Segmentation of spectroscopic images of the low solar atmosphere by the self-organizing map technique

F. Schilliro ⓘ★ and P. Romano ⓘ★

Osservatorio Astrofisico di Catania, INAF, via S. Sofia n.78, I-95123 Catania, Italy

ABSTRACT

We describe the application of semantic segmentation by using the self-organizing map technique to an high spatial and spectral resolution data set acquired along the H α line at 656.28 nm by the Interferometric Bi-dimensional Spectrometer installed at the focus plane of the Dunn solar telescope. This machine learning approach allowed us to identify several features corresponding to the main structures of the solar photosphere and chromosphere. The obtained results show the capability and flexibility of this method to identifying and analysing the fine structures which characterize the solar activity in the low atmosphere. This is a first successful application of the SOM technique to astrophysical data sets.

Key words: methods: data analysis – Sun: chromosphere – Sun: photosphere – sunspots.

1 INTRODUCTION

With the rise in big data, machine learning has become particularly important for solving data science problems in different areas (Finance, Image processing, Computational biology, Automotive, Aerospace, Natural language processing) (Aggarwal 2018). Machine learning algorithms and techniques allow scientists to automatize computer learning in order to make classifications, decisions, and predictions without the continuous interaction with systems under analysis. In general machine learning, as well semantic segmentation in particular, use two types of approaches: supervised learning, almost focusing on the 'deep learning' methodologies, e.g. based on Convolutional Networks (Shelhamer et al. 2017; Hausen & Robertson 2020), and unsupervised ones. The former algorithms train a chain of neural networks by processing a known input and correlated output data, so that it is possible to predict future outputs by learning from 'experiences'. Instead, less used unsupervised learning method (Barlow 1989; Zeljko et al. 2014), find hidden patterns or intrinsic structures inside the input data, allowing data classification or segmentation in a non 'a priori' organized fashion.

The application of machine learning techniques has also a great potentiality in the space weather field in general, and for solar images in particular, with the opportunity to identify new indicators for the forecast of eruptive events occurring on the Sun, such as flares and Coronal Mass Ejections (see Benz 2017), and producing effects in the near-Earth space environment, as well as at the Earth surface.

In fact, the scientific community has developed a growing interest for the automatic identification of the signatures characterizing the sources of disturbances whose radiative, field, and particle energy directly impact the Earth (e.g. Barnes et al. 2016; Florios et al. 2018; Kontogiannis et al. 2018; Falco, Costa & Romano 2019; Kim et al. 2019; Korsós, Yang & Erdélyi 2019). In particular, the characterization of solar active regions allows us to identify the topological configuration which are suitable for the occurrence of sudden, rapid, and intense solar eruptions (e.g. Romano et al. 2015, 2018; Romano, Elmhamdi & Kordi 2019).

In this paper, we consider for the first time the application of the semantic segmentation by using 'self-organizing maps' (SOM) technique (Kohonen 1995), to high-resolution spectroscopic observations of the low solar atmosphere. We used a data set taken during an observing campaign carried out on May 2015 at the NSO/Dunn Solar Telescope (DST). The data set showing a sunspot and its main structures visible at photospheric and chromospheric level is described in the next section. We describe the technique applied to get a skeletonization of the main structures in Section 3. Section 4 reports the main results and the advantages of the application of this new technique to solar images. The last section is dedicated to draw our conclusions.

2 DATA

We used data acquired by the Interferometirc Bidimensional Spectroscopic Instrument (IBIS; Cavallini 2006) which operated at the NSO/DST. The data set consists of high-resolution data, with a pixel scale of 0''.095, taken on 2015 May 18, in 17 spectral points along the H α 656.28 nm line (see Romano et al. 2017, for further details).

We selected the scan of the H α line taken during the best seeing conditions of that day of observation.

The original field of view (FOV) of IBIS data was 500×1000 pixels, but we considered in our analysis only a sub-FOV of 350×700 pixels centred on the preceding sunspot of the AR NOAA 12348. In this way it was possible to analyse the results obtained by the application of our method of analysis to the main structures visible in a sunspot and in the surrounding environment (Romano et al. 2020).

We also used broad-band images at 633.3 nm acquired simultaneously with the spectral frames, imaging the same FOV with the same exposure time, in order to restore the spectral images using the Multiframe Blind Deconvolution (MFBD; Löfdahl 2002) technique.

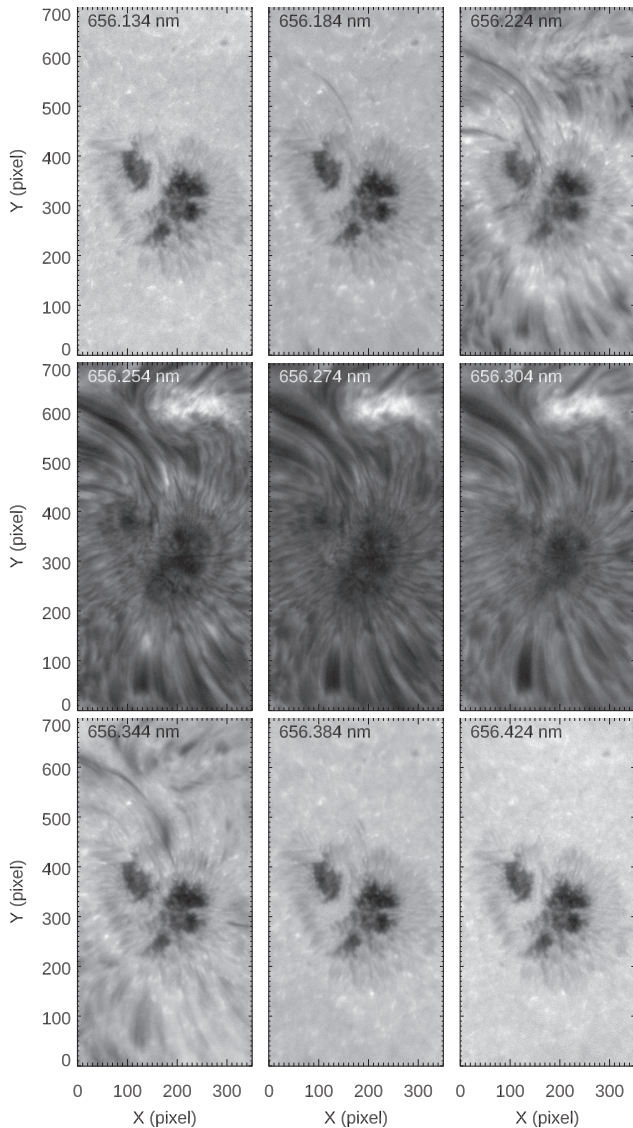★ E-mail: francesco.schilliro@inaf.it (FS); paolo.romano@inaf.it (PR)

Figure 1. Sample of the spectral images acquired along the Hα line by IBIS and used in our analysis. All the images have been acquired on May 18 at 14:42 UT.

The MFBD allowed us to reduce the seeing degradation, achiving a spatial resolution of about 0$''$3 at 656.28 nm.

In Fig. 1 we report some of the spectral images used for our analysis. We can see the different solar structures visible in our FOV varying the spectral points along the Hα line. In the spectral points near the continuum (top left and bottom right-hand panels of Fig. 1) and in the wings of the spectral line (top right and bottom left-hand panels of Fig. 1) we are able to identify the main fine structures of the sunspot and the surrounding environment typical of the photospheric level, such as the umbra, the penumbral filaments, a light bridge, the granular, and intergranular pattern (see Solanki 2003, for an overview of these structures). Near the centre of the Hα line (middle row panels of Fig. 1) we clearly see the chromospheric structures, such as the superpenumbra, some facular regions in the top right corner of the FOV, and a filament portion in the top left corner of the FOV (see Leka & Metcalf 2003).

