SOLAR RADIO SPECTRUM CLASSIFICATION WITH LSTM

Xuexin Yu[1], Long Xu[1], Lin Ma[2], Zhuo Chen[3], Yihua Yan[1]

[1]Key Laboratory of Solar Activity, National Astronomical Observatories,
Chinese Academy of Sciences, Beijing, China
[2]Tencent AI Lab, Shenzhen, China
[3]School of Computer Science and Engineering, Nanyang Technological University, Singapore
yuxuexin2008@gmail.com, lxu@nao.cas.cn, forest.linma@gmail.com,
chenzhuo.zoom@gmail.com, yyh@nao.cas.cn

ABSTRACT

This paper makes the first attempt to utilize long short-term memory (LSTM) for classification of solar radio spectrum. A solar radio spectrum is a gray-scale image representing solar radio radiation over multiple frequency channels and in a short time period. The vertical and horizontal dimensions of a spectrum correspond to frequency channel and time, respectively. Intrinsically, time dependence exists between columns of a spectrum, which indicates the slowly varying process of solar radio radiation. Thus, a spectrum can be treated as a time sequence instead of a general image losing its time information. Inspired by the big success of LSTM for time sequence processing, e.g., natural language recognition, it is reasonably believed that treating a spectrum as a time sequence would benefit its classification via LSTM. Thus, LSTM is employed to explore the sequential relations and interactions within a spectrum being treated as a time sequence. As such, LSTM learns the sequential properties and generates the representation of solar radio spectrum for classification. The experimental results demonstrate that LSTM can well capture the characteristics of solar radio spectrum, and thus achieves better classification accuracy.

Index Terms—solar radio spectrum; long short-term memory (LSTM); classification

1. INTRODUCTION

Solar radio astronomy belongs to an emerging interdisciplinary branch of radio astronomy and solar physics. The discovery of radio waves from the Sun provides a new method which obtains new information about the Sun. Recently, with the development of solar radio telescopes, fine structures in solar radio bursts can be detected. In this paper, we study the classification of solar radio spectrums obtained by Solar Broadband Radio Spectrometer (SBRS) of China [1]. The SBRS is with characteristics of high time resolution, high-frequency resolution, high sensitivity, and wide frequency coverage in the microwave region. The SBRS, which consists of five "component spectrometers" and works in five different wave bands (0.7-1.5, 1.0-2.0, 2.6-3.8, 4.5-7.5, and 5.2-7.6 GHz), can monitor solar radio bursts in the 0.7–7.6 GHz with time resolution of 1–10ms. Monitoring the solar radio bursts in every day, the SBRS produces massive data. It is exhausting for researchers to analyze and process massive data by manual operation every day. Besides, strong interference of noise in the observed data further cause a trouble to identify whether or not a spectrum contains bursts and thereby distinguish which type of burst it is. Therefore, automatically classifying the observed spectrums has great significance for solar radio astronomy study.

Nowadays, with the availability of massive data, deep learning [2] has been extensively explored to solve many traditional tasks of recognition, classification, regression and clustering. The methods of deep learning, such as convolutional neural networks (CNN), recurrent neural networks (RNN), deep belief networks (DBN) and etc., have demonstrated state-of-the-art performances in a wide variety of tasks, including visual recognition [3], audio recognition [4], and natural language processing [5]. These methods can directly learn useful features from the unlabeled or labeled data, instead of the need for hand-engineering, which will be helpful for automatic analysis of solar radio spectrum. In our previous works [6, 7, 8], several deep learning networks have been developed for classification of solar radio spectrums. Our efforts revealed that the DBN produces a better performance compared with traditional method of support vector machine (SVM) + principal component analysis (PCA) [6]. The multimodal neural networks (MNN) outperforms DBN [7, 8]. Moreover, it is shown that the classification accuracy of solar radio spectrums relates to the number of hidden layers for the deep MNN [8].

