IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 4, 2018, accepted November 13, 2018, date of publication November 16, 2018,
date of current version December 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2881832

Applying Deep Learning for Improving Image Classification in Nuclear Fusion Devices

GONZALO FARIAS[1], ERNESTO FABREGAS[2], SEBASTIÁN DORMIDO-CANTO[2], JESÚS VEGA[3], SEBASTIÁN VERGARA[1], SEBASTIÁN DORMIDO BENCOMO[2], IGNACIO PASTOR[3], AND ALVARO OLMEDO[2]

[1]Escuela de Ingeniería Eléctrica, Pontificia Universidad Católica de Valparaíso, Avenida Brasil 2147, Valparaíso 2362804, Chile
[2]Departamento de Informática y Automática, Universidad Nacional de Educación a Distancia, 28040 Madrid, Spain
[3]Laboratorio Nacional de Fusión, CIEMAT, 28040 Madrid, Spain

Corresponding author: Ernesto Fabregas (efabregas@bec.uned.es)

This work was supported in part by the Chilean Ministry of Education under Project FONDECYT 1161584 and in part by the Spanish Ministry of Economy and Competitiveness under Projects ENE2015-64914-C3-1-R and ENE2015-64914-C3-2-R.

ABSTRACT Deep learning has become one of the most promising approaches in recent years. One of the main applications of deep learning is the automatic feature extraction with auto-encoders (AEs). Feature extraction, one of the most important stages in machine learning, that can reduce drastically the dimensionality of the problem, making easier any subsequent process such as classification. The main contribution of this research is to evaluate the use of AEs for automatic feature extraction in massive thermonuclear fusion databases. In order to show the performance of AEs in a practical way, the problem of image classification of the TJ-II Thomson Scattering diagnostic has been selected. The classification has been performed by the algorithm of support vector machines and conformal predictors. The results show that the use of AEs produces the predictions faster, with more reliable models, and with higher success rates in comparison to the performance without using the deep learning approach.

INDEX TERMS Images classification, auto-encoder, future extraction, nuclear fusion.

I. INTRODUCTION

Deep learning is an emerging approach of machine learning that helps to create models composed of multiple processing layers with different levels of abstraction [1]. In recent years, deep learning has been used for many processing, recognition and classification applications, such as: images [2], video [3], audio (voice) [4], intelligent driving assistance [5], multi-sensors integration [6], medical disease analysis [7], etc. Some of the most powerful artificial intelligence methods in recent years involve the use of stacking sparse auto-encoders (AE) [8].

An auto-encoder is an artificial neural network (ANN) used for unsupervised learning of efficient codings [9]. The AE were introduced in the late 80's by Rumelhart *et al.* [11] and Baldi and Hornik [12] as a technique for dimensionality reduction, where the output of the encoder is the reduced representation and the decoder is tuned to reconstruct the initial input from the encoder's representation through the minimization of a cost function like in Rifai *et al.* [13].

On the other hand, feature extraction has become an increasingly critical stage in machine learning, due to the data growing in high dimensions. This increases the difficulty of proving the results due to the sparsity of meaningful data. Feature extraction can drastically reduce the dimensionality of the problem, making the process of classification easier.

For example, Yuan *et al.* [14] and Wen *et al.* [15] proposed feature extraction from electronic sensors to improve the behavior of the systems. However this process is usually performed manually and often requires multiple tests with a high level of computational resources, which consumes time and effort from the designer. This also implies that the designer needs to have a high level of expertise in the analyzed system. For that reason, we are proposing the use of auto-encoders to automate (i.e. without human manipulation) the feature extraction process.

In this context, Vincent *et al.* [16] and Muhammad *et al.* [17] presented the use of deep convolutional networks for remote sensing data analysis. In Masci *et al.* [18] and Romero *et al.* [19] the authors said that the convolutional neural network (CNN) require large quantities of labeled data (which is not a problem in nuclear fusion context). They proposed the stacked convolutional AE, which can map

images without any label information. In Liang *et al.* [20] the authors presented a review of the use of deep learning for feature extraction applied to text mining.

In the context of thermonuclear fusion of plasmas for energy generation, the process is nonlinear and complex and it requires a high degree of reliability. This process generates increasing amounts of data (signals, images, etc) that are stored in big databases. This data needs to be analyzed with different purposes, for example: patterns recognition and failures detection (disruption prediction) like in Makili *et al.* [21] and Cannas *et al.* [22]. In this scenario, the feature extraction has a crucial role to reduce the amount of data to analyze, like in the aforementioned examples. That is why it is important to automate this stage.

In the literature it can be found some works related to these issues. For example, some years ago Jionghua and Jianjun [23] presented a methodology for diagnostic system of fault detection using waveform signals and wavelet transform as feature extraction method (where the parameters are tuned manually). In Vega *et al.* [24] and Makili *et al.* [21] the authors presented two versions of an automatic classification system based on Support Vector Machines (SVM) and Thomson Scattering (TS) diagnostic. In Rattá *et al.* [25] the authors developed an algorithm for real time disruption prediction, where the feature extraction was based on Genetic Algorithms. Other approach for images classification using deep learning is presented in Matos *et al.* [26], where the authors introduced a pre-trained CNN in plasma diagnosis. More recently, in Farias *et al.* [27] we presented the advances in deep learning approach applied to two problems of nuclear fusion classification. In addition, in Farias *et al.* [28] the authors proposed the use of Adaboost algorithm to classify TS images of the TJ-II fusion device.

This article addresses the assessment of including AE for an specific case of automatic feature extraction in the massive thermonuclear fusion databases. In order to show the performance of AE in a practical way, the problem of the classification of a set of 981 images of the TJ-II TS diagnostic has been selected. TS is a technique used for the measurement of the electron temperature T_e and density n_e in very hot fusion plasmas like in Giudicotti *et al.* [29].

Following the cross-validation technique to evaluate predictive models, the set of images is partitioned into two subsets (training and test sets). Similar to other pattern recognition problems like in Dormido-Canto *et al.* [30] and Makili *et al.* [31], in this case there are two main stages: i) the pre-processing of the plasma signals using Wavelet Transform (WT) to reduce the dimensionality without loss of relevant information; and ii) the application of a classification algorithm to get a model. Thus, the AE is just added between both stages to extract features from the pre-processed images, which means that the AE stage can be applied to an existing classifier to extract features of the input images of the classifier in an automatic way. This is the main difference of this approach in respect to the aforementioned works that use deep learning (ANN, CNN, etc.). The classification stage has been performed by the algorithm of SVM like in Hearst *et al.* [32] in combination with Conformal Predictors like in Vovk *et al.* [33] and Saunders *et al.* [34] and Vovk *et al.* [35]. But, it should be clear that the selection of a different method for this stage should produce similar results.

The main contribution of this research is to evaluate the performance of the application of auto-encoders for automatic feature extraction in massive thermonuclear fusion databases. In this sense, the research rises three relevant questions of the use of AE for the image classification in the context of nuclear fusion: i) are the models with AE more accurate? ii) are the models with AE better fitted for new images (not used for training)?, and iii) are the predictions computed faster when AE is used?.

Thus, the motivation of this work raises the following hypotheses, that are evaluated and discussed in detail in next sections:

H1: the use of AE provides more reliable models and with better success rates.

H2: the use of AE reduces potential over-fitting.

H3: the use of AE decreases the prediction times.

The paper is structured as follows. Section II describes the Methodology followed in the image classification approach. The Methodology includes: 1) the database from the Thomson Scattering diagnostic of the TJ-II fusion device; 2) image classification with the particular attributes of the AE; and 3) SVM with Conformal Predictors. Section III discusses the main experimental results based on the Hypotheses raised at the beginning of this research. In the last section, the main conclusions and future works are discussed.