In order to better describe the potentiality of the SOM technique in the segmentation of our spectral data set, we also considered

some physical parameters which can be deduced by the analysis of the Hα line profile. We reconstructed the profile of the Hα line in each spatial pixel by fitting the corresponding signals obtained in the monochromatic images with a linear background and a Gaussian shaped line. Using this reconstructed profile, we performed an estimation of the line depth (LD) by

$$\frac{(F_c - F_o)}{< I >},\tag{1}$$

where F_c and F_o are the maximum and the minimum intensities of the line profile, taken in the continuum and in the centroid of the Gaussian fit, respectively, while $< I >$ is the average intensity of the quiet Sun, measured in a box of 10×10 pixels taken in the southern portion of the FOV, outside of the sunspot where few chromospheric structures were visible. Using the full width at half-maximum (FWHM) of the Gaussian fit we estimated the spectral line width (LW). We also measured the value of velocity along the line of sight (LOS) in photosphere from the Doppler Shift (DS) of the centroid of the reconstructed line profiles in each spatial point with respect to the reference wavelength (656.28 nm).

3 THE SELF-ORGANIZING MAP TECHNIQUE

SOM is an important class of neural networks capable of unsupervised autonomous learning from data, adapting according to rules of synaptic plasticity in response to external stimuli. SOM algorithms are inspired by neurobiological studies and indicate that different sensory inputs (motor, visual, auditory, etc.) are mapped on to corresponding areas of the cerebral cortex in an orderly fashion. This form of map, known as a topographic map, has two important properties:

(i) At each stage of representation, or processing, each piece of incoming information is kept in its proper context/neighbourhood.

(ii) Neurons dealing with closely related pieces of information are kept close together so that they can interact via short synaptic connections.

A particular kind of SOM known as a Kohonen Network (see Kohonen 1995, for an overview), belongs to the class of vector-coding algorithms; they offer a mapping topological which places a fixed number of vectors (code words) in an input space oversized, thus allowing data compression, data clustering, or data classification. This ability to extract information from the data and understand how these group spontaneously into cluster, is allowed by a neural structure with a single computational layer arranged in rows and columns, in which each neuron is fully connected to other and to all the source nodes in the input layer. The self-organization algorithm is composed in general of four major process that involves input layer, as source of data to be processed and neural fully connected lattice, that adapts its structure and morphology in order to replicate the input organization of data according to their neighbourhood. In our study, this lattice morphology is used to semantic grouping pixel in regions, the features, that are detected according to the distance of the multispectral set of input data collected for each pixel and the normalized random initial formatting of the net, as depicted in Fig. 2.

(1) Initialization: Input pattern is a normalized data-cube composed by a K-dimensional vector of spectral values for all the N $\times$ M pixels, as described in Fig. 2. All the K-dimensional connection weights for each neuron and for each input value, w_{ji}, are initialized with normalized random values.

(2) Competition: For each input pattern, the neurons compute their respective values of a discriminant function which provides
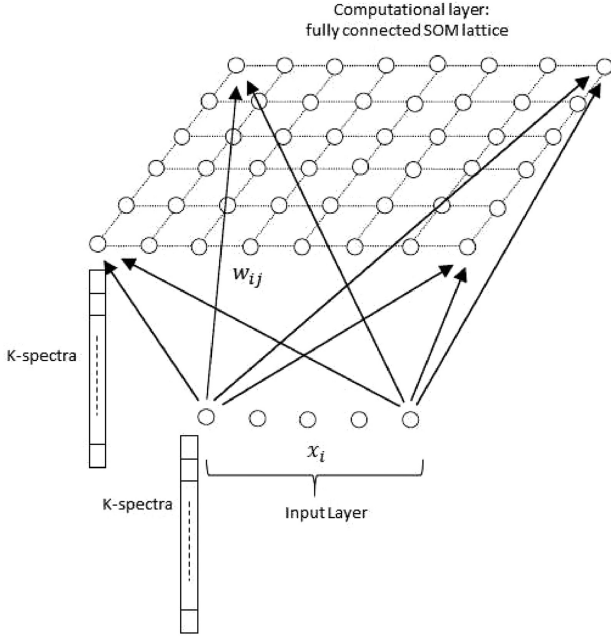
Figure 2. Neural lattice of a Kohonen Network: the $N_1 \times N_2$ fully connected lattice compute K-spectra vector input layers as vector of spectral values associated to each pixel of image under processing.

the basis for competition. For the $D = N \times M$ pixel/neurons of K-dimensional spectral input space, discriminant function is calculated as the spectral Euclidean distance between each input pattern ($x = (x_i: i = 1, \ldots, D)$) and each connection weights between the input units i and the neurons j in the computation layer, ($w_{ji} = x_{ji}: j = 1 \ldots N_1 \times N_2, i = 1, \ldots, D$), where ($L = N_1 \times N_2$) is the neural lattice dimension.

$$\Delta_i(\bar{x}) = \sum_{i=1}^{D} (\bar{x}_i - \bar{w}_{ji})^2 \qquad (2)$$

The neuron $E(x)$ whose weight vector comes closest to the input vector (i.e. is most similar to it) is declared the winner, so the continuous input space can be mapped to the discrete output space of neurons by a simple process of competition between the neurons.

(3) Cooperation: The winning neuron $E(x)$ determines the spatial location of a topological neighbourhood of excited neurons, thereby providing the basis for cooperation among neighbouring neurons, what in neurobiological studies is the lateral interaction within a set of excited neurons. When one neuron fires, its closest neighbours tend to get excited more than those further away, and this decays with distance according a Gaussian law :

$$G_{\delta,E(x)} = e^{\frac{-S^2_{\delta,E(x)}}{2\sigma^2}}, \qquad (3)$$

where

    (i) $E(x)$ is the winning neuron according x input vector;

    (ii) δ is the general index for neuron of the lattice;

    (iii) $S_{\delta, E(x)}$ is the lateral distance between $E(x)$ and δ, calculated as the Euclidean distance on each spectral K-dimensional vector;

    (iv) σ is the size of topological neighbourhood, typically shrinking with time (expressed in number of iterations as well) as the function $\sigma(t) = e^{\frac{-t}{\tau_0}}$, with σ_0 and τ_0 constant set by the neural network designer and respectively the initial neural
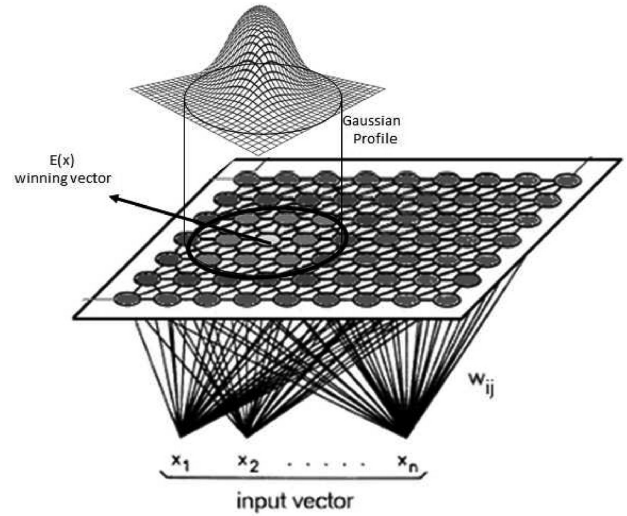

Figure 3. Group of cooperating neurons by application of Gaussian topological neighbourhood operator.

proximity parameter and the mean time of activation persistence by proximity.

Cooperation process is depicted in Fig. 3, where is highlighted the group of involved neurons by application of Gaussian topological neighbourhood operator.

(4) Adaptation: The excited neurons decrease their individual values of the discriminant function in relation to the input pattern through suitable adjustment of the associated connection weights, such that the response of the winning neuron to the subsequent application of a similar input pattern is enhanced. The adaptation of the SOM designed is the process responsible for the feature map formation, that links the input data to the self-organized neuronal lattice. The point of the topographic neighbourhood is that not only the winning neuron gets its weights updated, but its neighbours will have their weights updated as well, although by not as much as the winner itself. In practice, the appropriate weight update equation in term of weights updating is

$$\Delta w_{ji} = \eta(t) G_{\delta,E(x)}(t)(\bar{x} - \bar{w}_{ji}) \qquad (4)$$

$$\eta(t) = \eta_0 e^{\frac{-t}{\tau_n}} \qquad (5)$$

Upon repeated presentations of the training data, the synaptic-weight vectors w_j tend to follow the distribution of the input vectors because of the neighbourhood updating. The algorithm therefore leads to a topological ordering of the feature map in the input space in the sense that neurons that are adjacent in the lattice will tend to have similar synaptic-weight vectors, towards the input vector x.

