Cosmic Ray Background Detection and Removal from CCD Images using Neural Networks

I. M. Selim[1], Ibrahim El-henawy[2] & Mayada El-Sayed [2]
1- Dept. of CS Faculty of Computer and Artificial Intelligence, University of Sadat City
Dept. of CS Faculty of Computer and Information, Zagazig University

Abstract:

Over the last few years, Artificial Neural Networks (ANN) indicate favorable characterizations in accuracy and performance in different scientific fields, especially space since, Astrophysics-and physics. To obtain better quality of astronomical images the signatures of transient artifacts such as noise, cosmic rays, satellite trails and scattered cosmic light must be removed in the preprocessing operation. In this study we have developed an Artificial Neural Network (ANN) model for detection and removing the cosmic ray from the astronomical and space CCD images. The ANN proposed algorithm has been trained and tested using observational CCD image of data came from actual astronomical observations with the CCD camera system at kottamia 74 inch astronomical telescope in Egypt. We have been use this technique to identify, on a pixel-by-pixel level. The algorithm can be applied to any survey images. Also our result showing that the large amount of data give the more accuracy of results, it is also very fast in comparison with other algorithms found in astronomical data-processing.

Keywords: Artificial neural network; Cosmic rays; CCD images; image processing.

1. Introduction

At the end of the last century, the total amount of astronomical images more than 3 terabyte (TB). By the beginning of this century, the data size of Sloan Digital Sky Survey (SDSS) with multi-target has reached 40 terabyte (TB) [1] [2].

There are high-energy massive particles, high-energetic speed particle either an atomic nucleus or an electron that are generated outside the solar system. They are discovered by Victor Hess in 1912, who found that an electroscope discharged faster as ascended in a balloon to altitudes up to 5 km. Victor Hess concluded that called cosmic rays, it is arrived from outside our atmosphere and did not originate from decaying radioactive isotopes in the ground. A cosmic ray passing through the visual or infrared detector, can cause generation of a cluster of signal charge. These are energetic charged particles from outer space that travel at nearly the speed of light and accident with the earth at all directions. They are composed mainly of ionized nuclei, roughly 90 % protons, 10 % helium nuclei, and slightly under 1% heavier elements as well as electrons [3] [4]. Not only some of these particles originate from the sun, but also most come from sources outside the solar system. The most common problem of the cosmic ray is when the cosmic ray hits the CCD camera, it causes a bright spot on the image as shown in figure 1 (the field CCD image represent cosmic-ray map and the difference between cosmic ray and star in (a) and the profile of cosmic ray in (b)). Cosmic rays are usually easy to recognize because they are much sharper than stars (the high energy particle hits just a couple of pixels).

Planning to produce a good CCD image, it is very easy to clean out. However, removing them without damaging the real objects (stars or galaxies) can be trickier, but is possible. The cosmic ray must be detected and removed from the images without any distortion of the digital astronomical space image. The manually process of searching and removing of cosmic ray is an extremely tedious and difficult task because the discriminating between cosmic rays and stellar objects in CCD space images is very difficult. A relatively efficient approach to removing traces of cosmic rays from such images is to use multiple frames of the same object and then combine them using an algorithm for rejecting the outlying data. The space and astronomical digital images response over a large dynamic range of optical wavelength but the performance can be diminished, however, by the presence of cosmic ray in this image [5].

We planned to develop an automated Artificial Neural Network system capable to recognizing, locating, the cosmic ray in images and removing it. The proposed algorithm is based on the idea that cosmic rays deposit a portion of their energy in the pixels they hit, causing some extra signal in these pixels [6].