II. METHODOLOGY OF IMAGE CLASSIFICATION: METHODS AND MATERIAL

This section presents the methodology to carry out an automatic feature extraction method for pattern recognition problems. In order to evaluate the performance of the deep learning approach in nuclear fusion, the classification of five types of TJ-II Thomson Scattering images is carried out.

A. TJ-II THOMSON SCATTERING DIAGNOSTIC

The TJ-II is a highly flexible medium-size nuclear fusion experimental and stellarator device constructed at Spanish National Fusion Laboratory of the CIEMAT in Madrid, Spain, between 1991 and 1997 (see Fig. 1).

In TJ-II, the magnetic trap is obtained by means of various sets of coils that completely determine the magnetic surfaces before plasma initiation. The toroidal field is created by 32 coils. The three-dimensional twist of the central axis of the configuration is generated by means of two central coils: one circular and one helical. The horizontal position of the plasma is controlled by the vertical field coils. The combined action of these magnetic fields generate bean-shaped magnetic surfaces that guide the particles of the plasma so that they do not collide with the vacuum vessel wall. The plasmas in TJ-II are produced and heated with ECRH

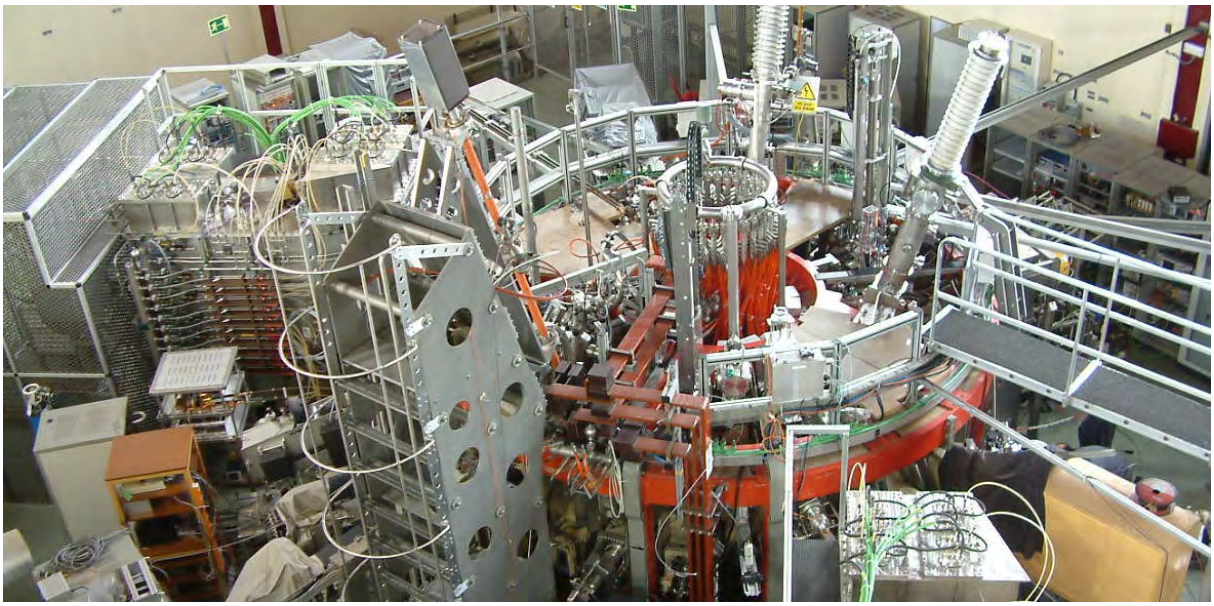

FIGURE 1. TJ-II experimental fusion device.

TABLE 1. Five type of TJ-II TS images and their corresponding pre-processed versions.

TS Image	Description
BKG	CCD camera background
COF	Cut-off density during electron cyclotron resonant heating
ECH	Electron cyclotron resonant heating
NBI	Neutral beam injection heating
STR	Stray light

(2 gyrotrons, 300 kW each, 53.2 GHz, 2nd harmonic, X-mode polarization) and NBI (300 kW) [36], [37]. In TJ-II a typical discharge last between 150-250 milliseconds, and depending on the sampling rate, the number of samples could be in the range of 4000-16000 per shot. These discharges produces big databases with lot of information that has to be analyzed.

One of the analysis that is carried out to study the plasma behavior is the well known Thomson Scattering diagnostic. This diagnostic in plasma consists in the re-emission of incident radiation (from very powerful lasers) by free electrons. Electron velocity distribution generates a spectral broadening of the scattered light (by Doppler effect) related to the electronic temperature. The total number of scattered photons is proportional to the electronic density. That is why this is the most used technique to determine the temperature and density of the plasma electrons.

TS diagnostic acquires five types of images (spectra of laser light scattered by plasma): CCD camera background (BKG), the cut-off density during electron cyclotron resonant heating (COF), images during electron cyclotron resonant heating (ECH), during neutral beam injection (NBI) and measurement of stray light without plasma or in a collapsed discharge (STR). Table 1 describes the five classes considered.

Next subsection explain the images classification with AE.

B. IMAGE CLASSIFICATION WITH AUTO-ENCODERS

Before to train an auto-encoder to extract features automatically, we first apply pre-processing to the TS images. To this end, we have selected Wavelet Transform in order to reduce the dimensionality of the problem. WT offers an efficient alternative to data processing and provides many advantages: (1) data compression, (2) computing efficiency, and (3) simultaneous time and frequency representation.

According to previous works and the experience of Makili *et al.* [21] and Farias *et al.* [38] the wavelet Haar has been used due to its simplicity (specially suitable for real-time operation) and low computational time. The parameters of the WT are defined as follows: *Approximation coefficient*; *Wavelet mother: Haar*; *Level of decomposition: 4*. As result of the application of WT, the image dimensions were reduced from the original 576x385 (221760) pixels to only 36x25 (900) pixels without loss of critical information, which makes easier and faster the subsequent steps.

Finally, in order to speed-up the experimental results, we have discarded upper and lower rows of the WT approximation coefficient to obtain square images of 25x25 (625) pixels. The results of this pre-processing stage for the five types of TS images can be observed in Fig. 2. Note that the use of another wavelet family (e.g. *db*, *sym*, *bior*, etc) could produce different results, specially from the dimensional reduction point of view.

AE is an ANN with an unsupervised learning that applies back-propagation algorithm (see Fig. 3). The AE is trained to fit an identity function, i.e., the AE is required to get targets or output values to be equal to the inputs. This seems to be very trivial and not interesting at all, but setting some constraints on the AE (e.g. number of neurons at the hidden layer, and sparseness) we could find out auto-