The learning-rate parameter $\eta(t)$ is responsible for speed of adaptation process, should also be time varying, as indicated in equation (4), and is time shrinking according the learning ability η_0 and its persistence in time τ_n.

4 CLUSTERING CRITERION AND SOM FEATURES SELECTION

The operation of image segmentation is basically the association of each pixel x_i of a frame, that represents a K-vector spectral data set, to a unique cluster C_l according to a method that in our case is a

SOM. A segmented region associated to each different cluster could be expressed as 'feature' $\rho_{k,i}$ as follows:

$$\rho_{k,i} = \begin{cases} 1 & \forall x_i \in C_l \\ 0 & \text{otherwise} \end{cases}$$

so that it is possible to define the number of pixel belonging to C_l as $|C_l| = \sum_{i=1}^{L} \rho_{l_i}$

In cluster analysis two strategies are possible in finding a possible optimal data classification, a supervised and an unsupervised approach. While the first one is adopted when segmentation of region in features is well specified and tested, the unsupervised one is a method in which the number of features is not previously specified, and it arises from data structure as a result of a data processing that produces different indices of clustering.

In our case the number of clusters or 'features' in not known 'a priori', because a class is directly connected to the presence of a region of interest that could be in the tile under study or not. Therefore, it could be useful to find inside the data structure, whether an optimal number of possible groups of data separated each others exists, each of which corresponding to different regions composing the image frame.

For the purpose of discovering the number of independent regions, that becomes important in the stage of design the SOM network, different clustering indices and parameters were investigated, as the Calinski-Harabasz (CH) index (see Calinski & Harabasz 1984), the Silhouette index (see Rousseeuw 1986), and Davies-Bouldin (DB) index (see Davies & Bouldin 1979). In particular, DB has been already used in literature for solar EUV data clustering (see Caballero & Aranda 2013), confirming that is a good and fast algorithm which converges with the analysed data. In our case, we found that DB index provided same results in the classification of optimal number of independent regions considering not only the clustering of the $N \times M \times K$ spectral data-cube, but also the clustering of the $N \times M \times 3$ data-cube containing the measured parameters of the spectral line (LW, LD, and DS).

In Fig. 4 we report the graphs of the DB index for the spectral data-cube and the parameters data-cube. They show the best result (i.e. a minimum) when the number of clusters is equal to 16, which can be considered the best number of regions for the segmentation and design of the SOM lattice. Anyway using a more (less) populated neural lattice with a number of features more (less) than 16, makes effective the semantic segmentation splitting (merging) the different regions in components that keep the same nature.

The algorithm complexity and relative time of execution is not hardly affected by SOM network complexity, that depends on the lattice structure. The choice of a number of features according to relative minima in the DB index is determined by the required detail level for the segmentation of the data; in our case we considered with particular attention the case of 16 regions designing a 4×4 SOM neural lattice. The choice of $N_1 = 4$ and $N_2 = 4$, complete the overall architecture for the segmentation algorithm, that really needs some trimmer for the values related to $\sigma_0, \tau_0, \eta_0, \tau_n$. This architecture is depicted in Fig. 5.

5 RESULTS

The algorithm defined in Fig. 5 runs after a phase of preparation of the input data cube which foresees the adaptation of the image to the appropriate size (in this case 350×700 pixels) and the normalization of the brightness values of each pixel for each wavelength of the data
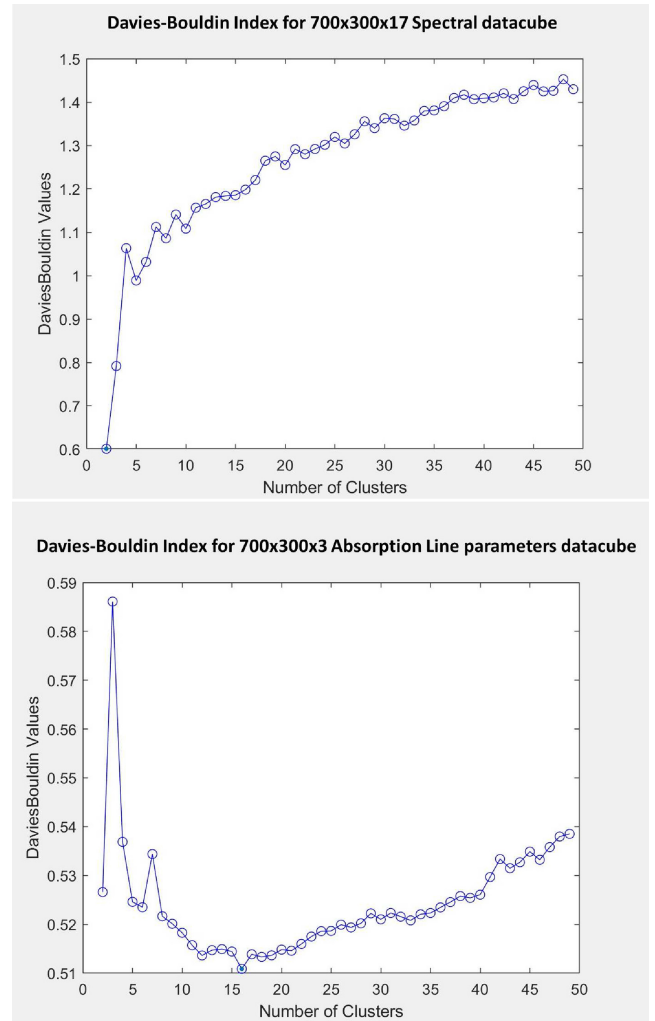

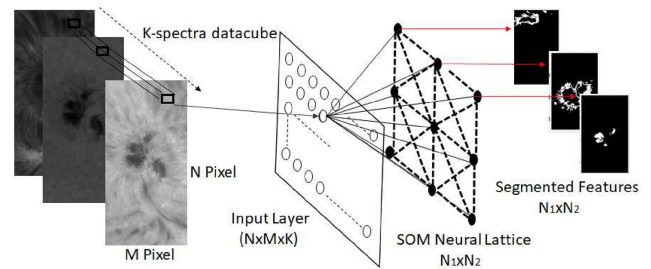**Figure 4.** Davies–Bouldin index for multispectral (up) and parameters (bottom) datacube.

Figure 5. Architecture for a semantic segmentation using self-organizing map with a 4×4 computational neural lattice.

cube. The SOM network is organized in such a way as to guarantee segmentation in complementary regions, so that the results of the elaboration are 2D maps, which display the features, i.e. semantically separated regions.