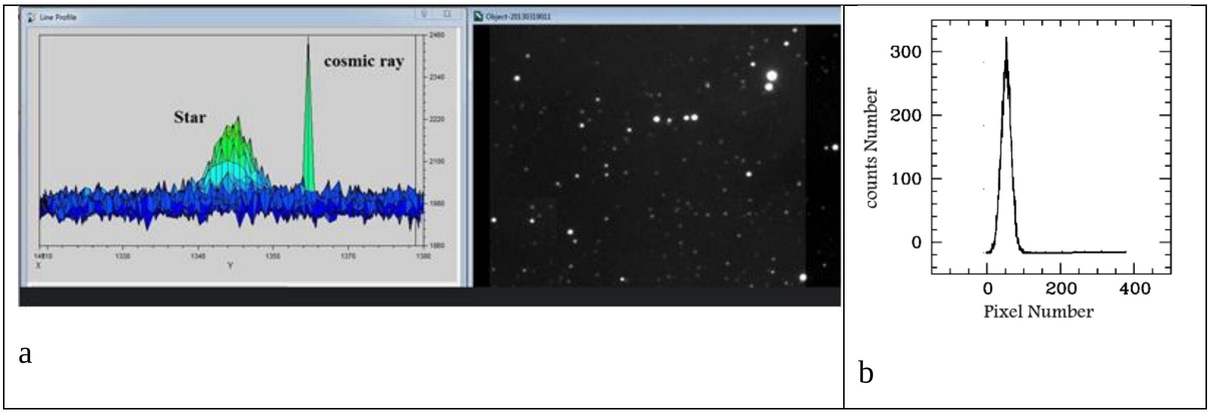

Figure 1. Example of field CCD image represent cosmic-ray map and the difference between cosmic ray and star in (a) and the profile of cosmic ray in (b)

The use of advanced computational models, particularly artificial neural networks, in the long-term forecasting of various environmental factors has recently seen a surge in increased application. We propose method for detecting and remove cosmic ray from space CCD images. Artificial Neural Network System can be learn from observing data set of images, as will as well suited to pattern recognition problems. Artificial Neural Network method is used as approximation useful tools for modeling complex, which means they modify themselves as they learn from initial training. These method help to estimate the effective and ideal methods for arriving at solutions and offers solutions for many problems.. An artificial neural network has been implemented to detecting and removing cosmic ray from astronomical space CCD images by python. We would like to stress here that our algorithm effectively removes cosmic rays while leaving almost all the image data untouched [7]. We have constructed a simple algorithm for detecting cosmic rays in CCD images. Our method need the model of the shape of the image features. We analyze the histograms of pixel counts in small sub frames in order to detect pixels deviating.

2. Related Work:

1- Srinadh Reddy et al. [8] have proposed a novel Dictionary Learning based framework to detect Cosmic Ray hits that contaminate the astronomical images obtained through optical photometric surveys. The unique and distinguishable spatial signatures of Cosmic Ray hits compared to other actual astrophysical sources in the image motivated us to characterize the Cosmic Ray patches uniquely via their sparse representations obtained from a learned dictionary. Specifically, the dictionary is trained on images acquired from the Dark Energy Camera observations. Next, the learned dictionary is used to represent the Cosmic Ray and Non- Cosmic Ray patches (e.g., each patch is with 11×11 pixel resolution) extracted from the original images. A Machine Learning classifier is then trained to classify the Cosmic Ray and Non- Cosmic Ray patches. They demonstrate that the proposed method provides additional guidance to the baseline models in terms of faster convergence rate and improves Cosmic Ray detection performance by 2% in the case of shallow models

2- Keming Zhang [9] they have found that Cosmic ray identification and replacement are critical components of imaging and spectroscopic reduction pipelines involving solid-state detectors. They present a deep-learning-based framework for Cosmic ray identification and subsequent image printing based on the predicted Cosmic ray mask. To demonstrate the effectiveness of this framework, they trained and evaluated models on Hubble Space Telescope images of sparse extragalactic fields, globular clusters, and resolved galaxies. They demonstrated that at a false-positive rate of 0.5%, achieves close to 100% detection rates in both extragalactic and globular cluster fields, and 91% in resolved galaxy fields, which is a significant improvement, compared with the best performing no neural technique tested. They presented their framework and the trained models as an open-source Python project.

