Hybrid Imaging and Neural Networks Techniques for Processing Solar Images

R. Qahwaji and T. Colak

Department of Electronic Imaging and Media Communications

University of Bradford

Richmond Road, Bradford BD7 1DP,

England, UK.

E mail: r.s.r.qahwaji@brad.ac.uk ; t.colak@bradford.ac.uk

Abstract

Solar imaging is currently an active area of research. A fast hybrid system for the automated detection of filaments in solar images is presented in this paper. The system includes three major stages. The central solar region is detected in the first stage using integral projections. Intensity filtering and image enhancement techniques are implemented in the second stage to enhance the quality of detection in the central region. Local detection windows are implemented in the third stage to detect the positions of filaments and to define various sized arrays to contain them. The extracted arrays are fed later to a neural network for verification purposes.

Keywords: Solar Imaging; Space Weather; Filaments Detection; Segmentation; Neural Networks.

1. Introduction

Solar filaments are formed in magnetic loops. These loops hold relatively cool and dense gas that is suspended above the surface of the Sun [1]. Filaments are located in corona, where their temperatures are hundred times lower and their densities are hundred times greater than the respective corona values. In H-alpha images, filaments appear as dark ribbons against their brighter background. At the limb, they become bright features against sky, and are called prominence [2].

The "Hydrogen-alpha" (H-alpha) filter is a common filter for observing the Sun as it passes only the 6563 A H-alpha line while blocking all other wavelengths. H-alpha line is in the red part of the visible spectrum and is convenient for observing the fine structures of active regions, solar flares and mostly sunspots and filaments [3,4]. Currently there are many ground observatories around the world that provide visible light observations. Meudon Observatory, Big Bear Solar Observatory (BBSO), and Mauna Loa Solar Observatory (MLSO) are quite important in this aspect with their publicly available visible light observations.

Filaments detection and tracking is important for space weather forecast. The most dramatic solar events affecting the terrestrial environment are solar flares and Coronal Mass Ejections (CMEs) [5]. Flares and CMEs are solar eruptions that can spew vast quantities of radiation and charged particles into space [*6*]. In order to design an automatic system that can predict Flares and CMEs, the solar features that are highly associated with these activities must be detected. The association among CMEs and prominence/filament eruptions is high to conclude that these phenomena are physically related. In [7], it was concluded that 88% of halo CMEs were associated with flares and more than 94% were associated with eruptive prominences/filaments, while 79% of the CMEs were initiated from active regions.

The detection of filaments in solar images is a good example of segmentation in vision systems. Segmentation is the process that subdivides an image into the desired Regions of Interest (RoI), which differ from one application to another. Many segmentation techniques exist [8]. Histogram-based techniques work well under small signal to noise ratios, and objects with known sizes and simple background. The edge-based techniques depend on edge information to detect objects. The region-based techniques detect the regions that satisfy certain predefined homogeneity criteria. Markov random field-based techniques require prior knowledge of the distribution of the Markov field for RoI. Finally, hybrid techniques depend on merging two or more of the previous

techniques [9]. Hybrid-based segmentation that incorporates edge-based and region-based techniques is used in this work.

In [10], an algorithm for the detection, classification and tracking of filaments is introduced. The algorithm is composed of four main modules: (1) Image Acquisition, (2) Image Processing, (3) Filament Detection and Characterization, and (4) Filament Tracking. This algorithm can determine the correct chirality in more than 70% of the cases, when compared to the man-made list provided by [11]. In [12], H-alpha images were used to detect filaments using morphological closing operations with multi-directional linear structuring elements to extract elongated shapes. Filaments detection using seed selection and region growing was introduced in [13]. In addition, Neural Networks were used in [9] for filament recognition in solar images. A brief survey for filaments detection can be found in [10, 13]. Most of the existing techniques give good results, but sometimes neglect the smallest, weakest or blob-like filaments [13]. In this paper, our aim is to design a robust, speedy, efficient and automatic detection system that can extract the positions and main details of filaments. The detected regions will be verified using Neural Networks to enhance the accuracy of detection.