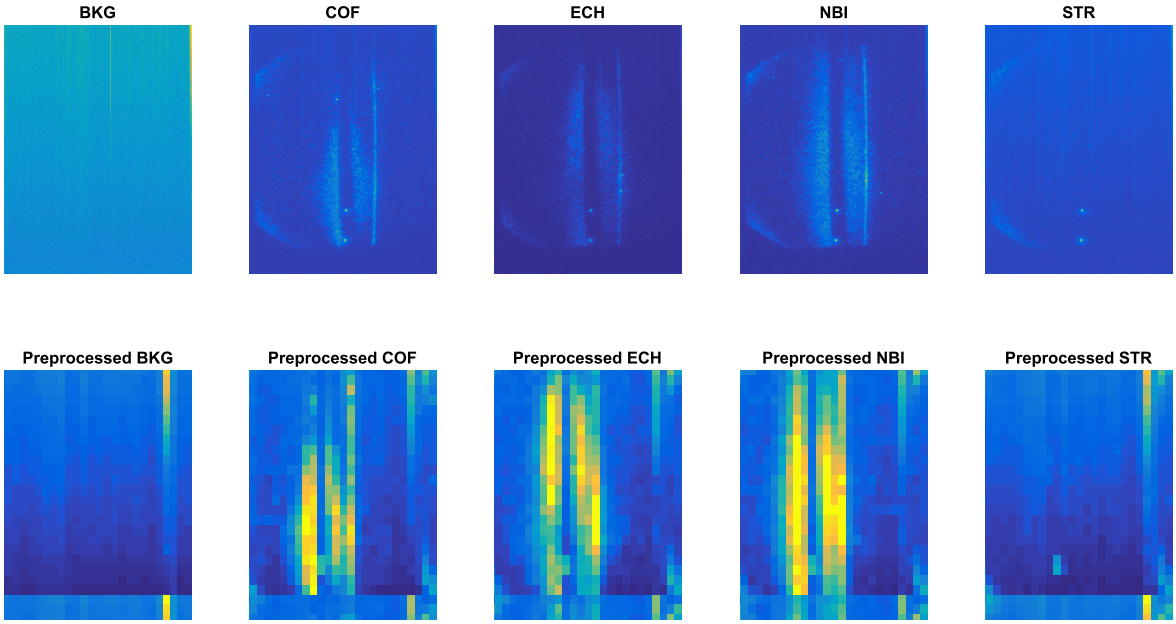

FIGURE 2. Five type of TJ-II TS images and their corresponding pre-processed versions.

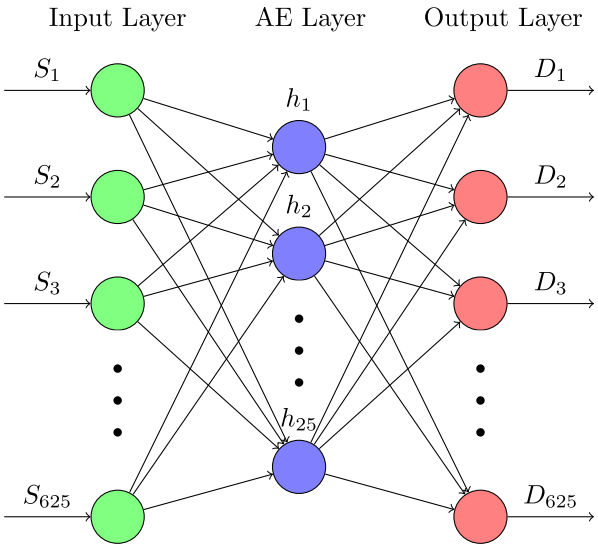

FIGURE 3. Auto-encoder.

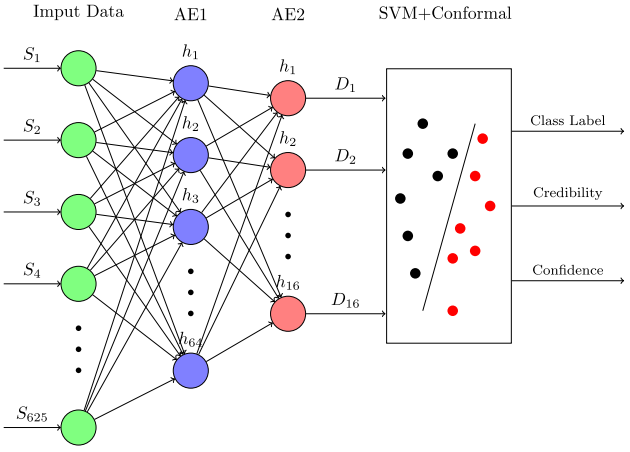

FIGURE 4. SVM classifier with two AE.

matically some useful features for a pattern classification problem.

For example, a small number of units or neurons in the AE layer could code input data in a simpler and reduced feature space. The sparseness is another constraint that could be very useful. A sparse AE is built to keep hidden neurons inactive most of the time. An inactive unit outputs values close to 0, on the contrary, an active unit outputs values close to 1. Thus, it should be expected that ideally only one unit of the hidden layer is active at time, when considering sparseness in the AE.

One advantage of AEs is the option to get stacked AE (deep networks) in a greedy layer-wise way, which means that each AE can be trained separately and then put them all together. This way avoids some well-known issues of multi-layer networks and back-propagation algorithm such as the vanish gradient problem [8]. Basically, such issue appears when trying to propagate back the error from the output layer to the input layer. Fig. 4 shows a simple deep network with two AEs.

An automatic image classification system based on SVM has been in operation for years in the TJ-II Thomson Scattering diagnostic, thus, in order to evaluate the automatic extraction of features by using AE, we have built a simple sparse AE. Based on our previous work [27], the ANN is configured

with one input layer of 625 units (each unit represents a pixel of the input TS image), one hidden layer (the AE layer), and one output layer of 625 units. In this work, we have modified the number of units of the hidden layer in order to test the performance of the classifier with a different number of features. We have also test the classifier with two auto-encoders (see Fig. 4). Note that the output of the first auto-encoder (AE1), which process the input image, feeds a second auto-encoder (AE2), which in turn, provides the input features for the SVM classifier. Thus, the original feature space could be reduced enormously. In this case, the input space starts in 625 features and it is reduced up to only 16 attributes. Finally note that a new extra layer, given by a conformal predictor, has been also added to the image classifier in order to evaluate better the performance of the entire system when the AE is used.

Next sections describe in detail how the AE, the SVM classifier and the Conformal predictor are combined in order to get a classification for each input image of the TS diagnostic.

C. SUPPORT VECTOR MACHINES

SVM is a very effective method for general purpose pattern recognition [39]. It classifies a set of elements into two different classes. To do it, the method builds a model that can learn from a marked training set of data. After that, the model is capable to assigns a category for new unknown input data. The model also maps the examples as points in space, divided by a clear gap that separates the categories. New inputs are placed in the same space and its category depends of the side of the gap where they fall [40].

If you can find two parallel hyper-planes that separate both classes, and these hyper-planes are far enough, it means that the training data set is linearly separable. The region marked by theses hyper-planes is called margin. This margin is maximum in the hyper-plane that represents the half distance between them. These hyper-planes can be described by the equation 1:

$$\begin{cases} \vec{w} \cdot \vec{x} - b = 1 \\ \vec{w} \cdot \vec{x} - b = -1 \end{cases} \tag{1}$$

All the points inside the hyper-plane including the border line represented by equation 1, is of one class with label 1. All the points below the hyper-plane including the border line represented by second expression of equation 1, belong to class -1.

The distance that separates the two hyper-planes is $\frac{2}{\|\vec{w}\|}$. This distance is maximum when $\|\vec{w}\|$ is minimum [41]. Fig. 5 shows a representation of the maximization of the distance between the hyper-planes. In this example Class 1 are the black points and Class 2 the white points. The red color represents the points that are into the limit of the hyper-plane of each Class.

If these hyper-planes can not be found, because straight lines that separate the data do not exist. It means that the training data in not linearly separable. In this case, SVM can map the data into high-dimensional feature space and it can build an optimal separating hyper-plane. The nonlinear

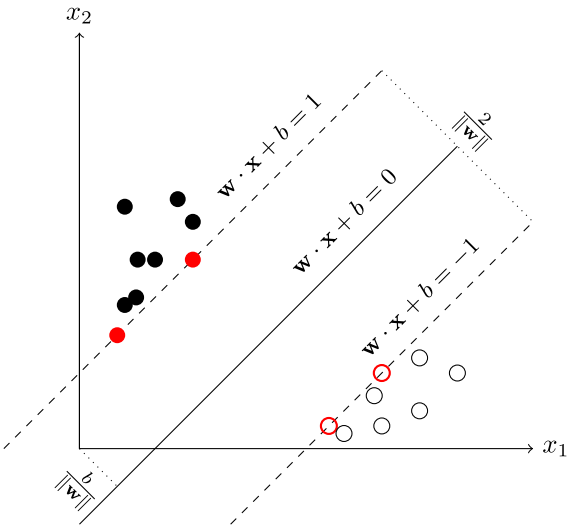

FIGURE 5. Maximum margin hyper-plane and margin for an SVM trained with samples from two classes.

mapping is performed by a kernel function $K(x, x')$ with the original constrains (see Fig. 6) [30], [42]. In this example, a not linearly separable data set is converted into a linearly separable data set through the kernel function $K(x, x')$.