2- Don Groom in [10] he has reached that Cosmic-ray muons make recognizable straight tracks in the new-generation CCDs with thick sensitive regions. Wandering tracks ("worms"), which we identify with multiply scattered low-energy electrons, are readily recognized as different from the muon tracks. These appear to be mostly recoils from Compton-scattered gamma rays, although worms are also produced directly by windows and field lenses.

4- Wolfram Freudling [11] he has proposed a method to find and remove cosmic rays from stacks of images which are not registered. Such dithered images obtained with under sampling cameras, such as the Wide Field and Planetary Camera 2 (WFPC2) on board the Hubble Space Telescope, can be used to recover some of the resolution lost by a large pixel size. The proposed method simultaneously cleans the images of cosmic rays. The output is a combined and restored image and a list of cosmic-ray hits for each of the input images. The final lists of cosmic-ray hits are useful even if a restoration of the images is not desired.

5- Shantanu et, al. [12] they have found that Astronomical images from optical photometric surveys are typically contaminated with cosmic rays. They have developed and tested an algorithm that removes these cosmic rays using a deep, artifact free, static sky coadd image built up through the median combination of point spread function, overlapping single epoch images. Transient artifacts are detected and masked in each single epoch image through comparison with an artifact free point spread function matched simulated image, model fitting catalog from the artifact of the single epoch image. This approach works well not only for cleaning single epoch images with worse seeing than the PSF homogenized coadd, but also the traditionally much more challenging problem of cleaning single epoch images with better seeing.

3. Proposed Method

An artificial neural network is a group of interconnected computing elements, or neurons, a neuron computes the sum of its inputs (which are either the outputs of other neurons or external inputs to the system) and each neuron has only one output but this output is multiplied by a weighting factor if it is to be used as the input to another neuron and passes this sum through a nonlinear function, such as a hyperbolic tangent or a hard threshold. The cosmic rays deposit a portion of their signal in the pixels they hit, causing some extra energy in these pixels. The image may have a large range of signal levels in different areas. The neurons are divided into layers, and each neuron is only connected to neurons in adjacent layers, which is shown schematically in Figure. 2. The actual neural network architecture makes use of a single hidden layer. The input layer contains 225 neurons, and the output layer has only one neuron, which provides the decision is identified the object whether a stellar object or cosmic ray. Signals from pixels belonging to the training and testing sub frames were presented to input neurons starting from the left corner along the x axis and successively row by row of image. The system is sufficient to solve the problem of discrimination between objects and have little time consuming during learning and operating phases. Furthermore, there are no feedback loops connecting neurons, so the signal produced by an external input move in only one direction (from left to right as Shown in Figure. 2) the output of the system is only the output of the last neuron in the network [9].

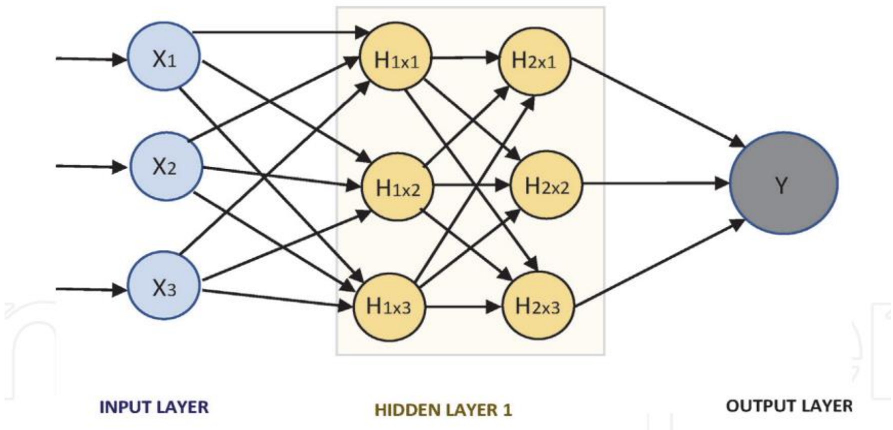

Figure2. Schematic diagram of an artificial neural network system [5].