Long short-term memory (LSTM) is a special kind of RNN, which has been widely employed to learn the representation in various tasks, such as speech recognition [9], machine translation [10]. LSTM tries to explore the relations and interactions of sequential data. Thus, LSTM is

978-1-5386-0560-8/17/$31.00 ©2017 IEEE

very suitable for representation learning of data related to time. Also, there are many variations of LSTM, such as LSTM adding peephole connection [11], the gated recurrent unit (GRU) [10], and popular variants of LSTM are compared and tested in some tasks [12,13]. The achievements provide the help for further analysis of solar radio spectrum related to time, such as clustering, classification, and so on.

In this paper, the LSTM is employed to learn the representation of the solar radio spectrum. Based on the representation, the spectrums are further classified into different categories. The main contributions of the paper are as following:

- The first attempt is made to employ LSTM to learn the representation of solar radio spectrums.
- The first attempt is made to seek time dependence within a solar radio spectrum whose structure is organized to make it more suitable for LSTM processing.
- A group of pre-processing methods, including channel normalization, down-sampling and spectrum normalization, are raised for the representation learning and classification task of solar radio spectrum.

The rest of the paper is organized as following. In Section II, the LSTM network is introduced to learn the representation of solar radio spectrums and thereby classify spectrums. Section III introduces the organizing of spectrum structure and the pre-processing procedures of spectrum to explore time dependence with a spectrum. Section IV gives the experimental results and analysis on spectrum classification. And the final section concludes the paper.

2. LSTM FOR SOLAR RADIO SPECTRUM CLASSIFICATION

The LSTM network architecture is illustrated in Fig. 1, where $(x_1, x_2, \cdots, x_T)$ is an input sequence which is consist of all columns of a spectrum. We employ the LSTM layer to learn the representation of the solar radio spectrum for classification. Then, a softmax layer, which is stacked on the top of the LSTM, takes the learned representations as the input, and outputs the classification results for each spectrum. Due to limited number of labeled solar radio spectrums, only one LSTM layer is employed in this work to avoid overfitting

In LSTM cell, there is a memory block to store the status of the cell in order that this status can be transferred to next time step. Thus, the LSTM layer mainly explores interactions and relationship of sequential data. Assumed $f_{LR}(\cdot)$ is the mapping function of LSTM, h_T represents the hidden state of LSTM cell at T time, the process of input to LSTM layer is represented by

$$h_T = f_{LR}(x_1, x_2, \cdots, x_T). \quad (1)$$

With LSTM, we can generate the representations of the input solar radio spectrums. Afterwards, the representation is fed into classifier for the classification. Assumed $\hat{L}$ is the output of the classifier, $\varphi[\cdot]$ is the classification function, the global function of the LSTM network is described by

$$\hat{L} = \varphi\left[f_{LR}(x_1, x_2, \cdots, x_T)\right]. \quad (2)$$

2.1. LSTM layer

As afore mentioned, there are many variations of LSTM. For the version of LSTM used in this paper [14], the basic structure of an LSTM cell is illustrated in Fig. 2.

In LSTM layer, the mapping function from an input sequence $x = (x_1, x_2, \cdots, x_T)$ to an output sequence $h = (h_1, h_2, \cdots, h_T)$ is precisely specified by

$$i_t = \sigma(W_{ix} x_t + W_{ih} h_{t-1} + b_i), \quad (3)$$

$$f_t = \sigma(W_{fx} x_t + W_{fh} h_{t-1} + b_f), \quad (4)$$

$$o_t = \sigma(W_{ox} x_t + W_{oh} h_{t-1} + b_o), \quad (5)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{cx} x_t + W_{ch} h_{t-1} + b_c), \quad (6)$$

$$h_t = o_t \tanh(c_t), \quad (7)$$

where the σ is the logistic sigmoid function, and i, f, o and c are respectively the input gate, forget gate, output gate, cell and cell input activation vectors, all of which are the same size as the cell output activation vector h. The W_* and b_* represent weight matrixes and bias vectors, severally.

2.2. Softmax layer

In this paper, softmax layer is employed for the classification, which is defined as:

$$softmax(x_i) = \frac{exp(x_i)}{\sum_i exp(x_i)}, \quad (8)$$

$$y_t = softmax(W_{sh} h_t + b_s), \quad (9)$$

where W_{sh} and b_s are the weight matrix and the bias vector, respectively. The output of softmax y_t denotes the probability of a spectrum belonging to each of types of spectrum.