One of the main advantage of the SOM technique is its flexibility in the choice of the number of features. In fact, according to the needs of the user it is possible to segment the data set with different level of accuracy. As mentioned above, on the base of the DB index we decided to consider the results obtained for 16 clusters (Fig. 6)
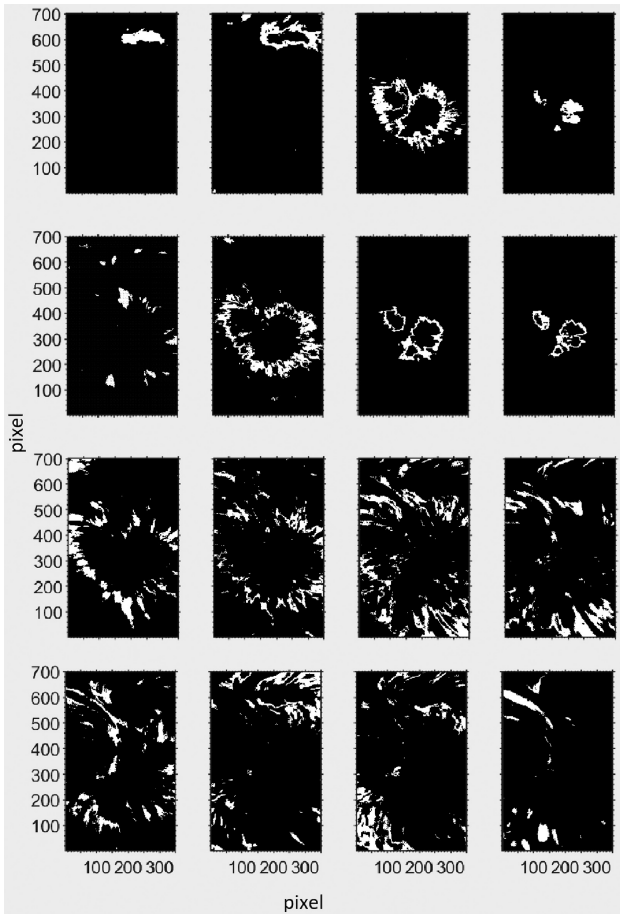
Figure 6. Feature lattice elaborated by 4×4 SOM and performing the semantic segmentation of 16 different regions.

Figure 7. Feature lattice elaborated by 3×3 SOM and performing the semantic segmentation of nine different regions.

and to describe in detail the different features found in this particular case. However, a different number of clusters, e.g. 9 (Fig. 7), results in the merging of some features obtained for a higher value (compare Figs 6 and 7).

We named each feature by a progressive number from 1 to 16 and we note that they are able to segment in great detail all the structures visible in the FOV at different wavelengths. Moreover, we found an interesting accordance of the features not only with the spectral properties of the segmented photospheric and chromospheric structures, but also with some of their physical parameters. For example, we can see that features n. 1 and 2 identify the chromospheric facula at North. The core of the facula is well segmented by feature n. 1, while feature n. 2 corresponds to the surrounding pattern, visible in the core of the Hα line (see the middle panels of Fig. 1), where the brightness of the facula is limited by the presence of the neighbouring fibrils (top left-hand panel of Fig. 8). In this case, we note that the differences between features n. 1 and n. 2 are not only in their values of intensity in the core of the Hα line, but also in the different values of the line depth (bottom left-hand panel of Fig. 8). Features n. 1 and n. 2 correspond to LD greater than 0.8 and 1.2, respectively. Instead, we do not see any correlation of these features with LW (top right-hand panel of Fig. 8 and the LOS velocity (bottom right-hand panel of Fig. 8) of the corresponding structures.

The umbra of the sunspot is well segmented by features n. 4, 7, and 8 (see the contours in Fig. 9). Combining these three features we are able to cover the whole umbra, while considering only some of them
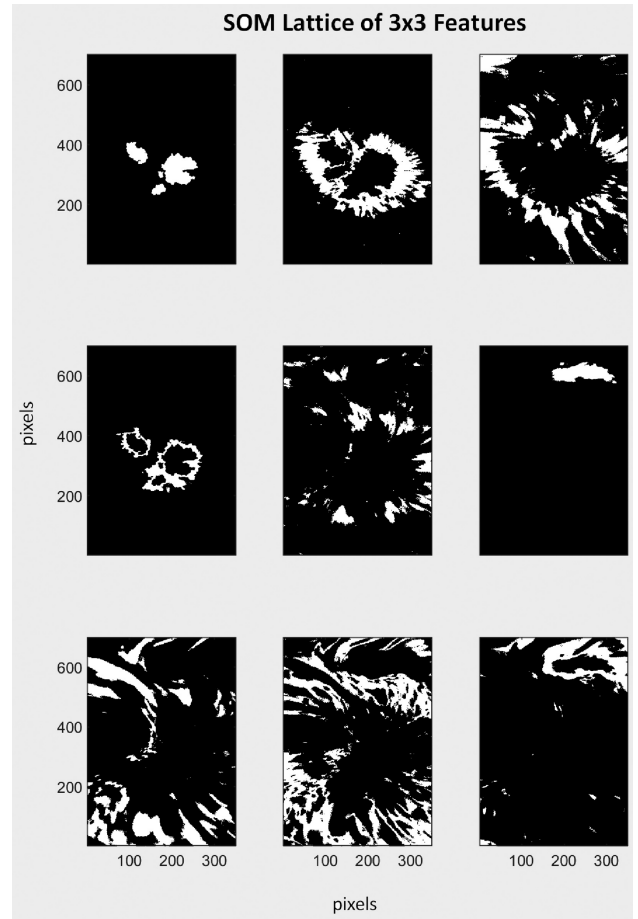
we are able to identify different subregions, gradually more extended. Features n. 4 corresponds to the umbra characterized by $Il/<I> < 1.0$ at 656.134 nm, while features n. 7 and n. 8 segment the portions of the umbra with $1.5 < Il/<I> < 1.7$ and $1.0 < Il/<I> < 1.5$, respectively. LD does not show any significant variation in the umbra (see the bottom left-hand panel of Fig. 9). Only LW and the LOS velocity seem to characterize these features. In particular, it is noteworthy that feature n. 7 corresponds to the portion of the umbra with a velocity of about 0.5 km s^{-1} (bottom right-hand panel of Fig. 9).

The penumbra of the sunspot is usually a more complicated structure to be segmented, due to the strong inhomogeneities, in the azimuthal direction as well as along the line of sight. It seems that the penumbra consist of dark filaments flanked by lateral brightenings, but also we can see bright filaments formed by a few penumbral grains that are radially aligned. Nevertheless, features n. 3, 6, 9, 10, 11, and 12 are useful to identify the different parts of the penumbra. In particular, features n. 9 and 10 seem to be useful to discriminate among dark and bright regions forming the fine structure of the penumbra at photospheric level, while feature n. 11 and 12 identify the superpenumbra at chromospheric level (see Fig. 10). Taking into account the physical properties corresponding to these features, we can see that the features n. 3, 6, ad 9 correspond to different levels of LW, LD, and LOS velocity, although due to the Evershed flow (Evershed 1909) we cannot associate specific values of LW and DS to each feature. Only the LD shows different levels for each feature, in fact features n. 3, 6, and 9 are characterized by LD around 0.4,
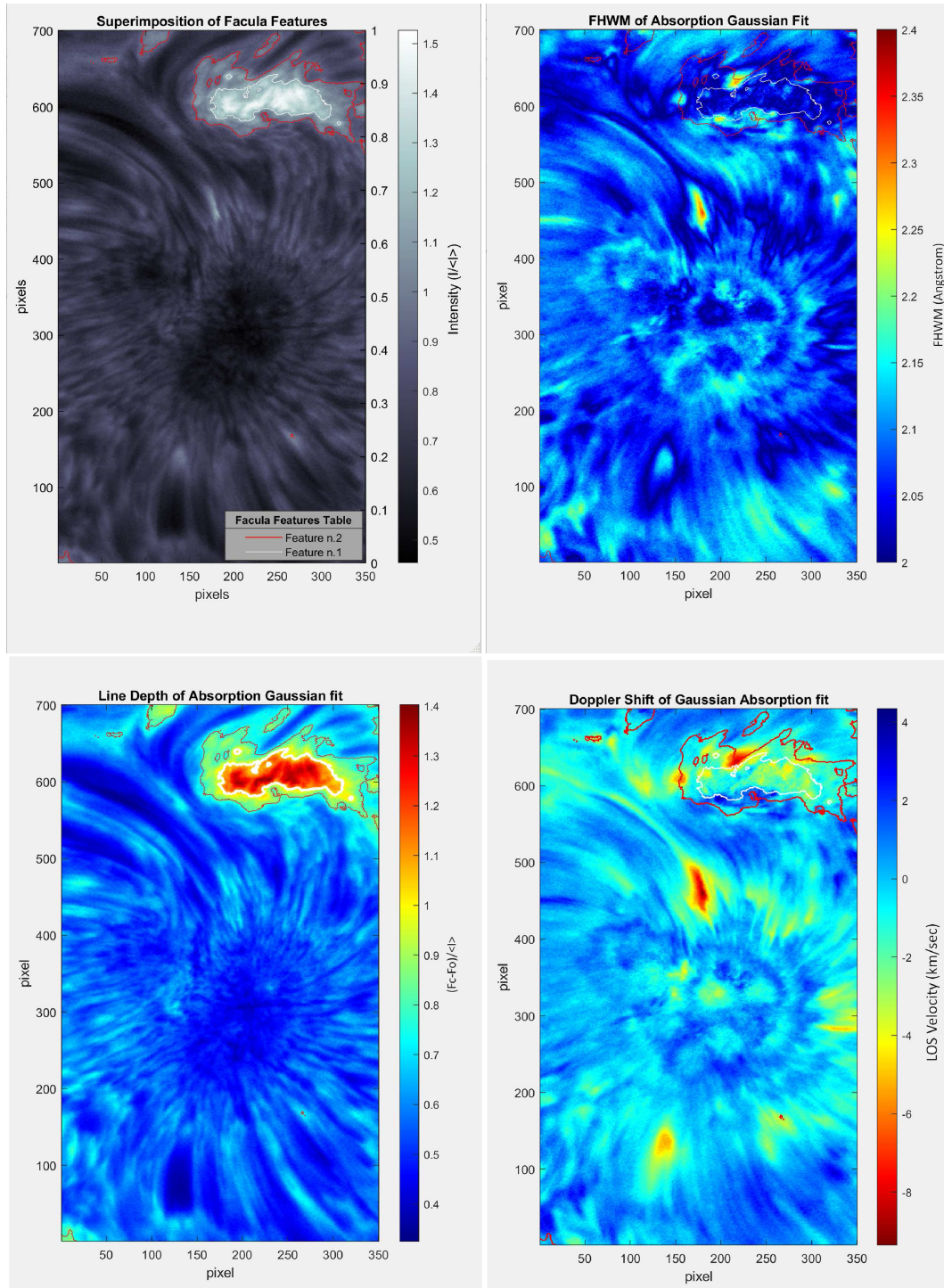
Figure 8. Superimposition of the contours of the features corresponding to the facula over the spectral image taken at 656.274 nm.

