP transformation of our proposed framework.

For our model, we encode the 1D IR signatures as 2D texture images using RP, which explores richer relations between features not only in time domain but also in space.

To further evaluate the classification performance of our method, the receiver operating characteristic (ROC) curves and the areas under the ROC curves (AUC) of the above three methods are given in Fig. 8. ROC is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied and is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. We randomly selected one set of data for displaying. It is noticed that the ROC curve of our method is above the other two methods, and our method has the highest AUC.

C. Experiments about Influencing Factors

In order to obtain a better classification model, we conduct experiments to investigate the influence of factors that are important to the performance of RP-CNN. The factors include data length of the IR signature, RP parameters, e.g., embedding dimension and time delay, and down-sampling image size.

1. Influence of Data Length

As the IR imaging system for detecting exo-atmospheric motion objects is a dynamical system, IR signatures are dynamically collected. Thus, the data length of the IR signature is variable, which requires a framework flexible enough to deal with input with different lengths. The applications of RP and down-sampling enable the proposed framework to process data with any length. This subsection aims to validate the effect of data length on RP-CNN classification performance. We test the RP-CNN classifier with different lengths (10 s, 20 s, and 30 s). The classification results are shown in Fig. 9. We can see that the framework achieves a higher performance when data length is larger than 20 s. The superiority of RP is in capturing recurrence patterns of time series. It is less than two cycles when data length is 10 s (the micro-motion periodicities are set in the range of 6 s to 8 s). The RP-CNN has superior performance for processing data that are longer than two cycles.

2. Influence of RP Parameters

The embedding dimension $m$, embedding time delay $\tau$, and binarization threshold parameter $th$ are key factors in the application of RP. Figure 9 shows the performance comparisons of RP-CNN using different RP parameters. IR signature transformation using different RP parameters, e.g., $(m:3, \tau:10)$, $(m:1, \tau:5)$, $(m:3, \tau:5)$, $(m:4, \tau:5)$, $(m:1, \tau:5, th:0.5)$, and $(m:1, \tau:5, th:0.2)$, is shown at the top in Fig. 10 from left to right, respectively. At the bottom, we can see that the
performance of RP-CNN for a binary image is lower than the gray-level image. This is mainly because the binarization of the RP images loses some information. As for gray-level images, the relatively high performance can be obtained when \( m/0.0136, \tau/0.0136 \).

3. Influence of Down-Sampling Image Size
For training of the CNN, a fixed-size window as input layer is required. There are many factors affecting the size of the RP image, e.g., data length of the IR signature and the RP parameters. In this work, we adopt down-sampling to acquire a fixed image size as the input of CNN. In addition, by down-sampling the input images, it can improve the computation efficiency of CNN to a certain degree. Figure 11 shows the classification performance comparisons of RP-CNN using different down-sampling image sizes \((28 \times 28, 20 \times 20, \text{and } 10 \times 10 \text{from left to right, respectively})\). Down-sampling images are shown at the top in Fig. 10. At the bottom, we can see that the performance of RP-CNN using an image size of \(10 \times 10\) is lower than the other two, and the performance of RP-CNN using an image size of \(20 \times 20\) is slightly worse than using an image size of \(28 \times 28\). Too small an image size has limited discriminative ability, while a large size is likely to increase computation cost. For the consideration of appropriate performance and computation efficiency, the image size of \(20 \times 20\) is chosen in this application.

Nevertheless, there are some limitations to this work. First, the training of deep neural networks is time consuming, since model parameters are determined by extensive experiments and the model still can be further optimized; second, the proposed method lacks actual flight data for testing. But we believe that once trained, this framework can be applied to real-data classification examples. Hence, in future work, we plan to study and extend our framework for IR objects classification on more data sets and parameter settings.

5. CONCLUSION
In this paper, we proposed a RP-CNN framework for exoatmosphere objects classification and evaluated its performance with extensive experiments. The results demonstrated the effectiveness of our method, especially in conditions with strong noise and long observing distance. The RP-CNN transforms the 1D IR signatures into 2D texture images using RP as the input of CNN, which explores richer relations between features not only in time domain but also in spatial domain. It overcomes the shortcoming of previous works, in that they learn features only in single time scale. In addition, the applications of RP and down-sampling enable framework flexibility to process IR signatures at any length. The proposed framework shows promise as a tool for objects classification in the exoatmosphere. For future work, we will further optimize our framework using more data sets.
REFERENCES