CONFORMAL PREDICTORS

Conformal Predictors are among the most accurate machine learning methods and, unlike other state-of-the art methods, they do not provide only a prediction. In addition to the class label, they provide a level of reliability for the prediction like in Vovk *et al.* [33] and Saunders *et al.* [34]. Conformal Predictors are underlying algorithms that can complement the predictions with two extra values: confidence and credibility. These values are in the range [0-1] and they are used to estimate the goodness of the prediction of each new sample. A high level of credibility means that the predicted label is very likely. A high confidence means that all labels except the predicted one are unlikely. Thus, If both values are (very) close to 1, the prediction reliability is (very) high. Note that any traditional classification method can be used as an underlying algorithm. It is particularly important to be able to apply conformal predictions to problems dealing with high-dimensional data sets with large number of samples.

Conformal algorithm uses past experience of examples to determine precise levels of confidence in new predictions. For each new sample it needs to measure how different is the new one from the old examples using the nonconformity measure like in Vovk *et al.* [35].

Conformal predictions use an alternative framework that is transductive (TCP) and online: each new prediction is based on all the previous samples instead of use a same rule from a fixed bag of samples. On the other hand, inductive conformal prediction (ICP) allows to obtain a general rule that is called decision rule.

Fig. 7 shows six samples that belong to three different classes: squares, triangles and circles. The feature vectors

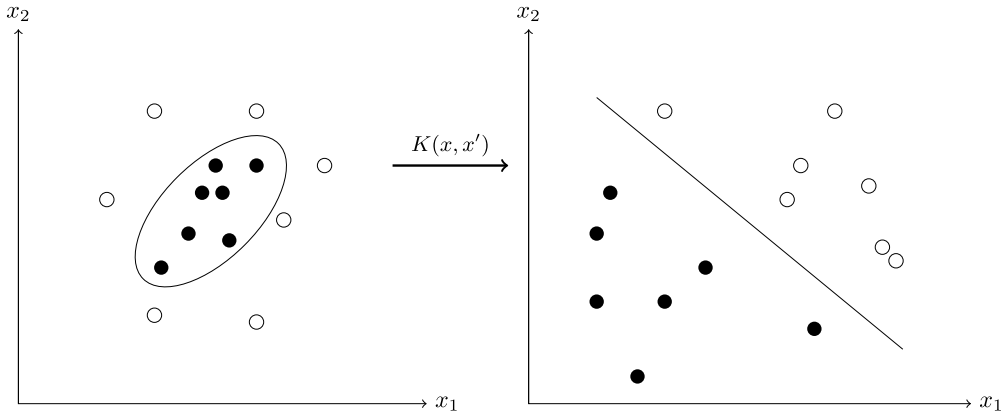

FIGURE 6. Training data into a higher-dimensional feature space via kernel K.

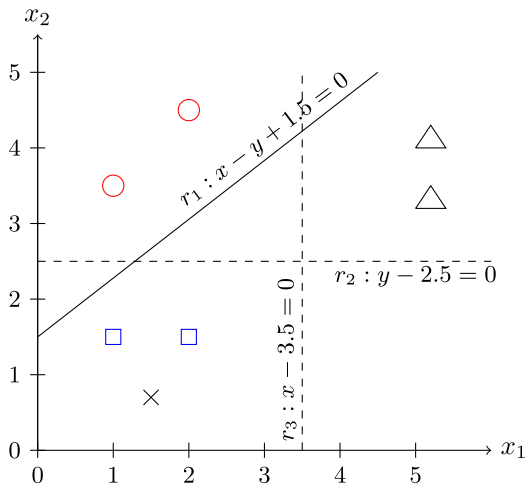

FIGURE 7. Example of transductive on-line prediction.

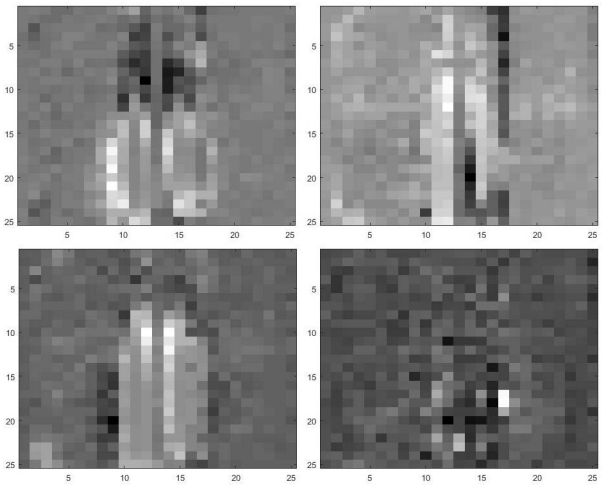

FIGURE 8. Input images that activate each one of the 4 units of the auto-encoder trained with the TS images.

are the respective Cartesian coordinates. In this case, a new sample (cross) has to be classified into one of the previous classes and the prediction has to be qualified with confidence and credibility.

The first step consists of choosing an underlying classification system to build the conformal prediction. As it is known, SVM has been designed to deal with classification problems of two classes fundamentally. To use it for multiple class it is necessary to extend its performance based in two approaches: 1) ''one vs. the rest'', and 2) ''one vs. one''. In the first procedure each classifier is trained to separate one class from the rest using M classes and M auxiliary classifiers. While the second approach uses one classifier for each pair of classes, which implies $M(M-1)/2$ different classifiers.

Using the ''one versus the rest'' approach in the example of Fig. 7, three SVM classifiers are needed: (squares) vs (circles, triangles), (circles) vs (squares, triangles) and (triangles) vs (squares, circles). In each one of the classifiers, the individual class is taken as class -1 and the rest samples are taken as class $+1$. The cross object is tested in each one of the three classifiers and in all cases it is assumed to belong to the class -1 (i.e. it is considered to be in the individual class). For a transductive on-line prediction, a nonconformity measure has to be defined to predict with confidence and credibility. The nonconformity measure has to provide how strange the cross object is in respect to the rest of examples. In SVM the prediction is more reliable when the distance to the decision function is greater, which means that this distance can be used as measure of strangeness of the sample.

An interesting result is to understand what the AE is trying to learn in the training phase. This example has been taken from our previous work Farias *et al.* [27]. In this case, the input data (e.g. TS image) that activates one of the features is used. So, the feature that is being sought for each unit of the sparse AE can be observed. For instance, Fig. 8 shows the TS images that an AE with four features has learned to look for. Note that these images show particular features of the TS image classes. In particular, features from NBI/ECH

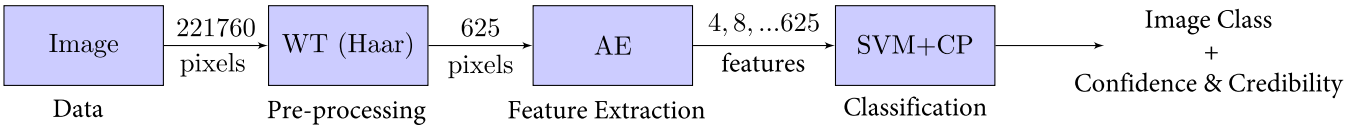

FIGURE 9. Block diagram of the system.

and COF images are clearly observed in the two lower input images.