0.5, and 0.7, respectively. For the superpenumbra we can see also that features n. 10, 11, and 12 correspond to different level of LW (the top right-hand panel of Fig. 11), LD (bottom right-hand panel of Fig. 11) and LOS velocity (bottom left-hand panel of Fig. 11).

Features n. 13, 14, and 15 correspond to the fibrils observed in the surrounding region around the sunspot, as depicted in Fig. 12. In particular, feature n. 13 is characterized by an FWHM of the H α line of about 2.0 Å, an LD of about 0.4 and a positive LOS velocity (around 1 km s^{-1}). Instead, features n. 14 correspond to the portion

of the fibrils with LW $\sim$ 2.1 Å and LD $\sim$ 0.7. Positive values of the LOS velocity characterize features n. 15 (see the bottom right-hand panel of Fig. 12).

Finally, feature n. 16 can be considered to study the filament portions visible in the FOV. In fact, we clearly recognize the elongated shape of a filament in the top left-hand corner of the bottom right-hand panel of Fig. 13. Instead, feature n. 5 seems to identify the brighter regions corresponding to the footpoints of the filaments, the so-called umbral filaments, i.e. structures where the increase of the
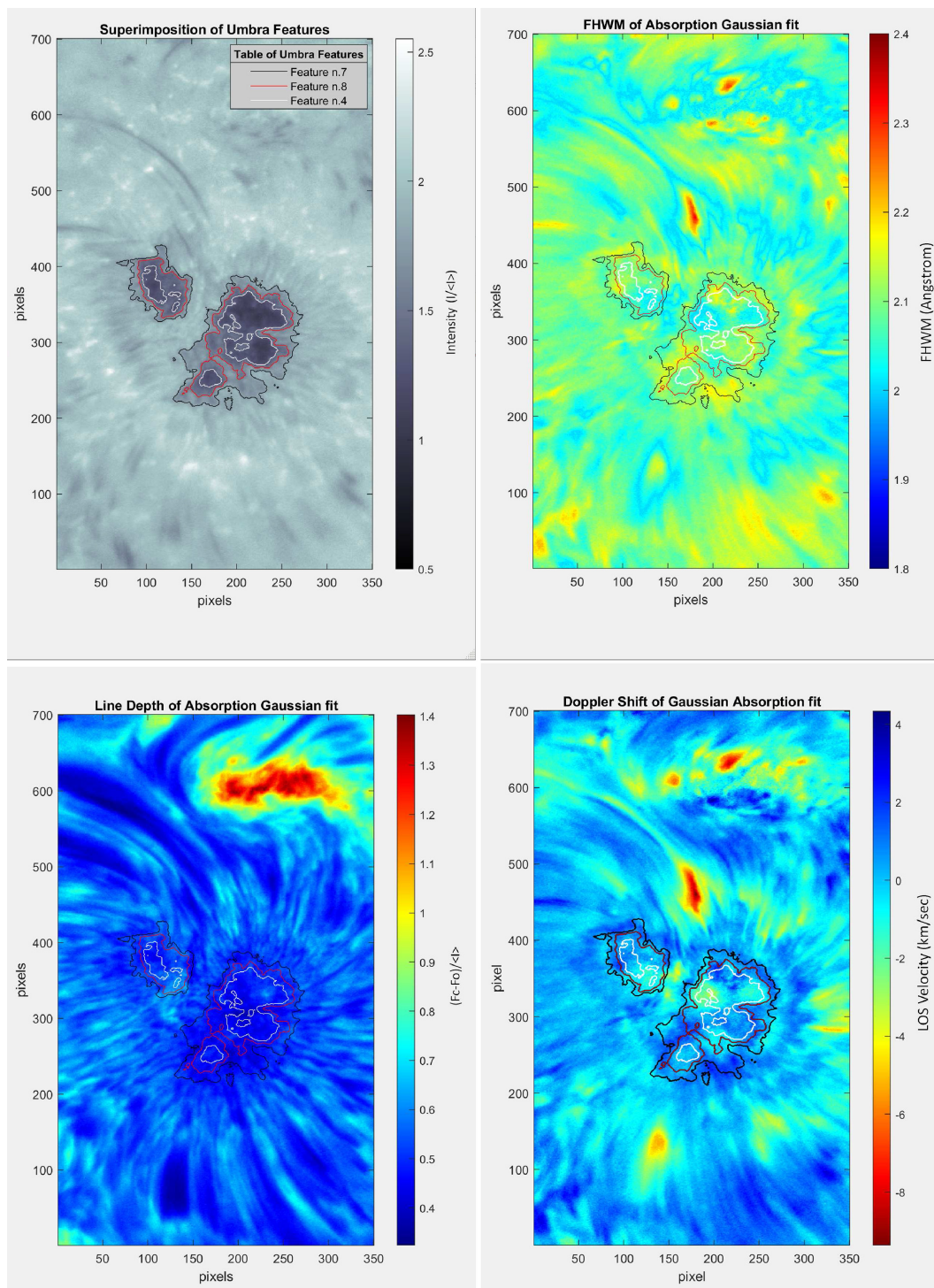
Figure 9. Superimposition of the contours of the features corresponding to the sunspot umbra over the spectral image taken at 656.134 nm.

plasma emission can be due to the accumulation of plasma coming from higher solar atmospheric levels into the photosphere and the low chromosphere (see Guglielmino, Romano & Zuccarello 2017; Guglielmino et al. 2019, for more details). This is confirmed by the LOS velocity map (bottom right-hand panel of Fig. 13), where we see that the black contours (feature n. 5) correspond to the stronger red shifted areas with velocity lower than $-8\,\mathrm{km\,s^{-1}}$, while the red contours (feature n. 16), indicating the upper parts of the filaments, are characterized by LOS velocities closer to $0\,\mathrm{km\,s^{-1}}$.

