ARTICLE IN PRESS

Pattern Recognition ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Multimodal deep learning for solar radio burst classification

Lin Ma [a,b], Zhuo Chen [b], Long Xu [b,c,*], Yihua Yan [b]

[a] Huawei Noah's Ark Lab, Hong Kong, China
[b] Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China
[c] School of Computer and Software, Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science and Technology, Nanjing, 210044

ARTICLE INFO

Article history:
Received 31 January 2016
Received in revised form
20 April 2016
Accepted 22 April 2016

Keywords:
Multimodal learning
Solar radio spectrum
Classification

ABSTRACT

In this paper, multimodal deep learning for solar radio burst classification is proposed. We make the first attempt to build multimodal learning network to learn the joint representation of the solar radio spectrums captured from different frequency channels, which are treated as different modalities. In order to learn the representation of each modality and the correlation and interaction between different modalities, autoencoder together with the structured regularization is used to enforce and learn the modality-specific sparsity and density of each modality, respectively. Fully connected layers are further employed to exploit the relationships between different modalities for the joint representation generation of the solar radio spectrums. Based on the learned joint representation, solar radio burst classification is performed. With the validation on the constructed solar radio spectrum database, experimental results have demonstrated that the proposed multimodal learning network can effectively learn the representation of the solar radio spectrum, and improve the classification accuracy.

1. Introduction

Solar radio astronomy is an emerging interdisciplinary field of radio astronomy and solar physics. The discovery of radio waves from the Sun provides a new window to exploit and investigate the solar atmosphere, as new information about the Sun can be obtained. With proper devices, the properties of the solar corona are much more easily depicted with the captured signals at radio wavelengths. As solar radio telescopes have improved a lot in recent years, fine structures in solar radio bursts can thus be easily and accurately detected. In this study, in order to analyze the solar burst behavior, we use the data obtained by solar broadband radio spectrometer (SBRS) of China [1] which is a solar dedicated radio spectrometer for capturing solar radio strength along time over multiple frequency channels in the microwave region. Its functionality is to monitor the solar radio bursts in the frequency range of 0.7–7.6 GHz with time resolution of 1–10 ms. It consists of five "component spectrometers", which work in five different wave bands (specifically 0.7–1.5, 1.0–2.0, 2.6–3.8, 4.5–7.5, and 5.2–7.6 GHz wave bands). As SBRS monitors the solar radio bursts in daytime, it produces massive data about the solar radio information. However, the solar activity researchers are only interested in the data reflecting the burst activity of the Sun in the massive data. However, the data reflecting the Sun burst activity is very rare (1% of the captured data). Moreover, the data is always accompanied with the interference during the capturing process. As such, it is of heavy labor for human to identify whether the data contains burst information or not timely. To the end, analyzing the captured data automatically (burst or not) are highly demanded and beneficial to the solar radio astronomy study.

Nowadays, with the available massive data, especially visual data including images and videos, many algorithms have been developed to learn the representation with unsupervised and supervised methods for the tasks of visual, classification [14,18], localization [24], and so on. Recent progresses on deep learning [2] have demonstrated state-of-the-art performances in a wide variety of tasks, including visual recognition [3,4], audio recognition [5,6], natural language processing [7], cross modality relationship [17,15,19], and so on. These techniques are super powerful because they are capable of learning useful features directly from both unlabeled and labeled data to avoid the need of hand-engineering. For solar radio spectrums, we also have massive data. Firstly, a large amount of the captured solar radio spectrums are unlabeled, most of which do not contain the burst information of Sun. Secondly, the professional experts who have the knowledge of solar physics are employed to label some of our captured data, which is

* Corresponding author.
E-mail addresses: forest.linma@gmail.com (L. Ma), chenzhuo@nao.cas.cn (Z. Chen), lxu@nao.cas.cn (L. Xu), yyh@nao.cas.cn (Y. Yan).

http://dx.doi.org/10.1016/j.patcog.2016.04.013
0031-3203/

time intensive and labor intensive. In this paper, we apply the deep learning, specifically the multimodal deep learning on the massive data of solar radio spectrums. The powerful ability of the deep network is expected to learn the inherent structural information of the solar radio spectrums for an effective and automatic classification of solar radio spectrums.

There are several kinds of disciplines for realizing unsupervised learning of deep neural networks for the massive data, such as Boltzmann machine, autoencoder (AE). AE is an unsupervised learning algorithm that applies back-propagation by setting the target values to be equal to the inputs. AE tries to learn a function which makes the input similar to the output of the function. In other words, it is trying to learn an approximation to the identity function, so as to output of the network that is similar to the input. The identity function seems a particularly trivial function to be trying to learn. But by placing constraints on the network, such as by limiting the number of hidden units, interesting structure about the data can be learnt. Therefore, AE is very helpful for representation learning of data, including visual data. The variances of AE, such as denoising AE [8] and stacked AE (SAE) [9] were also developed widely. In [10], the authors proposed an automatic dimensionality reduction to facilitate the classification, visualization, communication, and storage of high-dimensional data through an adaptive, multilayer encoder network to transform the high-dimensional data into a low-dimensional code and a similar decoder network to recover the data from the code. Using random weights as the initialization in the two networks, they can be trained together by minimizing the discrepancy between the original data and its reconstruction. Then the representation can be learned in an unsupervised manner. The network is also named as deep belief network (DBN). With the achievements of these learning methods, we can learn the representation of the solar radio spectrum even better than [11], which will be employed for further solar radio spectrum analysis, such as clustering, classification, and so on. However, Both AE and DBN, as well as their variants, treat the input signals equally. If the network takes different signals as the input, the characteristics between different input modalities cannot be distinguished. Thus the interaction and contributions between different modality inputs cannot be well exploited and captured.