III. EXPERIMENTAL RESULTS

Figure 9 shows the block diagram of the system architecture. The first block (from left to right) represents the input image to process (221760 pixels). The second block represents the pre-processing stage, where the WT is used to reduce the dimension of the image up to 625 pixels. The third block represents the feature extraction step, where the AE receives the reduced image to extract the features in an automatic way. Note that the number of features varies from 4 to 625 to perform the experimental analysis. The last block shows the classification stage, where the SVM and Conformal Predictor methods are applied to obtain the class of the image and the Confidence and Credibility of the prediction.

As it was mentioned before, the problem of classification is applied to a set of 981 images of the TJ-II Thomson Scattering diagnostic. Following the cross-validation technique to evaluate predictive models, the set of images is divided randomly into two subsets: 1) a training set of 588 images (60%) to train the model; and 2) a test set of 393 images (40%) to assess the developed system. Each experiment was repeated 100 times by changing the training and validation subsets. The results show the average of these 100 experiments. In this section the obtained results are discussed based on the hypotheses raised at the beginning of this work.

A. H1: THE USE OF AE PROVIDES MORE RELIABLE MODELS AND WITH A BETTER SUCCESS RATE.

Table 2 shows the results of the experiment of feature extraction from TS images database with a simple AE. To assess the whole classifier, the number of features has been modified. The parameters of the SVM are the following: *Kernel Radial basis function*, $\sigma = 8$, $C = 60000$. These parameters for the SVM classifier are constant during all the experimental evaluation. The performance was separated in two criteria: i) the success rate: success cases over total cases; and ii) the support vector rate (SVR): support vectors over total training samples.

The first column represents the input features. The second column represents the extracted features. The third column represents the success rate (%) which represents the average of the success classification of all classes. Regarding to the support vectors rate (%), it represents the percentage of training data considered as support vector by the SVM model, i.e., this value is associated with the complexity of the model.

TABLE 2. Results of one AE classifier for TS images.

Inputs Features	Extracted Features	Success rate (%)	Support vectors rate (%)
625	**625**	**98**	**48**
625	**256**	**99**	**15**
625	128	98	11
625	64	98	9
625	**32**	**97**	**7**
625	16	98	7
625	8	92	13
625	4	78	32

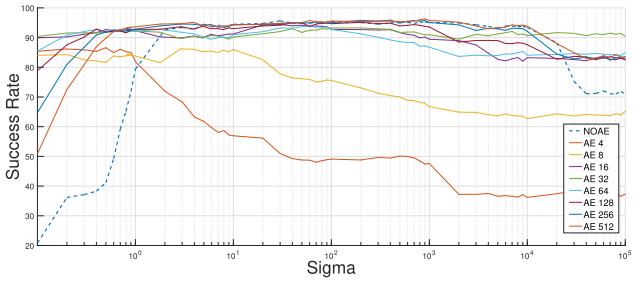

FIGURE 10. Success rate for a simple AE.

Therefore, the lower is the support vectors rate the lower risk of over fitting exists.

Note that the first row of the Table 2 represents the case without feature extraction (i.e., no AE is used in that case). Note also that the AE with 32 features reaches the lowest support vector rate (7%) and with a high success rate (97%) for the TS image classifier. In this latter case, note that feature space is reduced to less than 0.02% (32/221760) from the original one.

Fig. 10 shows the success rate with a simple AE as a feature extraction method. The number of features has been modified to assess the whole classifier. Thus, for instance, the case with 64 features corresponds to AE 64. The kernel of the SVM classifier is *Radial Basis Function*. The regularization parameter (C) is set in 60000. Finally, the parameter σ of SVM has been evaluated for a wide range (from 10^{-1} to 10^{6}). The evaluation also includes the case without using AE (NO AE − blue dashed line). Other tests was carried out for different parameters (e.g. C = 80000). These experiments showed the use of AE for 512 and 256 features, improves the number of times that the success rate is surpassed in comparison with NOAE (blue dashed line) 70.33% and 57.91% respectively. For these cases the values of the success rate remain high for more values of σ in comparison with NO AE. But, for

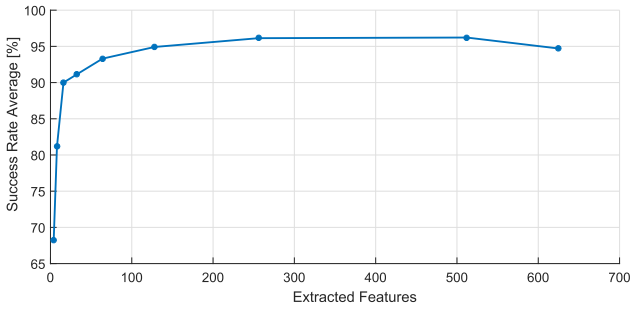

FIGURE 11. Success Rate Average vs. Features.

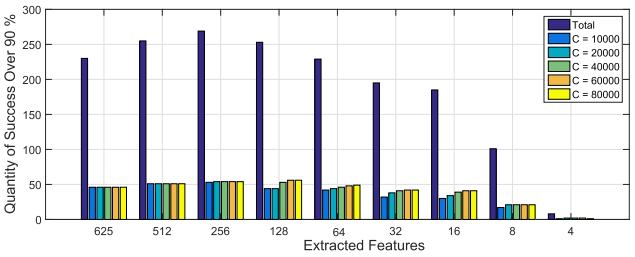

FIGURE 12. Success Rate vs. Features.

the rest of cases of features (4, 8, 16, 32, 64 and 128) the results are different. Which means that the use of AE not always improves the success rate for the same values of the parameters C and σ.

Fig. 11 shows the average of success rate vs. the number of features. As it can be seen, the success rate is higher for AE of 512, 256 and 128 features. Nevertheless for 64 features the result is similar to the case without AE and for 16, 8 and 4 the value is lower. The behavior of the best success rate for each number of hidden layers is similar to the success rate average (better results for AE of 512, 256 and 128 features).

Fig. 12 shows the success rate (higher than 90%) vs. the features for different values of the C parameter. As it can be seen, for 512, 256 and 128 features, there are more cases where the success rate is higher than 90% for different values of C (from 10000 to 80000). This indicates that the uses of AE provides models with better success rates. Note that the case of NOAE is for 625 features.

Fig. 13 shows the reliability average vs. features. As it can be seen, the reliability of AE with 512, 256, 128, 32 and 16 features (30% - 31%) is higher than NOAE (blue dashed line); and for the cases of 8 and 4, it is lower (17% and 24%) than NOAE. This indicates that the use of AE provides more reliable models if the cases of AE 8 and 4 are not taken into account due to their low success rate.

Fig. 14 shows the reliability rate vs. the support vectors rate. As it can be seen, the higher values of support vectors rate have better reliability. But for different values of σ (from 10^{-1} to 10^{6}) you can obtain bad values of reliability for high values of support vectors rate. Which means that a high value of support vectors rate does not grant high reliability. This not include that a model with low support vector rate provide a more reliable classification.

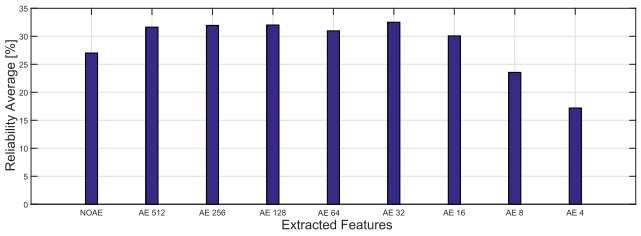

FIGURE 13. Reliability Average vs. Extracted Features.