6 CONCLUSIONS

The use of machine learning methods becomes increasingly important in the world of big data, for the ability to process data in a massive and intelligent way by increasingly performing processors.

In this work, we applied for the first time the SOM method to high resolution images of the solar atmosphere. We exploited the main advantage of a bi-dimensional spectrometer based on Febry–Perot devices, such as IBIS. Using high spectral and spatial resolution
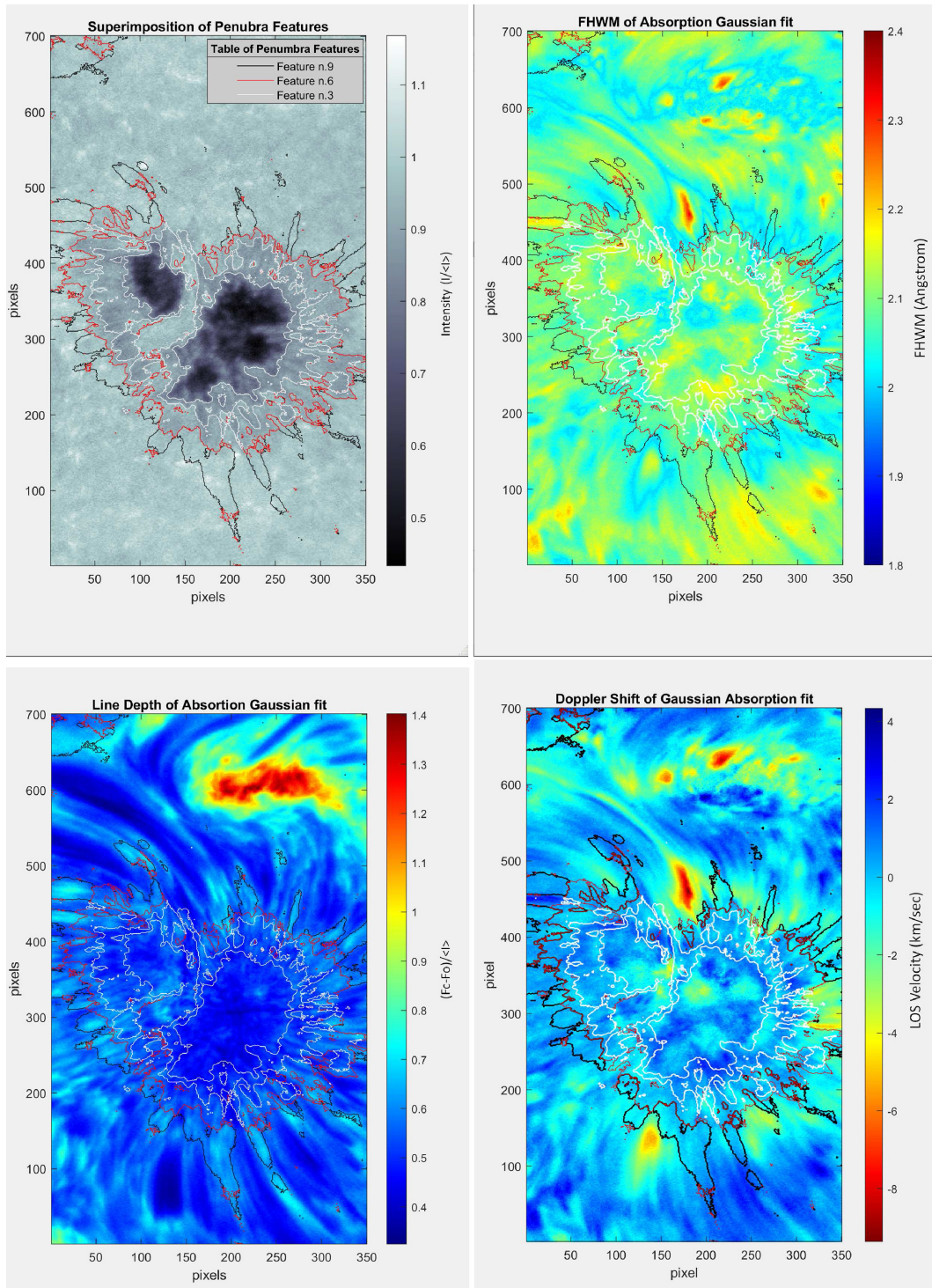
Figure 10. Superimposition of the contours of the features corresponding to the sunspot penumbra over the spectral image taken at 656.274 nm.

images taken at different wavelengths along the same spectral line (in our case the H α line), we were able to identify different structures of the solar atmosphere that depend on the plasma and magnetic field properties. The application of the SOM technique was perfectly suitable to segment the different images and to isolate among these the structures for a subsequent analysis. We remark that this method seems to obtain also a good correspondence between the features and the physical properties of the identified structures. In fact, we found

some differences in terms of physical or spectral properties (e.g. line core depth, line width, and Doppler shift) exhibited by features segmenting different portions of the same solar structure.

The advantages of the algorithm described in this paper are certainly manifold. First of all the developed algorithm is very efficient: using our data set, formed by 17 images of 350×700 pixels, its computation requires no pre-elaboration of data, took less than one minute for a SOM calculation with 200 epochs, using an
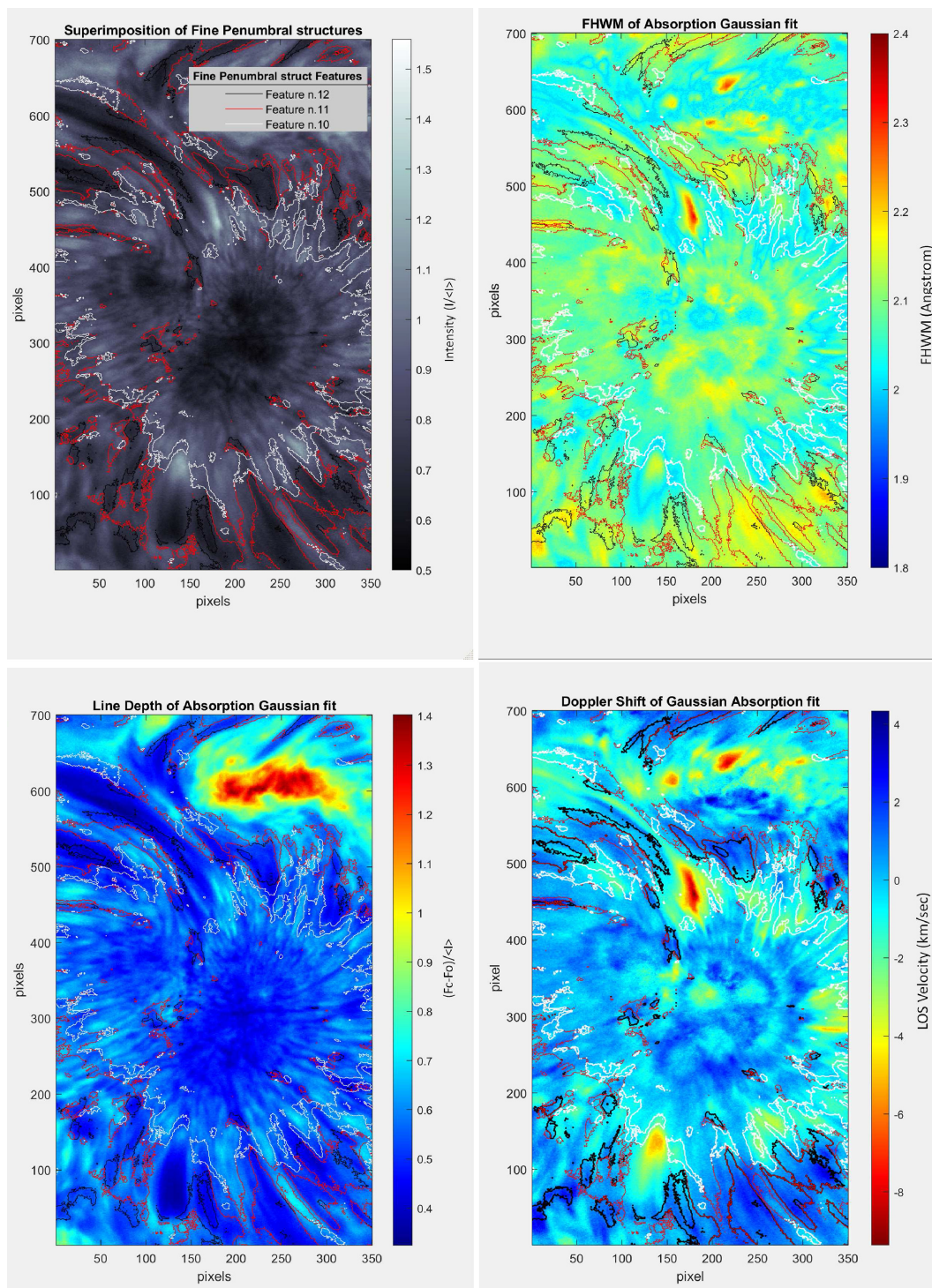
Figure 11. Superimposition of the contours of the features corresponding to the superpenumbra over the spectral image taken at 656.274 nm.

