論文 / 著書情報
Article / Book Information

論題	
Title	Astronomical Image Subtraction for Transient Detection Using CNN
著者	LONG YAN, 井上 中順, 篠田 浩一, 谷津 陽一, 伊藤 亮介, 河合 誠之
Author(s)	Yan Long, Nakamasa Inoue, Koichi Shinoda, Yoichi Yatsu, Ryosuke Itoh, Nobuyuki Kawai
出典	第21回 画像の認識・理解シンポジウム論文講演集, , ,
Citation	, , ,
発行日 / Pub. date	2018, 8

Astronomical Image Subtraction for Transient Detection Using CNN

Yan LONG[1,a)] Nakamasa INOUE[1,b)] Koichi SHINODA[1,c)] Yoichi YATSU[1,d)] Ryosuke ITOH[1,e)] Nobuyuki KAWAI[1,f)]

Abstract

As the interest in astronomical transient event has increased in recent years because understanding of them help scientists know the mechanism of how our universe generates and expands, transient detection currently is becoming increasingly important in astronomical survey. HOTPANTS is one of the most widely used method for astronomical image subtraction which works by taking two aligned images of the same filed and in different time, dividing them into several regions and calculating convolutional kernels to match point spread functions (PSF) for each region. However, only one kernel is used in HOTPANTS, which is improper to deal with stars of various sizes. In this paper, we propose a convolutional neural network (CNN) model consisting of multiple layers and kernels of different windows sizes to solve that problem. Our method uses Akeno dataset, which is collected in Akeno, Japan throughout multiple years. The experimental results reveal great performance of our method over HOTPANTS from 20.83% to 6.48% in terms of FNR.

1. Introduction

Transient, abbreviation of astronomical transient event, refers to astronomical object or phenomenon whose duration ranges from seconds to days, weeks or even several years. Difficulty in finding transient exists in its short duration and scarcity. The traditional method of transient detection is based on difference images where most non-transients are removed. A difference image is obtained by subtracting one image called science image from another one called template image which is taken in the same place but earlier than science in time. Direct subtraction is not an ideal solution because even though template and science images are taken by the same device on the same field of sky, they still differ in PSF (response of an imaging system to a point source or point object) caused by change of atmosphere at different time, which means even the same star may look slightly different on different images. Consequently, PSF matching is the most important task in astronomical image subtraction. Many approaches have been proposed to solve this problem in past years, and HOTPANTS[1], the implementation of Alard algorithm[2,3] for image subtraction, is the most popular one among them.

The goal of HOTPANTS is to find a proper convolutional kernel to best match PSFs on science (S) and template (T) images. It is described mathematically as: $(\sum(T(x,y)\bigotimes K(u,v)-S(x,y))^2)_{min}$. K can be decomposed to sum of several weighted basis functions K_n:

$$K(u,v)=\sum_{n=1}^{N}a_n(x,y)K_n \tag{1}$$

Then this problem is transformed into a linear-squares problem and can be solved by substituting stars to equations. (x,y) and (u,v) denote pixel-location in images and kernel respectively. $\bigotimes$ means convolution operation. a_n is coefficient of each basis function to be calculated in HOTPANTS and basis function, K_n, is Gaussian function:

$$K_n(u,v)=e^{-(u^2+v^2)/2\sigma_k^2 u^i v^j} \tag{2}$$

HOTPANTS is an effective method for image subtraction and has been used for a long period of time but it also has some disadvantages. First, we checked the remaining on the image produced by HOTPANTS and found over 50% remaining stars are of small sizes. We think it is

1 Tokyo Institute of Technology
a) yan@ks.cs.titech.ac.jp
b) inoue@ks.cs.titech.ac.jp
c) shinoda@c.titech.ac.jp
d) yatsu@hp.phys.titech.ac.jp
e) itoh@hp.phys.titech.ac.jp
f) nkawai@phys.titech.ac.jp

because one kernel is not able to extract features well of stars of various sizes. Second, it is difficult to optimize and tune the parameters. In HOTPANTS, the number of Gaussian functions, the σ of each Gaussian must be pre-defined by users. Therefore, the parameterization is not general. Users have to adjust parameters unceasingly until getting the best results for different inputs.

In this paper, we propose a CNN model which is composed of multiple layers for transient detection. Kernels in different layers could have windows of different sizes, allowing them to extract features of stars of all-kinds of sizes. We begin with some related works using machine learning methods in Section 2. We explain how we make training data from given dataset in detail in Section 3. We introduce our CNN architecture in Section 4 and then describe our experiments and show its result in Section 5. Finally, we summarize our work and give a conclusion in Section 6.