This paper is organised as follows: the Projection algorithm, which is implemented for the detection of the central region of the solar disk is described in Section 2. The image filtering and enhancement techniques are introduced in Section 3. Section 4 is devoted to the detection of filament regions. The verification using neural network is described in Section 5, while the practical implementation and evaluation of the detection algorithm is reported in Section 6. Finally, the concluding remarks are given in Section 7.

2. The Detection of the central Solar Disk

Most solar images suffer from limb darkening which is the gradual decrease in brightness of the solar disk as observed from its centre to its edge, or limb. One possible solution is to apply limb

darkening removal, in a manner similar to [14]. However, a simpler solution is to detect the central region of the solar disk in order to avoid the limb darkening effect and the detection of incomplete filaments that extend to or around the solar limb. The Filling algorithm presented in [15] could be applied for the detection of the central solar region [14]. The method presented in this work, which is based on the integral projections, can provide the same performance as in [14] but it is simpler in terms of its computational complexity and is faster.

The method explained in [16, 17] is implemented to find the integral projections. The vertical integral is computed by finding the average energy contained in every row, while the horizontal integral is computed by finding the average energy of each column. The vertical integral projection describes the vertical distribution of data in the image, while the horizontal integral projection describes the horizontal distribution of data. The average energy value for every row is represented graphically in the vertical position corresponding to that row as shown in Figure (1-a). The stages and results of applying the projection method are also shown in Figure (1). It is worth mentioning that on a 2.8 GHz machine, the filling algorithm takes on average 0.24 seconds to detect the central solar region when modified to handle 1024×1024 solar images, while the method presented here takes on average only 0.15 seconds.

Pseudo Code for the Projection Method

1. The input image is processed and the average energy for every row and column is found.

2. If the calculated energy is smaller than a threshold (A threshold value between 5 and 20 is suitable for most of the images) then its associated row or column is converted to black (zero values).

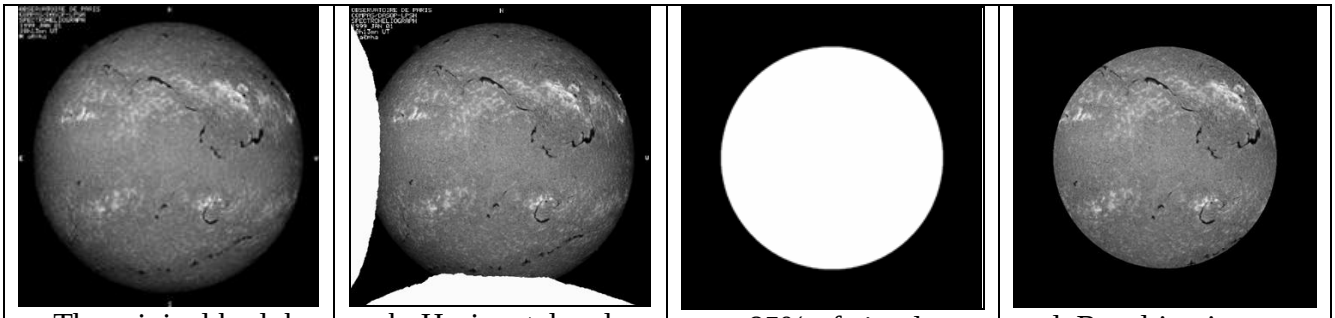
3. The centre of the solar disk and its radius are calculated as follows:

a. The vertical coordinate of the centre is equal to the centre of non-zero rows and the vertical radius is equal to half the number of these rows.

b. The horizontal coordinate of the centre is equal to the centre of non-zero columns and the horizontal radius is equal to half the number of these columns.

c. The radius is equal to the average of the horizontal and vertical radiuses.