octa-core i9 processor with 2.3 GHz clock. Moreover the algorithm can be parallelized and can be exported on heterogeneous calculation platforms having greater computational efficiency (e.g. GPU, FPGA) than the platform used to find the prototype results achieved. In addition, this algorithm can be implemented on front-end devices that allow a complete automatic analysis of the data collected and its processing in an almost real-time regime (Hughes et al. 2019), so monitoring of the regions of interest could be automatic

and independent of the operator. In comparison to other methods reported in literature, we note that the weakness of this method concerns the different position of a feature into the SOM lattice, mainly after the elaboration of subsequent data sets. This issue has some consequences when the user needs to compare or elaborate features in different frames in order to perform their tracking or to measure the evolution of their physical properties. For this reason, other segmentation algorithms, like Yet Another Feature Tracking
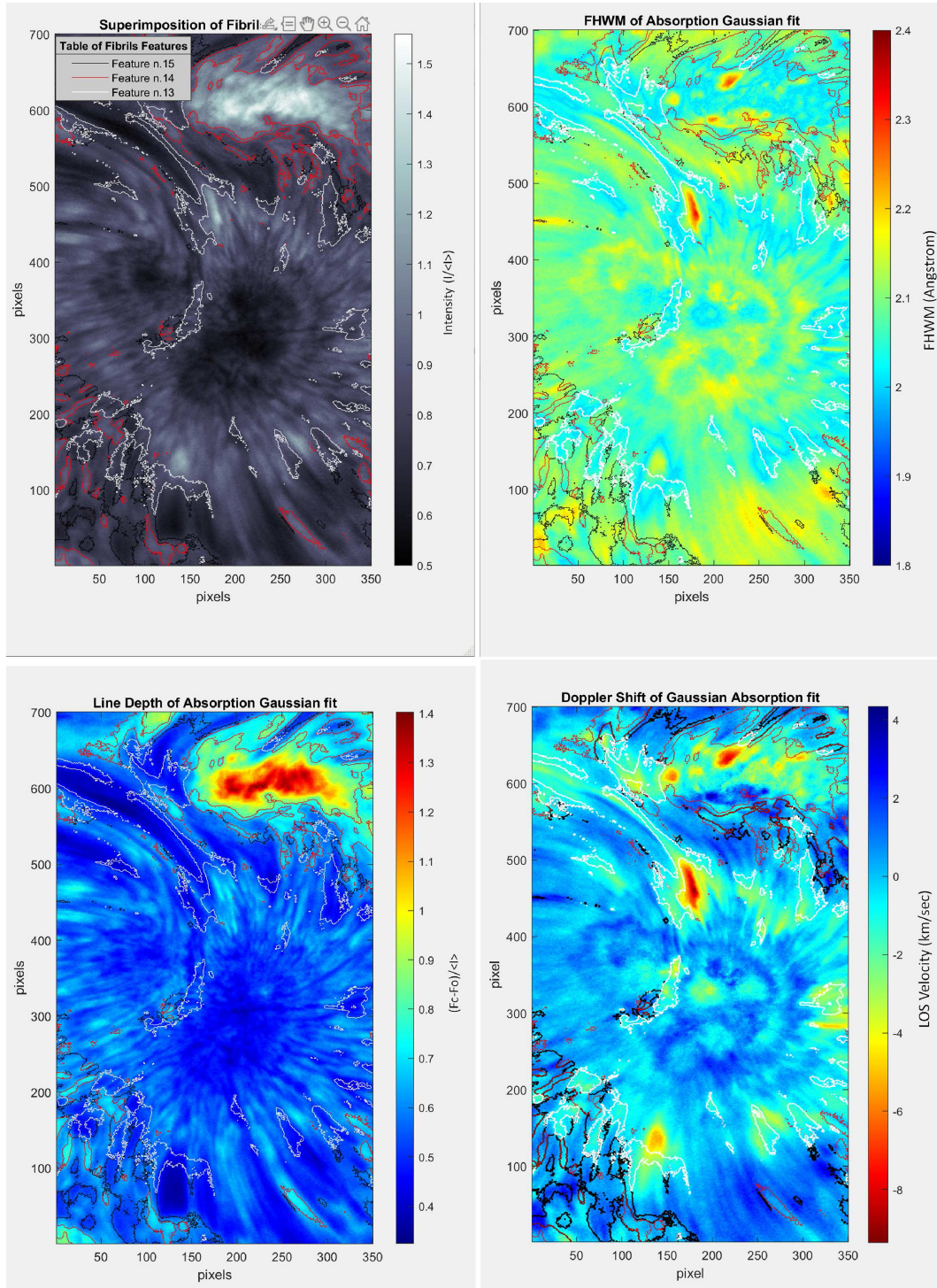
Figure 12. Superimposition of the contours of the features corresponding to the fibrils over the spectral image taken at 656.274 nm.

Algorithm (YAFTA; Welsch et al. 2004), provided certainly better performances for this scope. However, research activity in order to identify some topics of the feature and run a next cascade supervised and fast network is in progress.

Moreover we believe that the segmentation carried out is very accurate and is based on the determination of possible clusters emerging from the data itself. Unlike other methods (Zharkova et al. 2005; Aschwanden 2010; Perez-Suarez 2013; De Visscher

et al. 2015; Schad 2017) the segmentation reveals all the possible structures in one run and it is verified on the basis of objective parameters that make the segmented regions complementary and physically well connoted, for this reason it is known as semantic segmentation.