2. Related work

Historically, transient detection requires huge human effort because transient only accounts for less than 0.1% among all the celestial objects. Accordingly, subtraction between template and science image is done to reduce the number of non-transients, then the scientists check remaining candidates on difference image one by one. Traditional method greatly depends on the effectiveness of difference image. A variety of algorithms have been presented over last decade, but the basic idea of them are to find a convolutional kernel to match PSFs on two input images by least-square solutions for equations. Image subtraction algorithm was first proposed by Alard & Lupton, later Alard improved their work by replacing constant kernel with a space-varying kernel. This is what HOTPANTS bases on. Other methods include Bramich [4], ZOGY [5] and PTFIDE [6].

The term machine learning is created by Arthur Samuel in 1959 and began to flourish in the 1990s thanks to increasing availability of digital information. Astronomers did some attempts to apply machine learning to transient detection, such as [7], [8] and [9]. These works use an engineering feature as input, which is designed by professional astronomers based on their expertise and experience during the process of visual inspection. Usually, designing and editing engineering features require a large quantity of time and effort.

With the rapid development of deep learning, researchers in the area of astronomy begin utilizing deep learning to solve all sorts of astronomical problems. For example, [10] and [11] have used RNN and CNN for supernovae classification. In this paper, we exploits CNN to detect transient, with the input of only two channels from science and template image respectively.

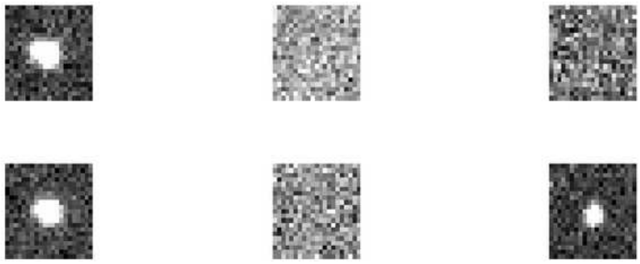

Fig. 1 Examples of training data. The first column is positive, the second column is background and the third column is transient candidate.

3. Data

Here we introduce Akeno Astronomical Image Dataset which is collected in Akeno, Japan throughout multiple years from 2008 to 2016. After removing files which were taken in bad weather, totally we have 183 astronomical images of 4 fields in R, G, I band.

First of all, we align two time-continuous images in the same band. This can be done in a way with the help of SExtractor, which is a toolkit to build a catalogue of all the objects from an astronomical image. We use it to get x, y coordinate of each star on the template and science image. Then star pairs are built by matching RA and DEC from two images and calculate the x, y shift between each star pair. Finally the average vertical and horizontal shift is obtained.

As explained above,the template and science image differ in many aspects. As a consequence, regardless of transient, even an invariable non-transient may look a little different on them. Luckily, the change of a non-transient is more regular than transient which has various types. Hence in this paper what we seek is non-transient for its simplify and regularity instead of transient. Though in our experiment we may meet three situations: (1) non-transient, (2) transient candidate and (3) background, we still insist to regard this problem as a binary classification problem because we don't perform any operation on (2) and (3). Once non-transient is detected, it is removed from input image while (2) and (3) is kept. We label non-transient positive and other cases negative. Figure 1 shows the examples of positive and negative.

In the experiment, we don't input the whole image to

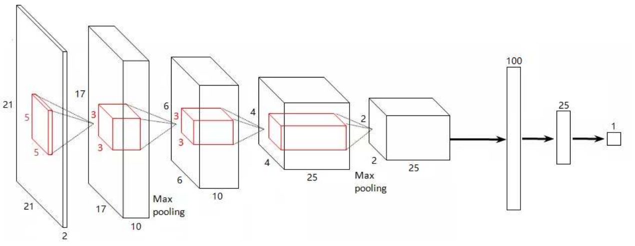

Fig. 2 Architecture of our proposed method.

CNN. Instead, we use a fixed 21 * 21 region pairs extracted from science and template images. We assume that all the stars on input image pair are positives, since most (99%) of stars on images are non-transient as mentioned previously. We are aware that some negatives may be regarded as positives, but it won't have much impact on the final result. Positives are generated by locating stars on science image and crop the regions from both science and template images on the same location. For each star on a science image, we produce a positive sample of 2 * 21 * 21 size tensor.