4. Using the data for the centre of the disk and radius, %85 of the associated circular area is extracted as the central solar disk.



a. The original h-alpha image taken on 01/01/99 | b. Horizontal and vertical integral projections | c. 85% of circular regions | d. Resulting image

Figure 1: Detecting the central solar region:

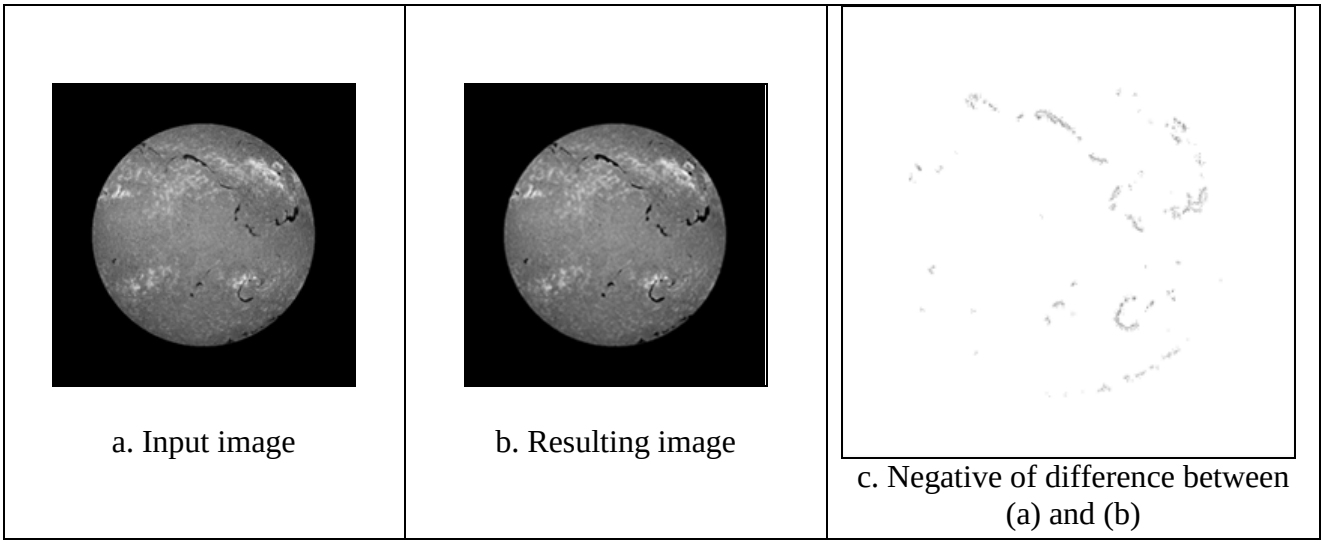
3. Initial Detection of Filaments using Local Enhancement and Intensity Filtering

After the central region is detected, image enhancement is applied (Figure 2), followed by local intensity filtering (Figure 3-a). Solar image enhancement aims to facilitate the detection process by increasing the grey-scale differences between filaments and their background using local sliding windows on the detected solar disk. Local enhancement is applied by dividing the image to number of sub images and then analysing the contents of these sub images. A window is centred on every sub image and the corresponding mean is calculated and then used to enhance its corresponding sub image. Later, intensity filtering is applied. The filaments and filaments-like regions are darker in colour; this enables an intensity filter with a low threshold to eliminate the background and active regions. This stage provides an image that contains candidates of filaments.

Pseudo-code of Enhancement

1. A 7 × 7 window is centred on each pixel and the mean for each window is calculated.
2. If the value of any pixel within the window is smaller than 60% of the mean value, then this pixel will be enhanced using Equation (1). This threshold ratio is found empirically.

$$pixel_{new} = pixel_{old} - \frac{pixel_{old} \times (mean - pixel_{old})}{mean} \quad \textbf{(1)}$$



a. Input image

b. Resulting image

c. Negative of difference between (a) and (b)

Figure 2: The implementation of the enhancement stage.

Pseudo-code of the Local Intensity Filtering

1. The enhanced image is divided into 32 × 32 sub images.
2. The mean value of each sub image is calculated.
3. If the intensity value of the pixel is smaller than 75% (empirically found) of the local mean value then it is given a false colour to indicate that they could be possible filaments.