3. PRE-PROCESSING OF SOLAR RADIO SPECTRUMS

SBRS records the Sun's activity over 120 frequency channels and with a certain time resolution. Taking 8ms spectrum file for example, it simultaneously records the Sun over 120 frequency channels in each 8ms to form a spectrum file. The stored file of 8ms spectrum contains right and left circular polarization parts, each of which is a matrix of 120×2520. After visualization, the spectrum of SBRS can be represented

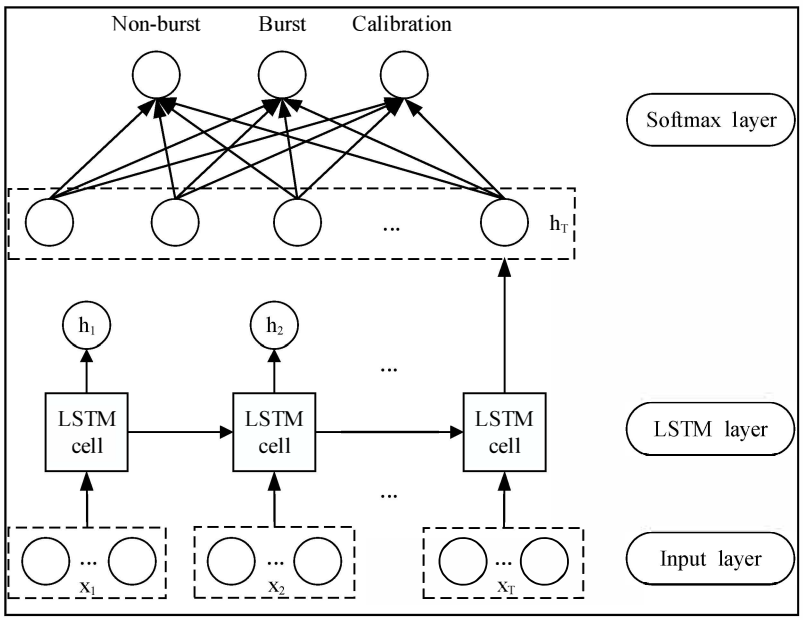

Fig. 1. LSTM network for solar radio spectrum classification

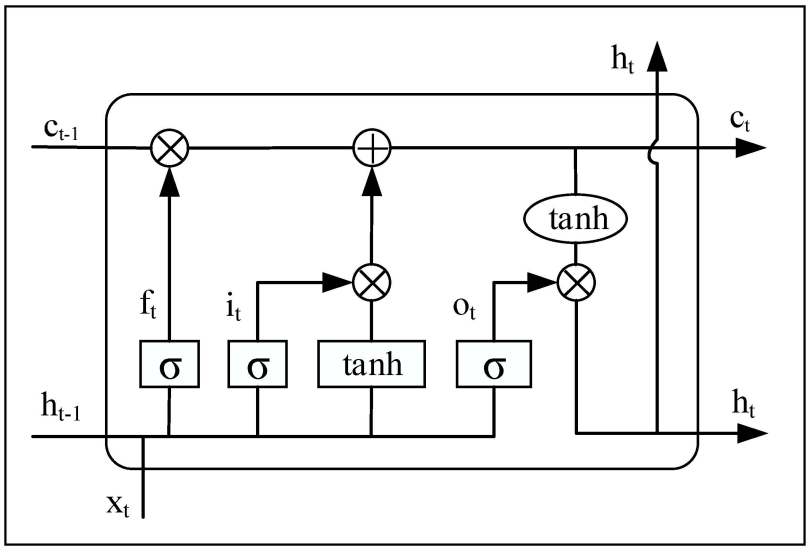

Fig. 2. The basic structure of an LSTM cell

by a gray-scale image as illustrated in Fig. 3. The intensity of each pixel represents the amount of solar radio radiation at a certain frequency channel and at a certain time point. The whole image illustrates solar radio radiation over multiple frequency channels in a short time period.