The neurons of the adjacent layer, and this connection have been characterized by a weight w. Usually the neuron j in the hidden layer derives a weighted sum of all inputs applied to it (the vector x) and, through a function, produces an output value y = f (w, x) which is passed to the neurons in the next layer, and so on up to the last hidden layer (i = N - 1) which produces an output vector. In the cosmic ray detection, the output of the network will be a flag pixel at the point (X, Y) in the image. The output of the system is computed using equation 1, [13]:

$$Y^i = f \sum_i w_i x_i \qquad (1)$$

Where $\boldsymbol{Yi}$ is the output of the i th neuron of the N-1 level. The output of the i th hidden unit is obtained first by forming a weighted linear combination of the d input values, and then by adding a bias. And d is the number of the input, W (1)j,i denotes a weight in the first layer (from input i to hidden unit j). Note that wj,0 denotes the bias for the hidden unit j, and f is a activation. Note that there is only one output neuron in our network. The output of neuron j, would be as in equation 2.

$$out_j\left(x_j\right) = \frac{e^{x_j} - 1}{e^{x_j} + 1} = tanh\left(\frac{x_j}{2}\right) \tag{2}$$

Each neuron in our method computes a hyperbolic tangent function of its inputs. The input to neuron j, xj, is defined by equation 3 [11].

$$x_j = \sum_i w_{ij}\, out_i \tag{3}$$

Where $\boldsymbol{wij}$ is the connection weight from neuron í to neuron j, and out is the output of neuron i (or input i if neuron j is in the first layer). With the error ϵ is defined by equation 4 [9].

$$\epsilon = \frac{1}{2}\sum_p \left[out_p(w) - target_p\right]^2 \tag{4}$$

Where the summation is taken over all the training patterns which have been presented to the network in the training set.

4. Experimental Results and Discussion

4.1 Dataset Description

The data have been used in this work came from actual astronomical observations with the CCD camera system at kottamia 74 inch telescope in Egypt. Although in principle to train the network we chose to preprocess the data and extract discrete features from it. Because the observation is carried out under different conditions, background levels and average counts are varied from one image to another image. This means that a network can find cosmic ray in one image might not be able to find them in other images [12]. The preprocess is simplify the training system and use the data to calculate various attributes or features of the image at each point **(x, y)** and use these quantities as the inputs to the network. The coordinates of all the cosmic ray in the image region were found by visual inspection of the image and were then used as inputs to the network. 100 randomly chosen coordinates in that region of images were also supplied to the network as a positions which contain cosmic rays.

4.2. Result and discussion

The limitation of the presented method for these types of images comes from the fact that direct images usually have large count variations within small scales. This produces a large standard deviation of the counts and prevents cosmic rays in the neighborhood of bright objects from being detected. Feature extraction significantly ensure reduces the dimensionality of the input space, which reduces the number of trainings required to achieve good detection performance. Before calculating the feature values to be used by the network, the data were normalized to the average counts in the image according to equation 5:

$$N(X,Y) = 1 - \frac{c(X,Y)}{(c)} \qquad (5)$$

We analyze relatively small sub frames to work with a more concentrated local distribution of counts. Select small sized sub frames that cover the whole frame, with substantial overlap. The area of sub frame is **A (3x3)** (this area of thee pixel in row and three pixel in column is equal 9 square pixels). The signal coming from cosmic rays does not have a Gaussian distribution. The label on the current pixel is (**X, Y**), where **X** is the column number and **Y** is the row number. This should reflect in the distribution of counts in the sub frame of image affected by cosmic rays. The features described here are calculated for each sub frames of pixels in the image. The normalized intensity at each point of **N(X, Y)** is **I= N (X, Y)**.and the features are labeled from feature1{ **F_1)**to feature2 (**F_3**.) Each feature is labeled to indicate which network is used in.