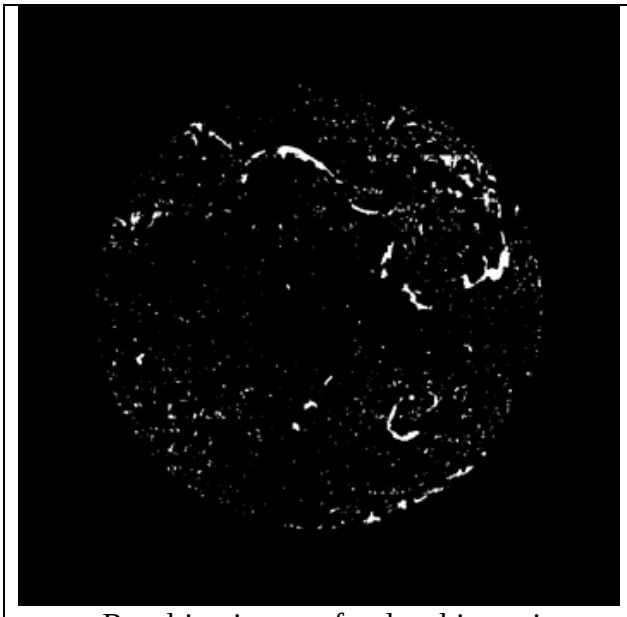
4. The Detection of the Filament Regions

The third stage aims to detect the exact locations of filaments with accuracy and speed. The window-based detection is supposed to filter the non-filament regions that were detected as possible

filaments by the second stage (Figure 3-b). After this process is completed, the detected filaments are stored in rectangular arrays.

Pseudo-code of Window-based Detection

1. In this method a window of size 11 × 11 is placed on each candidate pixel.
2. The number of other candidate pixels within the same window is calculated.
3. If this number is smaller than the thresholds of Equation (2), then the pixel in the centre is cleaned, other wise all the pixels within the window are given a false colour.



a. Resulting image after local intensity filtering

b. Resulting image after applying window based detection

Figure 3: Detection exact locations of filaments.

The threshold is calculated based on the size of the window as illustrated in Equation (2). In this case it is taken to be 25.

$$\text{Threshold} = \left[\frac{(\text{Size of Window}-1)}{2}\right]^2 = 25 \quad \textbf{(2)}$$

At the end of this process, the group detection algorithm defined in [14] is used to detect the regions. The detected rectangular regions are extracted from the image and represented in a data file that can be fed to neural network for verification. The detected rectangular regions could contain normal filaments or filaments that have complicated shapes (i.e. branches, curls, etc). This is due to the fact that the shape of the filament does not affect the detection. The detection algorithm is driven by the colour, size and location of filaments.

5. Verification of Filaments using Neural Networks

A verification stage is implemented on all the detected regions to enhance the accuracy of the detection. The verification is carried out using a neural network (NN) with back propagation training algorithm. Neural networks are chosen because of their known advantages, which can be found in [19, 20]. Compared to decision trees, neural networks would usually provide better performance for our application. As illustrated in [21], the hidden units in neural networks create new types of features that are used to separate classes. A comparison between the performance of neural networks and decision trees is carried out in [21], where it was found that the neural network is 3% more accurate than QUEST decision tree and 13% more accurate than usfC4.5 decision tree. The generalisation ability of neural networks makes it more suitable for predicting the correct solution for new input samples.

The choice of the training algorithm depends mainly on the performance of the error function during the training process [16]. A suitable learning algorithm should produce minimum error values and reach convergence quickly [16]. The back propagation training algorithm is widely used to train the multi-layer perceptron (MLP) networks and it was used in our application because it provides high degrees of robustness and generalisation in classification [16].

The NN training vector is constructed by extracting statistical features representing 100 filaments and non-filaments regions. Feature extraction is added to, simplify and speed the learning and testing phases, and to enhance the generalisation of the NN [17]. The extracted statistical features are: Skewness, kurtosis, standard deviation, mean and ratio of dark pixels to overall pixels. Skewness and kurtosis are the third and fourth moments respectively and they can capture the phase structure in the image (i.e. positions of edges and objects). These dimensionless measures would become zero for pure Gaussian distributions. Skewness can be considered as an indicator for the difference between the mean and the median of a dataset. On the other hand, kurtosis is based on the size of a distribution's tails relative to Gaussian [22]. Kurtosis provides a rough measure of image edginess. This is true because edges produce extreme band pass filter outputs, leading to heavy tailed (highly kurtotic) distributions [23].