In our previous works [6, 7, 8], some pre-processing methods, including channel denoising and normalization, and down-sampling, were investigated. These pre-processing methods were designed for better visualization of solar radio spectrum. However, in order to explore time dependence within a spectrum and be adapted to LSTM, different pre-processing procedures are developed as follows.

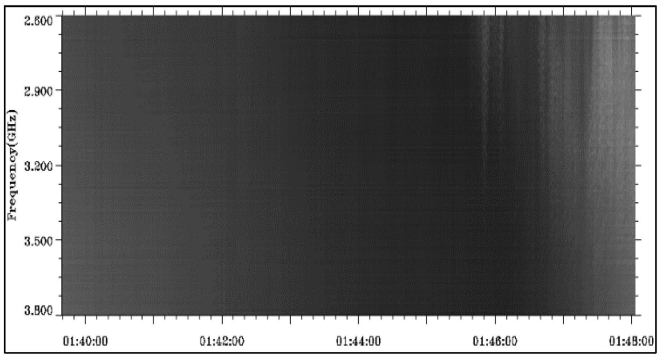

Fig. 3. The solar radio spectrum

3.1. Time sequence of spectrum feeding to LSTM

Referring to Fig. 3, each column of a spectrum is regarded as a sequential data. As such, a spectrum is transformed into a time sequence of columns oriented. Thus, LSTM can explore the interactions between columns at different time points. There are two reasons for arranging spectrum into time sequence in columns oriented instead of rows oriented. First, each column is recorded at the same time point, so the columns form a time sequence intrinsically. Second, the columns oriented sequence of a spectrum can better represent the variation of a solar radio burst. Within a spectrum, the burst can be identified by the variation of each row independently. Thus, a time sequence of columns oriented can represent a process of spectrum variation. However, the rows oriented sequence of a spectrum may not be able to represent a process of spectrum variation since all rows are synchronous during bursts and independent of each other.

3.2. Channel normalization

To make more distinct variation of solar radio radiation at each frequency channel of a spectrum, we propose one method of channel normalization, which is formulated as following:

$$\widehat{D} = D - D_{CM}\,, \tag{10}$$

where D is the matrix of a solar radio spectrum, $\widehat{D}$ is the matrix after performing the channel normalization, and D_{CM} is the minimum value of each channel within a spectrum.

3.3. Down-sampling and spectrum normalization

Moreover, in order to reduce the burden of network training and maintain characters of raw data, the original spectrum matrix with the size of 120×2520 is reduced to 120×120 with 1×21 weighted mean filter in each channel. Simultaneously, we make spectrum normalization, which is formulated as following:

$$\widehat{d} = \frac{d - min\,(D)}{max\,(D) - min\,(D)}\,, \tag{11}$$

where D represents a matrix of a solar radio spectrum, d denotes a data element in the spectrum matrix, and $\widehat{d}$ is a data element after performing the spectrum normalization.

Fig.4 illustrates a channel of a spectrum before and after pre-processing. It can be observed that the dynamic range of red curve is significantly larger than that of blue curve. This makes features of the solar radio burst more distinguishable, and is conducive to learn them by LSTM network.

4. EXPERIMENTAL RESULTS AND ANALYSIS

To evaluate the LSTM network for spectrum classification, we implement it on Tensorflow library by using python. The solar radio spectrum database established by us in [7, 8] is used.

The database includes 8816 spectrums and corresponding labels given by experts among three candidate types: “burst”, “non-burst” and “calibration”. As afore mentioned, a solar radio spectrum is a 120×120 matrix after pre-processing. The detailed information about the database is illustrated in Table 1.

The training and testing are performed on the split of the database. 800 “burst”, 800 “non-burst” and 800 “calibration” are randomly selected from the dataset for training, and the rest are for testing. The details about training set and testing set are listed in Table 2.