Feature 1: Intensity in central sub frame

$$F_1 = N(X, Y). \qquad (6)$$

Feature 2: Intensity variance in maximum nearest pixels is computed using equation 6 [9]

$$F_2 = \left(\frac{1}{25}\sum_{i=-2}^{2}\sum_{j=-2}^{2} N^2(X+i, Y+j)\right)^{\frac{1}{2}} \qquad (7)$$

Feature 3: Number of pixels with intensities clustered around central pixel intensity is computed using equation 7 [13].

$$F_3 = \sum_{i=-2}^{2}\sum_{j=-2}^{2} max\,[N(X+i, Y+j) - D^*, 0] \qquad (8)$$

Where $D^* = \left\{1 - [1 - D(c,r)]^{\frac{2}{3}}\right\}$

For the detection the cosmic ray in our system, network given three input features, 5 neurons in the first layer, 6 neurons in the second layer, and one output neuron. The output of the network ranged between -1 and 1, with all outputs greater than the specific value with minimum error interpreted to mean that a cosmic ray was presented at the current point in the image.

Adjusting the weights is equivalent to moving the boundaries of the decision regions around until appropriate decisions are made by the network. The sequence steps is as the follows:

1) Set the weights to random values between - 1 and 1.
2) Present network with several input vectors ("training patterns") and network responses.
3) Calculate the output of the network for each of these training vectors.
4) Calculate the sum of errors between the outputs and the responses of the network.
5) Depending on whether the sub frame is a part of a cosmic ray or not, the error ϵ is defined by equation 8 [9]

$$\epsilon = \frac{1}{2}\sum_{p}\left[out_p(w) - target_p\right]^2 \qquad (9)$$

Where the summation is taken over all the training patterns which have been presented to the network in the training set. This procedure defines a positive error value ϵ for each possible combination of weight values. To find the minimum of ϵ in weight space training method terminates when the network error is less than a previously determined convergence criterion. In our implementation we decided to substitute the cosmic rays with the average of the counts in the neighboring pixels. We found that the performance of the system of detection and removing cosmic ray from the astronomical CCD images based on neural networks was optimized when we used the more data found the accuracy of results shows that the result accuracy of method we note that it decreases of error as the amount of data is increased as shown in figure 3. Instead the result accuracy of improves for large data sets, due to the presence of neural network is usually used for complicated tasks, such as image classification, image recognition.