The filament regions that are used for training are detected by the detection algorithm and are verified manually with the synoptic maps, as will be explained in the section to follow. The non-filament regions are extracted manually and they represent background regions that contain no filaments. The NN has 5 input nodes, one node for every extracted feature. The output layer has two nodes to indicate whether the detected region is a filament or not. The number of hidden nodes is 7 and they are located in a single hidden layer. This number is determined empirically based on the method proposed in [16]. In general, the number of hidden nodes must be small enough to permit generalisation and large enough to form an adequate internal representation of the input data. The NN provides better overall performance if it is optimised. Optimising a NN means that the best hidden layer nodes and the optimum learning times are reached [24; 16].

The NN training vector is computed by combining the extracted features for every region with the class to which it belongs. For every region, 5 features are extracted followed by 2 real numbers, which indicate the theoretical neural output for the input features. The output node that

corresponds to the desired class is given a high value (0.9), while the other nodes are given a low value (0.1). The training of the NN is considered to be successful when the NN manages to converge to a normalised system error of 0.001.

After the successful training, the neural network is coupled with the detection algorithm. Every H-alpha image is processed by the detection algorithm to extract the variable sized arrays that represent the extracted filaments. Feature extraction is applied later to these arrays to extract the 5 statistical features that will be fed to the neural network to determine whether the detected region is a filament region or not.

6. The Evaluation of Performance

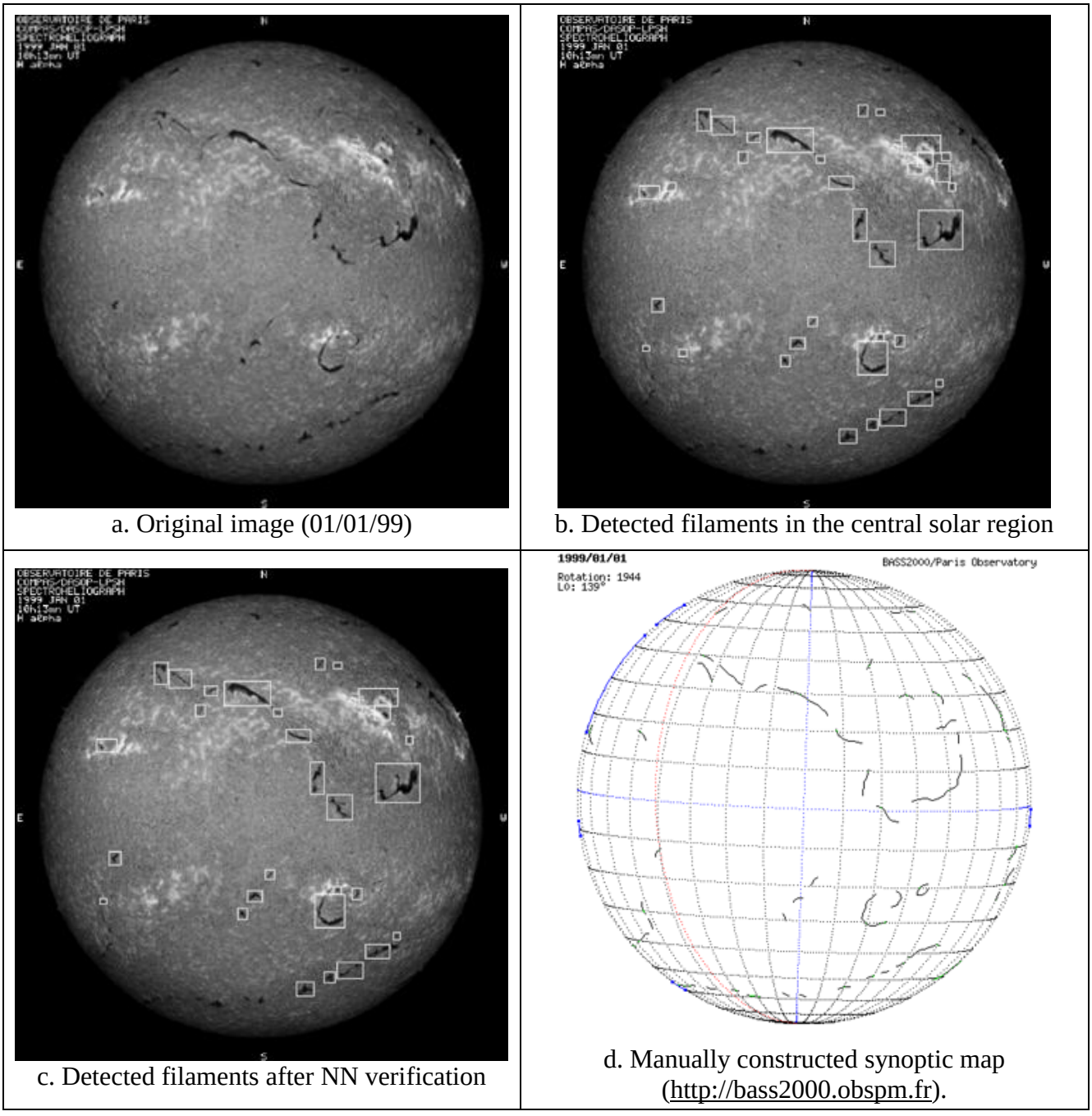
The detection technique is implemented fully in C++ and is applied to solar images that are obtained from Meudon observatory (http://bass2000.obspm.fr). The detection algorithm detects the filaments in a 1024 × 1024 image in about 1 second using P4-2.8 GHz PC with 1 GB RAM.