Negatives consist of background and transient candidates. Nevertheless, transient candidates are extremely rare in real data. To make a balance between negative and positive samples, we have to make some artificial transient by ourselves. Generation of background is similar to positive, the difference is that the background pair don't contain anything or at least nothing in their central region. The Generation of transient candidate is based on background pair: first take a background pair, then add a star from positive to the background from science image. Note that the center of star should overlap the center of the region.

Totally, we have 25,644 pairs of positive and 25,644 pairs of negative (half are transient candidates and another half are background). Notice that all the samples need to subtract the background (the median of the image).

4. Proposed method

Figure 2 shows the architecture of our proposed model, which is composed of 3 convolutional layers, 2 max-pooling layers and 2 fully connected layers. The activation functions used are all hyperbolic tangent function (f(x) = tanh(x)) The network receives a pair of regions of size 21 * 21 and it stack them forming a 2 * 21 * 21 tensor. The data go through the first convolutional layer consisting of 10 filters, each one of them is of size 5 * 5. After first convolutional layer, a max-pooling layer of 3 * 3 filters is used, outputting a tensor of 6 * 6 * 10.

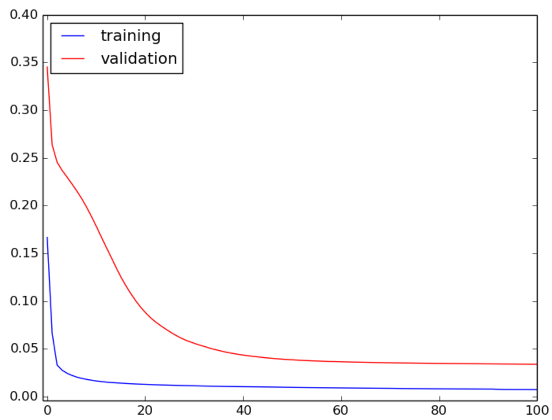

Fig. 3 Learning curve for training and validation sets.

Next, a second group of convolutional layer and max-pooling layer are used. The data from last layer passes through the second convolutional layer consisting of 25 filters of 3 * 3 size and second max-pooling layer, filter size of which is 2. And other setting is the same as previous layers.

Finally, tensor of 2 * 25 * 25 size from second pooling layer is reshaped to a vector of size 100 and input to last two fully connected layers of 100 units and 25 units respectively. The output is the probability representing the probability of non-transient.

5. Experiment

We used data described in Section 3 to train and test our model. As presented in Section 3, dataset includes 51,288 samples totally, half of which are positives and half are negatives. We split the dataset randomly into 40,620 samples for training, 5,334 samples for validation and 5,334 samples for testing. The model is trained by SGD and learning rate keeps invariant all the time with initial value of 0.00001. The loss function used during training is mean squared error (MSE). Figure 3 is the learning curve during training showing the MSE loss of the model as a function the number of epochs for training and validation. The model converges after about 60 epochs. Table 1 and table 2 show the results of our experiments with the test set and comparison with HOTPANTS.

We compared our CNN model with HOTPANTS. In fact, HOTPANTS is not a deep learning method, so we use all the training, validation and test set to HOTPANTS. In Table 2, FNR (false negative rate) denotes the non-transient that should be subtracted but not subtracted in practice. Lower FNR, better performance it is. The FNR