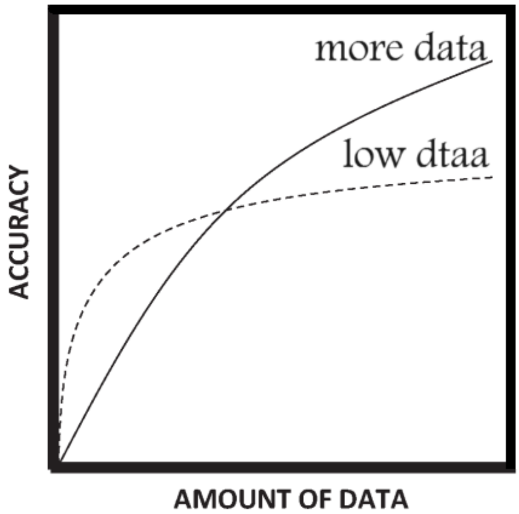

Figure 3. The accuracy of results with different set of amount of data

5. Conclusion

Recently, many digital sky surveys in a wide range of wavelengths are producing huge databases of astronomical image. Automated tool for analyzing astronomical data images and objects, such as galaxies, star, cluster, is very important tools, because rapidly collected data by modern projects is a need to create a good solution[14] [15]. This main that we required only three features to detect and removing cosmic ray from the astronomical space images. [16]. This is true for any computational or automated technique because neural networks (artificial intelligence) are thought as people often feel that they do not need to set up the problem. We have presented a cosmic-ray detection and removing from CCD images algorithm based on a artificial neural network and analysis of the images. The cosmic ray detection lists based on neural network is about 95% of all the cosmic rays. In a sense, the average difference in feature space between any cosmic ray and the background is larger than the difference between the cosmic ray and bright staller object The most important advantage of our method is that it may be can applied a priori to any type of images. We caution the reader that proper problem definition and setup are important for pattern recognition with neural networks, just as they are for any computational technique. However, it is important to be sure that the detection problem is well- posed and that the training set is as complete as possible [17]. Our method is to use the algorithm presented above for images data whenever multiple-image methods cannot be employed for cosmic-ray removal. The automated ANN detecting system can be quite effective for small exposure times. It begins by extracting relevant features from the data, and small probability of long cosmic ray signatures. It may be possible to check only images with exposure time of at least a few seconds and to find those images [18]. Recently, automated techniques coupled with computational developed astronomical system leading to better understanding of different stellar objects and their development process also leads to better understanding of the past, present, and future of our planet, solar system, galaxy and the universe [19]. We found that the performance of the system of detection and removing all cosmic ray from the astronomical CCD images based on neural networks was optimized when we used. By comparing the accuracy of results the graph shows that the result accuracy of method we note that it decreases of error as the amount of data is increased. Instead, the result accuracy of improves for large data sets, due to the presence of neural network is usually used for complicated tasks, such as image classification, image recognition.

6. References

[1] 1. S. G. Djorgovski, R. R. Gal, R. R. De Carvalho, S. C. Odewahn, A. A. Mahabal, R. Brunner, P. Lopes, and D. Team, The Digital Palomar Observatory Sky Survey (DPOSS):

General Description and the Public Data Release, Bulletin of the Astronomical Society of Brazil, vol. 34, no. 2, pp. 743, 2002.
[2] Mayada El-Sayed , Noha MM. Abdelnapi , I. M. Selim3 & Ibrahim El-henawy (IJCSIS), Vol. 19, No. 11, November 2021
[3] A. Hearn, M.F., Combi, M.R. Deep Impact at Comet Tempel 1. Icarus 187, 1–3, 2007.
[4] Alcock, G., et al. 1992, in Robotic Telescopes in the 1990s, ASP Conference Series, Vol. 34, ed. . V. Filippenko (San Francisco, ASP)
[5] I. M Selim, M Abd El Aziz IJAEBS vol1 no 1 pp1-8 2018
[6] M. Abd El Aziz, I. M. Selim, and S. Xiong, *Sci. Rep.*, vol. 7, no. 1, pp. 1–9, 2017.
[7] J. M. Caswell, Journal of Signal and Information Processing 5 (2014) 42-53.
[8] Vijay Laxmi & Harish Rohil. IJCA (0975 – 8887) Volume 99– No.8, August 2014
[9] NE Khalifa, MH Taha, AE Hassanien, I Selim 2018 International Conference on Computing Sciences and Engineering (ICCSE), 1-6
[10] Anik Daigle and Gilles Joncas Laval, Que´bec, QC G1K Marc-Antoine Miville-
[11] Desche^nes the Astronomical Society of the Pacific, 115:662–674, 2003 June
[12] L. Erguo, Y. Jinshou, Proceedings of the 4th World Congress on Intelligent Control and Automation, Shang Hai (2002) 2755-2759.
[13] Press, W. H., Flannery, B. P., Teukolsky, S.A., and Vetterling, W. T. 1986, Numerical Recipes: The Art of Scientific Computing (Cambridge, Cambridge University Press), Chap. 10
[14] M Abd El Aziz, I. M Selim, A Essam Experimental Astronomy 42 (1), 49-59
[15] F. Chollet, Xception: Deep Learning with Depthwise Separable Convolutions (2017), arXiv:1610.02357
[16] RA Ibrahim, M Abd Elaziz, AA Ewees, I. M. Selim, S Lu (2018) Journal of Astronomical Telescopes, Instruments, and Systems 4 (3), 038001
[17] Aggarwal CC. Neural Networks and Deep Learning. Cham: Springer International Publishing; 2018. DOI: 10.1007/978-3-319-94463-0
[18] Adams, C. (2021). Coreyjadams/cosmictagger: v1.0.1.
[19] I. M Selim, P Elkafrawy, W Dabour, M Eassa International Journal of Astronomy and Astrophysics, Astronomical Society of the Pacific 106: 532-541, 2021