To evaluate the detection performance, the following two error rates are introduced [18]:

- The false acceptance rate (FAR): the probability of a non-filament being detected as a filament.
- The false rejection rate (FRR): the probability of a filament not being detected because it is considered to be a non-filament object.

The detection algorithm was applied to all the available Meudon H-Alpha solar images for the period from 01/01/99 till 17/02/99. For every H-alpha image used, a corresponding manually constructed synoptic map that contains the locations of filaments exists. The FAR and FRR error parameters are established by comparing the filaments that are detected using our algorithm with those detected manually and recorded in the synoptic maps. The synoptic maps are drawn manually using the subjective analysis of solar observers. Subjective analysis depends mainly of the

experience of the human operator, but it is also affected by fatigue and other human-related factors. On the other hand, the objective analysis of solar images, which is carried out by the automated detection system, provides consistent performance but its accuracy is usually lower. The detailed implementation of the filaments detection and verification is shown in Figure 4.



a. Original image (01/01/99)

b. Detected filaments in the central solar region

c. Detected filaments after NN verification

d. Manually constructed synoptic map (http://bass2000.obspm.fr).

Figure 4: The detection and verification of solar filaments.

The results are shown in Table 1. The first column shows the date of every H-alpha image, while the total number of filaments that are manually detected is indicated in Column 2. The FAR and FRR are shown in Columns 3 and 4, respectively and are denoted FAR (D) and FRR (D). It is important to note that the filaments that are considered are the ones that are located inside the central solar region, as illustrated previously in Section 2. The average error rate for FAR (D) is 17.2%, while the average error rate for FRR (D) is 12.9%. The error rates that are obtained after the NN-based verification stage are called FAR (V) and FRR (V). From Table (1) it can be shown that FAR (V) has dropped by 74.4%, while FRR (V) has increased by 23.3%.

Table 1: The FAR and FRR values for solar images after detection (D) and verification (V).