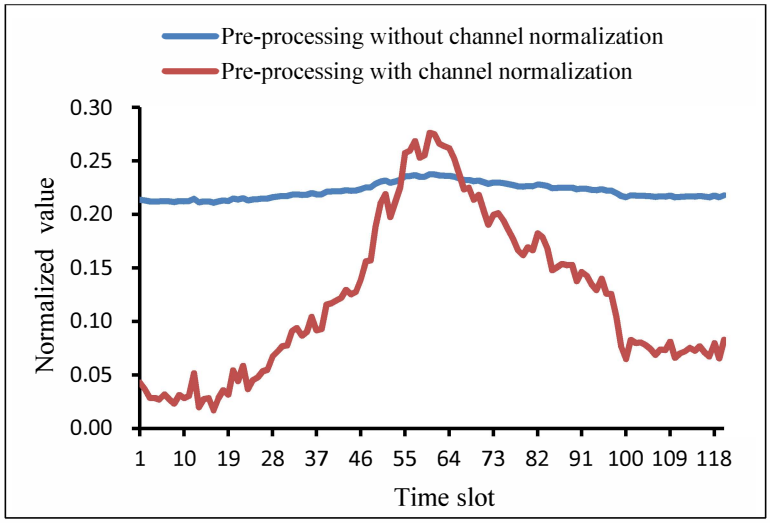

Fig. 4. A channel of a spectrum before and after pre-processing

Table 1. The details of the database

Spectrum Type	Non-burst	Burst	Calibration	Total
Spectrum Number	6670	1158	988	8816
Spectrum Size	120×120	120×120	120×120	120×120

Table 2. The details of training and testing data

Spectrum Type	Non-burst	Burst	Calibration	Total
Training Number	800	800	800	2400
Testing Number	5870	358	188	6416
Spectrum Size	120×120	120×120	120×120	120×120

Table 3. Performance comparisons

Spectrum Type	**LSTM**		Multimodal Deep		Multimodal		DBN		PCA+SVM	
	TPR	**FPR**	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR
Non-burst	**92.3%**	**8.2%**	83.3%	9.6%	80.9%	13.9%	86.4%	14.1%	0.1%	16.6%
Burst	**85.4%**	**6.7%**	82.2%	22.5%	70.9%	15.6%	67.4%	13.2%	52.7%	26.6%
Calibration	**96.2%**	**0.9%**	92.5%	1.7%	96.8%	3.2%	95.7%	0.4%	38.3%	72.2%

In order to efficiently assess the performance of the LSTM, we compare it with previous DBN method, a PCA+SVM model, multimodal learning and multimodal deep learning for classification of solar radio spectrum [6, 7, 8]. The receiver operating characteristic (ROC) [15] analysis is employed in performance analysis. True positive rate (TPR) is defined as the number of correct classification among all positive samples during the test. Thus, the larger the TPR value, the better the classifier. Similarly, false positive rate (FPR) is defined as the number of negative samples which are wrongly classified into positive category during the test, so the smaller the FPR value, the better the classifier. In this case, there are three categories of spectrums. For each category, one can compute TPR and FPR. The statistics of TPR and FPR are listed in Table 3.

From Table 3, TPR of LSTM is larger than those of others for each spectrum type, which means LSTM can identify "burst", "non-burst" and "calibration" more accurately than others. Actually, TPR acts as general accuracy index of classification, FPR represents the error that other types are wrong classified into a certain type. For example, FPR of 6.7% is for LSTM. It means 6.7% of "non-burst" and "calibration" are wrongly categorized into "burst". For Multimodal Deep, although its TPR is acceptable, its FPR is much worse than LSTM. It means that Multimodal Deep could wrongly categorize other types into "burst". In a word, the smaller the FPR is, the better the performance is. From Table 3, it can be observed that LSTM is better than others with respect to FPR on "burst". Thus, larger TPR and smaller FPR are for LSTM, especially for "burst". It means we can accurately figure out "burst" from others. This is most important in our daily task to select "burst" from massive data.

It can be observed that TPR of "burst" is always lower than others in LSTM. The first reason may be that the structure of "burst" is more complicated than the other two. The second reason may be that limited number of labeled "burst" makes network not to learn all the features. In the future, we would build larger database to further improve spectrum classification. For "calibration", it seems like that LSTM, Multimodal Deep, Multimodal and DBN are comparable with respect to TPR. They are all with bigger TPR. The reason is that "calibration" is with very specific characteristics in spectrum form. It can be easily figured out by all models.