This work represents a first step towards the application of the SOM technique to astrophysical data sets. Taking into account that other skeletonization methods are not able to identify so many
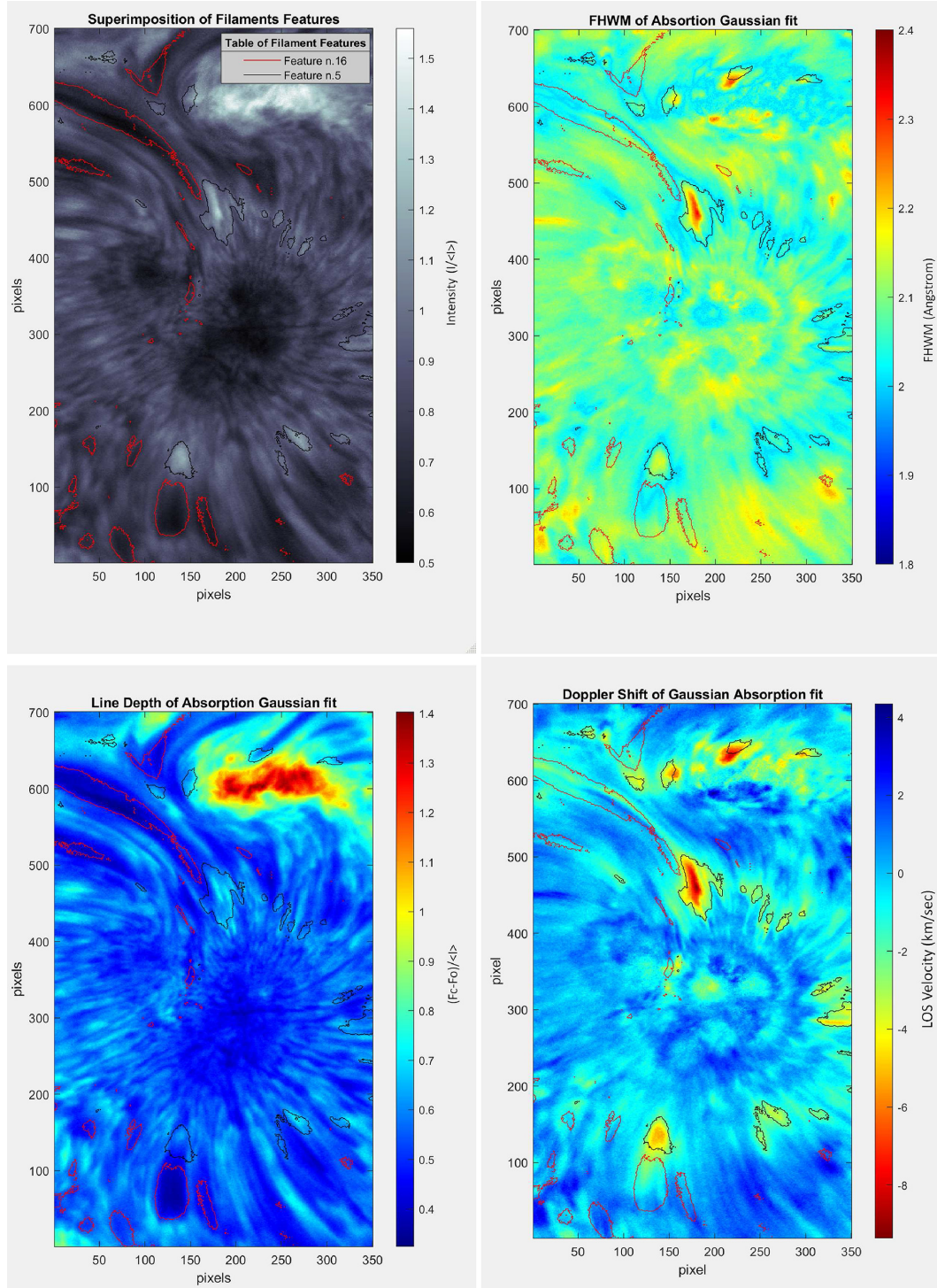
Figure 13. Superimposition of the contours of the features corresponding to the filaments over the spectral image taken at 656.274 nm.

structures at the same time and that such a result could not be achieved by a simple selection of several thresholds, we think that the SOM technique may find very useful applications to the analysis of this kind of data in the context of the space weather field. In fact, we also plan to apply the same technique to other solar data. We think that its application to full disk images taken at different wavelengths corresponding to different layers of the solar atmosphere, e.g. images taken by the Solar Dynamics Observatory

(SDO; Pesnell, Thompson & Chamberlin 2012), may reveal further potentialities for this kind of unsupervised approach (Verbeeck 2013).

Finally, the possibility to follow the evolution of those features identified by the SOM technique over several days of observation could be a useful tool to find new parameters which are able to forecast the occurrence of eruptive events in solar active regions.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The data set used in this Paper is available at http://ibis.oa-roma.inaf.it/IBISA/, in the archive containing data acquired with IBIS during some observing campaigns. The archive IBIS-A has been realized in the framework of the FP7 SOLARNET3 High-resolution Solar Physics Network that aims at integrating the major European infrastructures in the field of high-resolution solar physics, as a step towards the realization of the European Solar Telescope (Collados et al. 2017).

REFERENCES

Aggarwal C. C., 2018, Neural Networks and Deep Learning. Springer, Berlin
Aschwanden M. J., 2010, Sol. Phys., 262, 235
Barlow H. B., 1989, Unsupervised learn. Neural Comput., 1, 295
Barnes G. et al., 2016, ApJ, 829, 89
Benz A. O., 2017, Living Rev. Sol. Phys., 14, 2
Caballero C., Aranda M. C., 2013, Sol. Phys., 283, 691
Calinski T., Harabasz J., 1974, Commun. Stat., 3, 1
Cavallini F., 2006, Sol. Phys., 236, 415
Collados M. et al., 2010, AN, 331, 615
Davies D. L., Bouldin D. W., 1979, in Chien Y. T., Pavlidis T., eds, IEEE Transactions on Pattern Analysis, Machine Intelligence, Vol. PAMI-1, IEEE Computer Society, N.2, p. 224
De Visscher R. J., Delouille V., Dupont P., Deledalle C.-A., 2015, J. Space Weather Space Clim., 5, A34
Evershed J., 1909, The Observatory, 32, 291
Falco M., Costa P., Romano P., 2019, J. Space Weather Space Clim., 9, 22
Florios K. et al., 2018, Sol. Phys., 293, 42
Guglielmino S. L., Romano P., Zuccarello F., 2017, ApJ, 846, L16
Guglielmino S. L., Romano P., Ruiz Cubo B., Zuccarello F., Murabito M., 2019, ApJ, 880, 34
Hausen R., Robertson B. E., 2020, ApJS, 248, 20
Hughes M. J., Hsu V. W., Seaton D. B., Bain H. M., Darnel J. M., Krista L., 2019, J. Space Weather Space Clim., 9, A38
Kim T. et al., 2019, Nat. Astron., 3, 397
Kohonen T., 1997, Self-Organizing Maps, 2nd edn. Springer, Berlin
Kontogiannis I., Georgoulis M. K., Park S.-H., Guerra J. A., 2018, Sol. Phys., 293, 96
Korsós M. B., Yang S., Erdélyi R., 2019, J. Space Weather Space Clim., 9, A6
Leka K. D., Metcalf T. R., 2003, Sol. Phys., 212, 361
Löfdahl M. G., 2002, in Bones P. J., Fiddy M. A., Millane R. P., eds, SPIE Conf. Ser. Vol. 4792, Image Reconstruction from Incomplete Data. SPIE, Bellingham, p. 146
Perez-Suarez D., 2013, in Qahwaji R., Green R., Hines E. L., eds, Applied Signal, Image Processing: Multidisciplinary Advancements, 1st edn. IGI Global, p. 207
Pesnell W. D., Thompson B. J., Chamberlin P. C., 2012, Sol. Phys., 275, 3
Romano P. et al., 2015, A&A, 582, A55
Romano P. et al., 2017, ApJ, 837, 173
Romano P., Elmhamdi A., Falco M., Costa P., Kordi A. S., Al-Trabulsy H. A., Al-Shammari R. M., 2018, ApJ, 852, L10
Romano P., Elmhamdi A., Kordi A., 2019, Sol. Phys., 294, 1
Romano P. et al., 2020, ApJ, 899, 129
Rousseeuw P. J., 1987, J. Comput. Appl. Math., 20, 53
Schad T., 2017, Sol. Phys., 292, 24
Shelhamer E., Long J., Darrell T., 2017, IEEE Trans. on Pattern Analysis and Machine Intelligence, 39, 4
Solanki S. K., 2003, A&AR, 11, 153
Verbeeck C., 2013, Sol. Phys., 283, 67
Welsch B. T. et al., 2004, ApJ, 610, 1148
Zeljko I., Connolly A. J., VanderPlas J. T., Gray A., 2014, Princeton Series in Modern Observational Astronomy, Statistics, Data Mining and Machine Learning in Astronomy. Princeton Univ. Press, Princeton, NJ, p. 270
Zharkova V., Ipson S. S., BenkhalilS A., Zharkov S., 2005, Artif. Intell. Rev., 23, 209

This paper has been typeset from a TeX/LaTeX file prepared by the author.