Date	#	FAR(D)	FRR(D)	FAR(V)	FRR(V)
01/01/99	28	17.9	3.6	3.6	7.1
02/01/99	20	10.0	10.0	5.0	10.0
03/01/99	22	22.7	9.1	13.6	9.1
05/01/99	19	47.4	0.0	15.8	5.3
06/01/99	20	0.0	10.0	0.0	15.0
08/01/99	19	21.1	10.5	0.0	15.8
09/01/99	19	0.0	21.1	0.0	21.1
10/01/99	20	25.0	10.0	5.0	30.0
13/01/99	22	18.2	13.6	4.5	13.6
14/01/99	23	0.0	13.0	0.0	13.0
15/01/99	17	35.3	5.9	5.9	5.9
18/01/99	27	11.1	22.2	7.4	22.2
22/01/99	21	42.9	14.3	9.5	19.0
26/01/99	19	0.0	15.8	0.0	15.8
27/01/99	20	0.0	15.0	0.0	15.0
29/01/99	24	8.3	16.7	0.0	20.8
30/01/99	26	3.8	26.9	3.8	30.8
31/01/99	19	0.0	10.5	0.0	15.8
05/02/99	22	4.5	4.5	0.0	4.5
09/02/99	12	33.3	8.3	8.3	16.7
11/02/99	17	5.9	23.5	5.9	23.5
12/02/99	18	44.4	11.1	11.1	16.7
14/02/99	23	4.3	17.4	0.0	21.7
16/02/99	31	61.3	12.9	9.7	12.9
17/02/99	24	12.5	16.7	0.0	16.7
Average	21.3	17.2	12.9	4.4	15.9

7. Conclusions

In this paper, a fully automated and fast algorithm for the detection of filaments in solar images is described. The detection process consists of three major stages: the detection of the central solar region, initial detection using enhancement and intensity filtering and the final detection of filament regions.

The algorithm is tested on solar images that are obtained from Meudon observatory. These images cover the period from 01/01/99 till 17/02/99. The detection algorithm is fast and it provides an FAR error rate of 17.2% and FRR error rate of 12.9%, when compared against the detected filaments in the manually constructed synoptic maps.

The detection performance is enhanced by adding a verification stage using neural networks. The verification stage ensures that the detected regions truly correspond to filaments. It increases the robustness of the algorithm and reduces the FAR error rate. In our case, the FAR has dropped from 17.2% to 4.4%. The algorithm is implemented in C++ and can be applied to various computer platforms. In the near future, the detection algorithm will be modified to detect other solar features (i.e., sun spots and active regions) in real-time mode. In addition, we are currently working on updating this algorithm to enable it to track filaments over consecutive days. Such tool can be important for space weather forecast as it can provide early warning for the disappearance of filaments and the possible occurrence of the associated Coronal Mass Ejections (CMEs).

References

[1] D. Hathaway, "What are solar filaments and prominences?" interview at http://www.spaceweather.com/glossary/filaments.html, Last Access: September, 2005.

[2] J. Gao, M. Zhou and H. Wang, “A threshold and region growing combined method for filament disappearance area detection in solar images,” 2001 Conference on Information Sciences and Systems, The Johns Hopkins University, 2001.

[3] H. Kurokawa, G. Kawai, R. Kitai, Y. Funakoshi, Y. Nakai, S. Tsuneta, T. Kosugi, S. Enome, L. W. Acton, and Y. Ogawara, "Detailed Comparison between H-alpha and Yohkoh Soft X-Ray Images of a Confined Two-Ribbon Flare" Publ. Astr. Soc. Japan, 44, 1992.

[4] M. Steinegger , C. Denker , P. Goode ,W. Marquette , J. Varsik, H. Wang, W. Otruba, H. Freislich, A. Hanslmeier, G. Luo, D. Chen, and W. Zhang, “The New Global High-Resolution Halpha Network: First Observations and First Results” paper presented at The solar cycle and terrestrial climate, Solar and space weather Euroconference. Proceedings of the 1st Solar and Space Weather Euroconference., ESA, Tenerife, Spain, September, 2000.

[5] M. Pick, C. Lathuillere, and J. Lilensten, “Ground based measurements”, ESA Space Weather Programme Feasibility Studies, Alcatel-LPCE Consortium, 2001.

[6] D. Lenz, “Understanding and Predicting Space Weather”, The Industrial Physicist, vol. 9, pp.18-21, 2004.