The gain of the proposed model may come from three aspects. First, a spectrum is reorganized into a time sequence so that its inner structure can be exploited to benefit classification. Second, LSTM is employed to learn the relations and interactions of sequential data and generates the representation of solar radio spectrum for classification. Third, pre-processing of spectrums enhances variation of solar radio radiation at each frequency channel of a spectrum as illustrated in Fig. 4.

5. CONCLUSIONS

The paper makes the first attempt to employ LSTM network for classification of solar radio spectrums. The solar radio spectrums are pre-processed to enhance variation of solar radio radiation at each frequency channel of a spectrum, and thereby each of spectrum is reasonably arranged into a time sequence to be fed into the LSTM network. The experimental results demonstrated the superiority of the LSTM network on the solar radio spectrum classification task.

6. ACKNOWLEDGMENT

This work was partially supported by a grant from the National Natural Science Foundation of China under Grant 61572461, 11433006 and CAS 100-Talents (Dr. Xu Long).

7. REFERENCES

[1] Q. Fu, et al., "A new Solar Broadband Radio Spectrometer (SBRS) in China," *Solar Physics*, vol. 222, no.1, pp. 167-173, 2014.

[2] Y. Bengio, “Learning Deep Architectures for AI,” 2rd ed., vol. 1, Foundations and Trends® in Machine Learning, 2009, pp. 1-127.

[3] K. Sohn, D. Y. Jung, H. Lee, and A. O. Hero, “Efficient learning of sparse, distributed, convolutional feature representations for object recognition,” in *Int. Conf. on Computer Vision*, vol. 24, pp. 2643-2650, 2011.

[4] A. R. Mohamed, G. E. Dahl, and G. Hinton, “Acoustic modeling using deep belief networks,” *IEEE Transactions on Audio Speech and Language Processing*, vol. 20, no.1, pp. 14-22, 2012.

[5] R. Collobert, J. Wetern, Bottou, L&#, M. Karlen, K. Kavukcuoglu, and P. Kuksa, “Natural language processing (almost) from scratch,” *Journal of Machine Learning Research*, vol. 12, no.1, pp. 2493-2537, 2011.

[6] Z. Chen, L. Ma, L. Xu, C. Tan, and Y.Yan, “Imaging and representation learning of solar radio spectrums for classification,” *Multimedia Tools and Applications*, vol. 75, no.5, pp. 2859-2875, 2016.

[7] Z. Chen, L. Ma, L. Xu, Y. Weng, and Y. Yan, “Multimodal learning for classification of solar radio spectrum,” in *IEEE Int. Conf. on Systems, Man, and Cybernetics*, pp. 1035-1040, 2015.

[8] L. Ma, Z. Chen, L. Xu, and Y. Yan, “Multimodal deep learning for solar radio burst classification,” *Pattern Recognition*, vol. 61, pp. 573-582, 2016.

[9] A. Graves and N. Jaitly, “Towards end-to-end speech recognition with recurrent neural networks,” in *Proc. of the 31st Int. Conf. on Machine Learning*, vol. 32, pp. 1764–1772, 2014.

[10] K. Cho, et al., “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” *Computer Science*, 2014.

[11] F. A. Gers, N. N. Schraudolph, Schmidhuber, J&#, and rgen, “Learning precise timing with lstm recurrent networks,” *Journal of Machine Learning Research*, vol. 3, no. 1, pp. 115-143, 2003.

[12] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, “LSTM: A search space odyssey,” *IEEE Transactions on Neural Networks and Learning Systems*, 2016.

[13] R. Jozefowicz, W. Zaremba, and I. Sutskever, “An empirical exploration of recurrent network architectures,” in *Proc. of the 32 nd Int. Conf. on Machine Learning*, vol. 37, 2015.

[14] F. A. Gers, J. Schmidhuber, and F. Cummins. “Learning to forget: continual prediction with LSTM,” *Neural Computation*, vol. 12, no. 10, pp. 2451-71, 2000.

[15] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874, 2006.