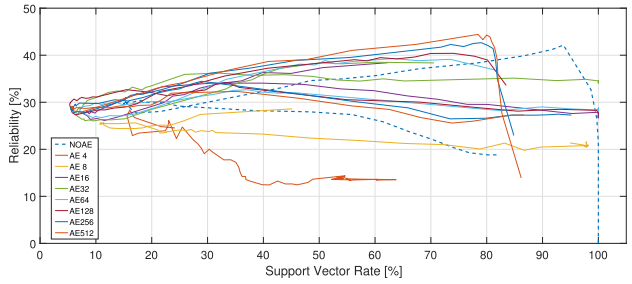

FIGURE 14. Reliability Average vs. Support Vectors Rate for C=500000.

Fig. 15 shows the results of Confidence and Credibility for the same range of Sigma. Note that the use of AE provides in general a higher success rate than the case without auto-encoder (NOAE). Regarding to the credibility, the use of AE keeps a high credibility (over 88%) for all evaluated σ. In the case of the confidence the use of AE maintains a high confidence (over 30%), which is higher than NOAE (blue dashed line) for many values of sigma.

B. H2: THE USE OF AE REDUCES OVER-FITTING.

Fig. 16 shows the average of support vectors rate vs. the success rate for the case of the parameter $C = 80000$. As it can be seen, for all the cases of features, the average of the success rate increases while the support vectors rate decreases. This indicates that a model with lower support vectors rate provides a model without over-fitting.

Fig. 17 shows the support vectors rate average vs. the features. As it can be seen, for AE 512, 256, 128, 64, 32, 16 and 8 the average of support vectors (around 25%) is less than the case of NOAE (blue dashed line) around 45%. This behavior is the same for the total data and different values of the parameter C. This indicates that the use of AE decreases the support vectors rate.

Fig. 18 shows the support vectors rate vs. parameter σ for C = 80000. As it can be seen the support vectors rate decreases when σ increases for the AE of 512, 256, 128, 64 and 32 features. But not in all cases of AE the support vectors rate is less than the case without AE. This indicates that the use of AE not always decrease the support vectors rate for the same values of C and σ.

C. H3: THE USE OF AE DECREASES THE PREDICTION TIMES.

Fig. 19 shows the test time added to the application time per image vs. the features for different values of C. As it can bee

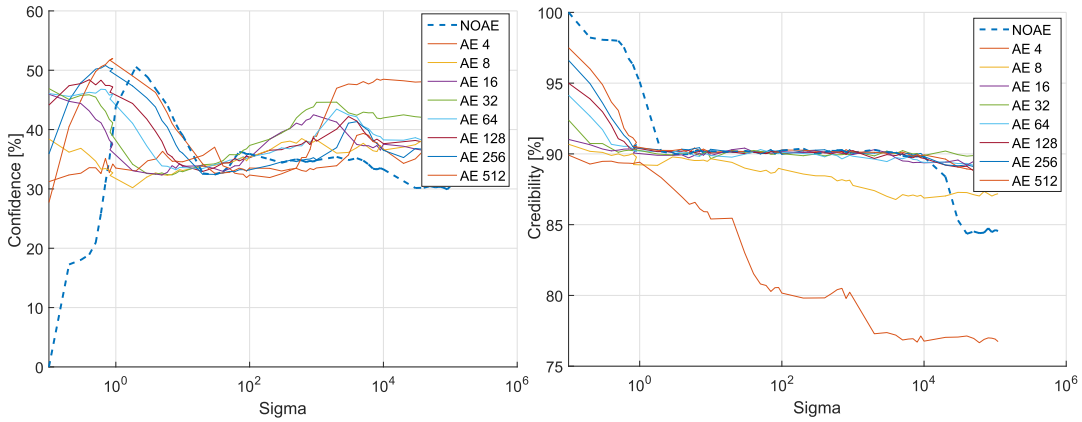

FIGURE 15. Confidence and Credibility for a simple AE.

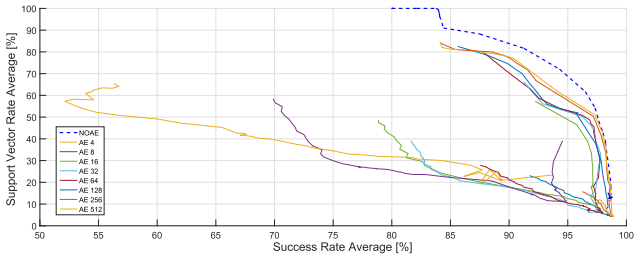

FIGURE 16. Support Vector Rate vs. Success Rate C=80000.

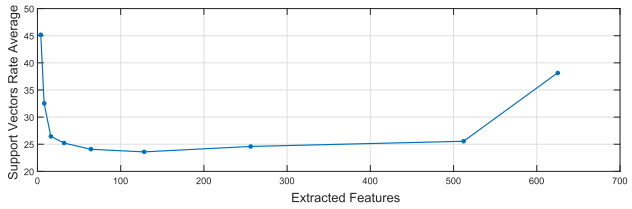

FIGURE 17. Support Vector Rate vs. Features.

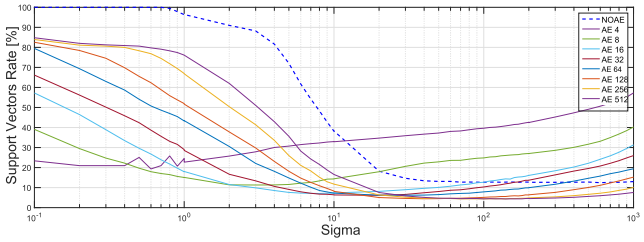

FIGURE 18. Support Vector Rate vs. Sigma.

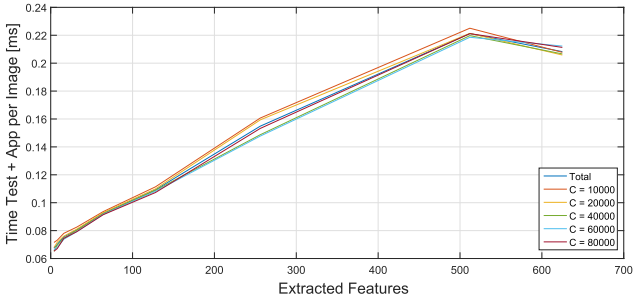

FIGURE 19. Processing and Classification Time vs. Features.

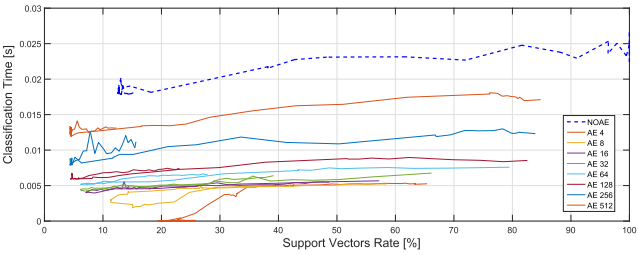

FIGURE 20. Classification Time vs. Support Vectors Rate.

seen, for one image in the cases from AE 4 to 256 features the times of the test and application are less than without AE. Only for AE 512 features these times are higher. This means that the use of AE decreases the prediction time in most of the cases.

Fig. 20 shows the test time vs. the support vectors rate for the same value of C. The increase of the classification time is more affected by the increases of the features than the number of support vectors (which has a little variation). This means that the use of AE decreases the prediction time only for the tests. Other tests were carried out for different values of parameter C. The results showed that time of the tests for different values of σ does not change in comparison with the increase of the extracted features.

A summarize of the hypotheses and their corresponding evidences are showed in table 3. First column represents the hypothesis which is related with the current evidence. Second column represents the evidence. Third columns represents figures related to the evidence. Fourth column represents the decision about the current evidence.

As it can be seen, the first two evidences are related to the first hypothesis (H1). Based in the carried out experiments

TABLE 3. Summary of hypotheses.