[7] G. Zhou, J. Wang and Z. Cao, “Correlation between halo coronal mass ejections and solar surface activity”, Astron. Astrophys., vol. 397, pp.1057-1067, 2003.

[8] K. Haris and S. Efstraiiadis, “ Hybrid Image Segmentation Using Watersheds and Fast Region Merging,” IEEE Transactions on Image Processing, vol. 7, No. 12, pp. 1684-1699, 1998.

[9] V. Zharkova and V. Schetinin, “A Neural Network Technique for Recognition of filaments in solar images," Proceedings of the seventh International Conference on Knowledge-Based Intelligent Information & Engineering Systems KES'03, Part I, pp.148-154, 2003

[10] P. Bernasconi, D. Rust and D. Hakim, “Advanced Automated Solar Filament Detection and Characterization Code: Description, Performance, and Results,” accepted for publication in Solar Physics, 2005.

[11] A. Pevtsov, K. Balasubramaniam, and J. Rogers, “Chirality of Chromospheric Filaments,” Astrophys. J., vol. 595, pp. 500, 2003.

[12] F. Shih, and A. Kowalski, “Automatic extraction of filaments in H-alpha images,” Solar Physics, vol. 218, pp. 99 – 122, 2003.

[13] N. Fuller and J. Aboudarham, "Automatic Detection of Solar Filaments versus Manual Digitization," the 8th International Conference on Knowledge-Based Intelligent Information & Engineering Systems (KES2004), Springer Lecture Notes in Computer Science, vol. 3, pp. 467-475, 2004.

[14] R. Qahwaji and T. Colak, “Automatic Detection and Verification of Solar Features”, accepted for publication in International Journal of Imaging Systems and Technology,2005.

[15] R. Qahwaji and R. Green, “Detection of Closed Regions in Digital Images,” International Journal of Computers and Their Applications, vol. 8, pp. 202-207, 2001.

[16] J. Kim, A. Mowat, P. Poole and N. Kasabov, “Linear And Non-Linear Pattern Recognition Models For Classification Of Fruit From Visible-Near Infrared Spectra,” Chemometrics And Intelligent Laboratory Systems, vol. 51, pp. 201-216, 2000.

[17] D. Horn, “Neural Computation Methods and Applications: Summary Talk of the AI”, Journal of Nuclear Instruments and Methods in Physics research, vol. 389, pp. 381-387, 1997.

[18] L. Hong and A. Jain, “Integrating faces and fingerprints for personal identification,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, No. 12, pp. 1295-1307, 1998.

[19] J. Mao and A. Jain, “Artificial Neural Networks for Feature Extraction and Multivariate Data Projection,” IEEE Transactions on Neural Networks, vol. 6, no. 2, pp. 296-317, 1995

[20] B. Lerner, H. Guterman, M. Aladjem and I. Dinstein, “A Compartive Study of Neural Network Based Feature Extraction Paradigms,” Pattern Recognition Letters, vol. 20, no. 1, pp. 7-14, 1999.

[21] L. Hall, X. Liu, K. Bowyer and R. Banfield, “Why are Neural Networks Sometimes Much More Accurate than Decision Trees: An Analysis on a Bio-Informatics Problem,” IEEE International Conference on Systems, Man & Cybernetics, Washington, D.C., USA, 2003.

[22] E. Reinhard, P. Shirley and T. Troscianko, "Natural image statistics for computer graphics", University of Utah tech report UUCS-01-002, March 2001.

[23] R. Dror, E. Adelson and A. Willsky, “Recognition of Surface Reflectance Properties from Single Image under Unknown Real-World Illumination,” Proceedings of the IEEE Workshop on Identifying Objects Across Variations in Lighting: Psychophysics & Computation Collocated with CVPR 2001. Kauai, Hawaii, December 2001.

[24] Q. Ma, A.Yan, Z. Hu, Z. Li and B. Fan, “Principal Component Analysis and Artificial Neural Networks Applied to the Classification of Chinese Pottery Neolithic age”, Analytica Chimica Acta Journal, vol. 406, pp.247-256, 2000.