For solar radio spectrums, the signals are captured from different frequency channels, which depict the Sun's activity from different perspectives. In this paper, we firstly employ the multimodal learning method, specifically the AE with the structured regularization, to learn the representation of the solar radio spectrum by distinguishing the contribution of each modality. By further stacking more fully connected (FC) layers, the joint representation of the solar radio spectrums is generated, which is input to the softmax layer for classification. By evaluating the constructed multimodal network on the solar radio spectrum database, the experimental results demonstrate that the multimodal learning method can effectively analyze the solar radio spectrum.

The rest of the paper is organized as following. In Section 2, a multimodal learning architecture is introduced. In Section 3, a deep neural network based on the multimodal learning architecture is proposed to classify the solar radio spectrum. Section 4 gives the experimental results on representation learning and classification. And the final section concludes the paper.

2. Multimodal learning architecture

We propose a multimodal learning architecture for the purpose of solar radio burst classification, which is illustrated in Fig. 1. The proposed multimodal learning architecture takes different numbers and types of modalities as the input and generates their joint representation for the targeted task, such as classification. The proposed multimodal learning model needs to adequately learn the representation of each individual modality. Most importantly, the inter-modality relationships and interactions need to be accurately captured to generate the joint representation. As illustrated in Fig. 1, our proposed multimodal learning model relies on AE with the structured regularization to model each modality individually and jointly capture their interactions. Afterwards, several FC layers are stacked and employed to nonlinearly transform the intermediate representation to the final joint representation for the specific target task. The benefits of the introduced FC layers are twofold. Firstly, the multiple FC layers with the nonlinear activation function will increase the nonlinearity of our proposed multimodal learning architecture, which will further make the final decision function (such as classification) more discriminative [18]. Secondly, the multiple modalities can interact more closely with each other through layers of FC nonlinear transformation. As such, the proposed multimodal learning model learns the multimodal abstractive representations from the detailed information contained in each modality.

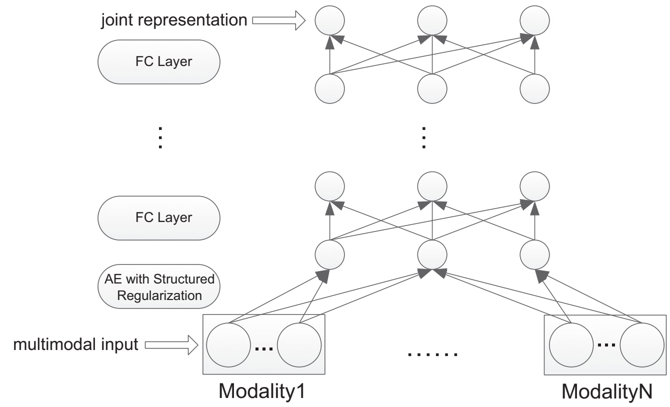

Fig. 1. The framework of the multimodal learning architecture.

We formulate the proposed multimodal learning architecture as:

$$\nu_{jr} = f_t^n\left(\cdots\left(f_t^1\left(f_{SR}(x_1, x_2, \ldots, x_m)\right)\right)\right) \quad (1)$$

where ν_{jr} is the final learned joint representation from the input with m different modalities $x_1, x_2, \ldots, x_m$. f_{SR} takes the input different multiple modalities and learns their intermediate representations. $f_t^1, \ldots, f_t^n$ are the following n FC layers, which are stacked together and transform the intermediate representation learned from f_{SR} to the final joint representation ν_{SR}. As illustrated in Fig. 1, f_{SR} is realized by the AE together with the structured regularization in this paper. AE aims at transforming the input signal into output signal with the smallest distortions. AE treats each node of the input signal equally by performing the mapping process from the input to the output. As such, the different contributions of different modalities to the nodes of the output signal cannot be well learned and captured. However, different modalities may contribute differently to the specific task. In order to overcome this limitation and fully exploit the contributions and interactions between different modalities, the structured regularization is introduced to AE, which makes the proposed multimodal learning network distinguish different modalities with individual treatments for the intermediate representation. Moreover, our proposed multimodal learning network can be trained greedily layer by layer, as such stacking architecture ensures the scalability of the learning ability. On one hand, as aforementioned more nonlinear transformation layers can help improve the nonlinearity

Please cite this article as: L. Ma, et al., Multimodal deep learning for solar radio burst classification, Pattern Recognition (2016), http://dx.doi.org/10.1016/j.patcog.2016.04.013