Hypotheses	Evidence	Figures	Evaluation
H1	AE provide models with better success rates	11, 12	Accepted
H1	AE provide more reliable models	13	Accepted
H2	Less SVR model provides more reliable classification	14	Rejected
H2	Less SVR model provides higher success rates (%)	16	Accepted
H2	AE decrease the SVR	17	Accepted
H2	AE decrease the SVR for the same C and σ	18	Accepted
H3	AE decrease the prediction time (including training)	19	Accepted
H3	AE decrease the prediction time (only test)	20	Accepted

TABLE 4. TS image classifier results with two AE.

Input Features	Extracted Features AE1	Extracted Features AE2	Success rate (%)	Support vectors rate (%)
625	256	64	98	7
625	**256**	**32**	**98**	**7**
625	**256**	**16**	**99**	**8**
625	256	8	97	12
625	256	4	86	26

represented by figures 11, 12 and 13, we can say that the uses of AE provides more reliable models with a better success rates. For these reasons, we can say that this hypothesis is accepted.

In the same way, the following four evidences are related to the second hypothesis (H2). With the obtained results, represented by figures 14, 16, 17 and 18, we can say that the uses of AE decrease the support vectors rate and provides models with higher success rates (more reliable classification). Based on these results we can say that this hypothesis can be accepted.

The last two evidences correspond to the third hypothesis (H3). Based on the obtained results, represented by figures 19 and 20, we can say that the uses of AE decrease the prediction time for tests and training. This imply that this hypothesis can be accepted.

A second experiment was performed to evaluate a deep network with two stacked AE for the TS image classifier. Results are shown in Table 4. Notice that it can be possible to have even better rates with a greater reduction in the feature space. Which is showed in second and third rows, where the support vectors rates are between (7% and 8%) while the success rates are between (98% and 99%) respectively.

IV. CONCLUSIONS

This paper assesses the use of deep learning, by means of special neural networks called auto-encoder, for automatic feature extraction in nuclear fusion databases. The feature extraction is an stage of image classification that is usually made manually. The use of AE allows to find suitable features for pattern recognition systems without human manipulation, which is a big advantage in the context of nuclear fusion where there are massive databases. The main disadvantage of this approach is that the parameters of the classifier (SVM in our case) should be re-tuned when the feature extraction is applied by using AE.

In this context and from the analysis, we can conclude that using AE can reduce the number of original features drastically, but with a high balance of accuracy and generality. The presented results support the following statements: a) The use of AE produces models with higher success rates, reaching in some cases up to 96% in average (2% over the performance without AE); b) The model predictions can be computed in less time when AE are used; c) The models can be twice faster than the cases without AE, although it could require parallel programming of AE for operation in real-+time; d) The classifiers with AE are more robust and better fitted for new images; and e) Classifiers with AE are able to make predictions for new images with up to 50% more of confidence and credibility.

Based on these promising results, we can say that this procedure could be applied to other kind of pattern recognition problems, such as the classification and information recovery of temporal evolution signals.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[2] R. Socher, B. Huval, B. Bath, C. D. Manning, and A. Y. Ng, "Convolutional-recursive deep learning for 3D object classification," in *Proc. 26th Adv. Neural Inf. Process. Syst. (NIPS)*, Lake Tahoe, NV, USA, Dec. 2012, pp. 656–664.
[3] S. Saha, G. Singh, M. Sapienza, P. H. Torr, and F. Cuzzolin, "Deep learning for detecting multiple space-time action tubes in videos," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, York, U.K., Sep. 2016, p. 133.
[4] A. Sehgal and N. Kehtarnavaz, "A convolutional neural network smartphone app for real-time voice activity detection," *IEEE Access*, vol. 6, pp. 9017–9026, 2018.
[5] R. A. Naqvi, M. Arsalan, G. Batchuluun, H. S. Yoon, and K. R. Park, "Deep learning-based gaze detection system for automobile drivers using a NIR camera sensor," *Sensors*, vol. 18, no. 2, p. 456, 2018.
[6] G. Psuj, "Multi-sensor data integration using deep learning for characterization of defects in steel elements," *Sensors*, vol. 18, no. 1, p. 292, 2018.
[7] C. Wu, C. Luo, N. Xiong, W. Zhang, and T. H. Kim, "A greedy deep learning method for medical disease analysis," *IEEE Access*, vol. 6, pp. 20021–20030, 2018.
[8] Q. V. Le, "Building high-level features using large scale unsupervised learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Vancouver, BC, Canada, May 2013, pp. 8595–8598.
[9] C.-Y. Liou, J.-C. Huang, and W.-C. Yang, "Modeling word perception using the Elman network," *Neurocomputing*, vol. 71, pp. 3150–3157, Oct. 2008.
[10] Z. M. Hira and D. F. Gillies, "A review of feature selection and feature extraction methods applied on microarray data," *Adv. Bioinf.*, vol. 2015, May 2015.
[11] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, p. 533, Oct. 1986.