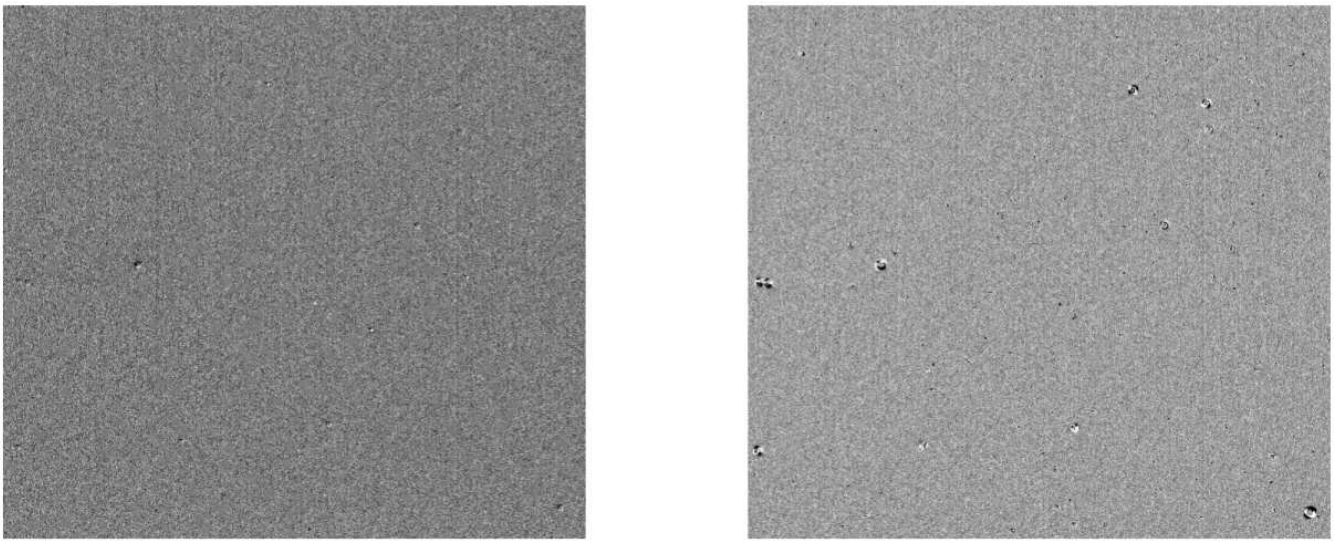

Fig. 4 Comparison between our method and HOTPANTS, left one is proposed method and right one is HOTPANTS.

Table. 1 Results of CNN model with test sets

	predicted positive	predicted negative	total
positive	2342	145	2487
negative	85	2500	2585
total	2427	2645	5072

Table. 2 Comparison of CNN model and HOTPANTS

	CNN model	HOTPANTS
FNR	6.48%	20.83%

of HOTPANTS is 20.83% while ours are only 6.48%, which is much better than HOTPANTS. As you can see in Figure 4, difference image of ours looks more clean than it of HOTPANTS.

6. Conclusion and Future work

We introduce a CNN model for transient detection using Akeno dataset. We perform image alignment on two time-continuous images of same filed and band, and make star pair for non-transient, transient candidate and background respectively. It is noted that transient candidate is scarce in real data, we produce a lot of them to balance the dataset. We also compare our method with a popular image subtraction method, HOTPANTS and it shows great improvement over it. Later, we consider revise the structure of the network to receive input of any size, since the shape of star varies greatly in real condition.

7. Acknowledgements

This work was supported by JSPSKAKENHI with grant number of 16K13783.

References

[1] B. Andrew. "HOTPANTS: High Order Transform of PSF and Template Subtraction", Astrophysics Source Code Library, record ascl:1504.004

[2] C. Alard, R. H. Lupton, "A Method for Optimal Image Subtraction", The Astrophysucak Journal, 503:325331, 1998 August 10

[3] C. Alard. "Image subtraction using a space-varying kernel", Astron. Astrophys. Suppl. Ser. 144, 363-370 (2000)

[4] D. M. Bramich. "A new algorithm for difference image analysis", Monthly Notices of the Royal Astronomical Society: Letters, Volume 386, Issue 1, 1 May 2008

[5] B. Zackey, E. Ofek, et al. "Proper image subtraction - optimal transient detection, photometry and hypothesis testing", The Astronomical Journal, 830:27 (23pp), 2016 October 10

[6] F. Masci, R. Laher, et al. "The IPAC Image Subtraction and Discovery Pipeline for the Intermediate Palomar Transient Factory", Publications of the Astronomical Society of the Pacific, Volume 129, Number 971

[7] D. Wright, S. Smartt, et al. "Machine learning for transient discovery in Pan-STARRS1 difference imaging", Monthly Notices of the Royal Astronomical Society, Volume 449, Issue 1, 1 May 2015, Pages 451466

[8] R. Romano, C. Aragon, C. Ding. "Supernova Recognition using Support Vector Machines", Machine Learning and Applications, 2006. ICMLA'06. 5th international Conference on. IEEE, 2006, pp.77-82

[9] M. Morii, S. IKEDA, N. TOMINAGA, et al. "Machine-learning selection of optical transients in the Subaru/Hyper Suprime-Cam survey", Publ. Astron. Soc. Japan (2016) 68 (6), 104 (18)

[10] T. Charnock, A. Moss. "Deep Recurrent Neural Networks for Supernovae Classification", arXiv:1606.07442

[11] A. Kimura, I. Takahashi, M. Tanaka, et al. "Single-epoch supernova classification with deep convolutional neural networks", 2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)