[12] P. Baldi and K. Hornik, "Neural networks and principal component analysis: Learning from examples without local minima," *Neural Netw.*, vol. 2, no. 1, pp. 53–58, 1989.
[13] S. Rifai, P. Vincent, X. Müller, X. Glorot, and Y. Bengio, "Contractive auto-encoders: Explicit invariance during feature extraction," in *Proc. 28th Int. Conf. Mach. Learn. (ICML)*, Bellevue, WA, USA, Jun./Jul. 2011, pp. 833–840.
[14] L. Yuan, Y. Yang, Á. Hernández, and L. Shi, "Feature extraction for track section status classification based on UGW signals," *Sensors*, vol. 18, no. 4, p. 1225, 2018.
[15] T. Wen *et al.*, "Feature extraction of electronic nose signals using QPSO-based multiple KFDA signal processing," *Sensors*, vol. 18, no. 2, p. 388, 2018.
[16] P. Vincent, H. Larochelle, Y. Bengio, and P. A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, Helsinki, Finland, Jul. 2008, pp. 1096–1103.
[17] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho, and S. W. Baik, "Convolutional neural networks based fire detection in surveillance videos," *IEEE Access*, vol. 6, pp. 18174–18183, 2018.
[18] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in *Proc. 21th Int. Conf. Artif. Neural Netw. Mach. Learn. (ICANN)*, Espoo, Finland, Jun. 2011, pp. 52–59.
[19] A. Romero, C. Gatta, and G. Camps-Valls, "Unsupervised deep feature extraction for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1349–1362, Mar. 2016.
[20] H. Liang, X. Sun, Y. Sun, and Y. Gao, "Text feature extraction based on deep learning: A review," *EURASIP J. Wireless Commun. Netw.*, vol. 2017, no. 1, p. 211, 2017.
[21] L. Makili *et al.*, "Upgrade of the automatic analysis system in the TJ-II Thomson Scattering diagnostic: New image recognition classifier and fault condition detection," *Fusion Eng. Des.*, vol. 85, pp. 415–418, Jul. 2012.
[22] B. Cannas, F. Cau, A. Fanni, P. Sonato, and M. K. Zedda, "Automatic disruption classification at JET: Comparison of different pattern recognition techniques," *Nucl. Fusion*, vol. 46, no. 7, p. 699, 2006.
[23] J. Jin and J. Shi, "Automatic feature extraction of waveform signals for in-process diagnostic performance improvement," *J. Intell. Manuf.*, vol. 12, no. 3, pp. 257–268, 2001.
[24] J. Vega *et al.*, "Application of intelligent classification techniques to the TJ-II Thomson Scattering diagnostic," in *Proc. 32nd EPS Conf. Plasma Phys.*, Tarragona, Spain, Jun./Jul. 2005, pp. 1–5.
[25] G. A. Rattá, J. Vega, A. Murari, and JET-EFDA Contributors, "Improved feature selection based on genetic algorithms for real time disruption prediction on JET," *Fusion Eng. Des.*, vol. 87, no. 9, pp. 1670–1678, 2012.
[26] F. A. Matos, D. R. Ferreira, P. J. Carvalho, and JET-EFDA Contributors, "Deep learning for plasma tomography using the bolometer system at JET," *Fusion Eng. Des.*, vol. 114, pp. 18–25, Jan. 2017.
[27] G. Farias *et al.*, "Automatic feature extraction in large fusion databases by using deep learning approach," *Fusion Eng. Des.*, vol. 112, pp. 979–983, Nov. 2016.
[28] G. Farias, S. Dormido-Canto, J. Vega, I. Martínez, L. Alfaro, and F. Martínez, "Adaboost classification of TJ-II Thomson scattering images," *Fusion Eng. Des.*, vol. 123, pp. 759–763, Nov. 2017.
[29] L. Giudicotti, M. Bassan, F. P. Orsitto, R. Pasqualotto, M. Kempenaars, and J. Flanagan, "Conceptual design of a polarimetric Thomson scattering diagnostic in ITER," *J. Instrum.*, vol. 11, no. 1, p. C01071, 2016.
[30] S. Dormido-Canto *et al.*, "TJ-II wave forms analysis with wavelets and support vector machines," *Rev. Sci. Instrum.*, vol. 75, no. 10, pp. 4254–4257, 2004.
[31] L. Makili, J. Vega, S. Dormido-Canto, I. Pastor, and A. Murari, "Computationally efficient SVM multi-class image recognition with confidence measures," *Fusion Eng. Des.*, vol. 86, pp. 1213–1216, Oct. 2011.
[32] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul./Aug. 1998.
[33] V. Vovk, A. Gammerman, and C. Saunders, "Machine-learning applications of algorithmic randomness," in *Proc. 16th Int. Conf. Mach. Learn.*, Bled, Slovenia, Jun. 1999, pp. 444–453.
[34] C. Saunders, A. Gammerman, and V. Vovk, "Transduction with confidence and credibility," in *Proc. 16th Int. Joint Conf. Artif. Intell.*, Stockholm, Sweden, vol. 2, Jul./Aug. 1999, pp. 722–726.
[35] V. Vovk, A. Gammerman, and G. Shafer, *Algorithmic Learning in a Random World*. New York, NY, USA: Springer Science & Business Media, 2005.
[36] C. Alejaldre *et al.*, "First plasmas in the TJ-II flexible heliac," *Plasma Phys. Controlled Fusion*, vol. 41, p. A539, Mar. 1999.
[37] T. Estrada *et al.*, "Sheared flows and transition to improved confinement regime in the TJ-II stellarator," *Plasma Phys. Controlled Fusion*, vol. 51, no. 12, p. 124015, 2009.
[38] G. Farias, R. Dormido, M. Santos, and N. Duro, "Image classifier for the TJ-II Thomson scattering diagnostic: Evaluation with a feed forward neural network," in *Proc. 1st Int. Work-Conf. Interplay Between Natural Artif. Comput.*, Las Palmas, Spain, Jun. 2005, pp. 362–381.
[39] V. Vapnik, *The Nature of Statistical Learning Theory*, 2nd ed. New York, NY, USA: Springer Science & Business Media, 2013.
[40] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, Sep. 1995.
[41] H. Bhaumik, S. Bhattacharyya, M. D. Nath, and S. Chakraborty, "Hybrid soft computing approaches to content based video retrieval: A brief review," *Appl. Soft Comput.*, vol. 46, pp. 1008–1029, Sep. 2016.
[42] B. Scholkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA: MIT Press, 2001.

GONZALO FARIAS received the degree in computer science from the University of La Frontera, Temuco, Chile, in 2001, and the Ph.D. degree in computer science from the Universidad Nacional de Educación a Distancia, Madrid, Spain, in 2010. Since 2012, he has been with the Electrical Engineering School, Pontificia Universidad Católica de Valparaíso. His current research interests include machine learning, pattern recognition, simulation and control of dynamic systems, and engineering education.

ERNESTO FABREGAS received the B.S. degree in automation and the M.S. degree in digital systems from Polytechnic José Antonio Echeverría, Havana, Cuba, in 2004 and 2008, respectively, and the Ph.D. degree in computer science from the Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain, in 2013. He is currently a Post-Doctoral Fellow at UNED. His current research interests include pattern recognition, machine learning, simulation and control of multi-agent systems, and mobile robot applied to engineering education.

SEBASTIÁN DORMIDO-CANTO received the M.S. degree in electronics engineering from the Universidad Pontificia de Comillas, Madrid, Spain, in 1994, and the Ph.D. degree in physics from the Universidad Nacional de Educación a Distancia (UNED), Madrid, in 2001. Since 1994, he has been with the Department of Computer Science and Automatic Control, UNED, where he is currently a Full Professor of control engineering. His research and teaching activities are related with pattern recognition, machine learning, parallel processing, and nuclear fusion.

JESÚS VEGA received the M.S. degree from the Universidad Complutense de Madrid, Madrid, and the Ph.D. degree from the Universidad Nacional de Educación a Distancia, Madrid. He is working on nuclear fusion with CIEMAT, Madrid, and is very involved in the EURATOM Research Program on fusion projects, mainly, in the TJ-II stellarator and in the JET tokamak. He is also the Head of the Data Acquisition Unit, Spanish Fusion National Laboratory by Magnetic Confinement. His previous research activities were plasma diagnostic techniques in soft X- ray radiation. His current research is focused on both remote participation systems and advanced data analysis methods.

SEBASTIÁN VERGARA was born in Viña del Mar, Chile, in 1992. He is currently pursuing Electrical Engineering degree with the Pontificia Universidad Católica de Valparaíso. He will pursue the M.Sc. degree in electrical engineering at the Pontificia Universidad Católica de Valparaíso. His research areas include machine learning, computer vision, and automatic control.

SEBASTIÁN DORMIDO BENCOMO received the B.S. degree in physics from Complutense University of Madrid, Madrid, Spain, in 1968, the Ph.D. degree in sciences from the University of the Basque Country, Bilbao, Spain, in 1971, and the Doctor Honorary degree from the Universidad de Huelva and the Universidad de Almería. In 1981, he was appointed as a Professor of control engineering with the National University of Distance Education, Madrid. From 2001 to 2006, he was the President of the Spanish Association of Automatic Control, CEA-IFAC. He is currently an Emeritus Professor at the Universidad Nacional de Educación a Distancia. He has authored or co-authored over 250 technical papers in international journals and conferences and has supervised over 35 Ph.D. theses. His scientific activities include computer control of industrial processes, model-based predictive control, hybrid control, and Web-based laboratories for distance education. He received the National Automatic Control Prize from the IFAC Spanish Automatic Control Committee.

IGNACIO PASTOR was born in Madrid, Spain, in 1962. He received the B.S. and Ph.D. degrees in physics from the Complutense University of Madrid, in 1985 and 1994, respectively. His Ph.D. thesis was on nonlinear experimental laser dynamics. Since 1990, he has been with CIEMAT, where he has mainly done optical plasma diagnostic development for the TJ-II Stellarator, in particular Thomson Scattering (since 1994) and more recently near infrared spectroscopy. He has also contributed numerical and theoretical studies on nonlinear system dynamics and nonlinear Thomson Scattering.

ALVARO OLMEDO received the System Engineering degree from UNEXPO Luis Caballero Mejías, Caracas, Venezuela, in 2003, the Professorship in commercial education from the Instituto Universitario Pedagógico Monseñor Arias Blanco, and the M.S. degree in electrical engineering from the Universidad Central de Venezuela, Caracas, in 2011. He is currently pursuing the Ph.D. degree in systems engineering and control with the Universidad Nacional de Educación a Distancia, Spain. His research interests are focused in deep learning, deep neural networks, machine learning, and pattern recognition.