Learning Control Coil Currents from Heat-Flux Images using Convolutional Neural Networks at Wendelstein 7-X

Fabio Pisano[1], Barbara Cannas[1], Alessandra Fanni[1], Giuliana Sias[1], Yu Gao[2], Marcin Jakubowski[2,3], Holger Niemann[2], Aleix Puig Sitjes[2] and W7-X Team[2]

[1] Department of Electrical and Electronic Engineering, University of Cagliari, Cagliari, Italy
[2] Max-Planck-Institut für Plasmaphysik, Teilinstitut Greifswald, Greifswald, Germany
[3] University of Szczecin, Institute of Physics, Szczecin, Poland

Abstract

An important goal of Wendelstein 7-X (W7-X), the most advanced operating fusion experiment of the stellarator line, is to demonstrate the ability of stellarators to perform steady-state discharges. In this respect, the monitoring and control of the heat loads on the plasma facing components, especially of the strike-lines in the ten island divertors, will be critical during next operation phase OP2. In this paper, it is shown that deep convolutional neural networks are able to learn the relationship between the heat-flux images, obtained by the analysis of thermographic data, and the applied control coil currents in standard magnetic configuration experiments. This study is carried out in view of understanding and modelling the relationship between the heat-flux distribution in the divertor strike-lines and the actuators influencing them.

Keywords: strike-line, control coil currents, convolutional neural networks

1. Introduction

Wendelstein 7-X (W7-X) is an advanced stellarator with modular superconducting coils which enable steady state plasma operation in order to explore the reactor relevance of the W7-X concept. W7-X is equipped with several diagnostic systems that allow monitoring of the machine during experiments. Among these diagnostic systems, there are video [1] and infrared cameras [2], which monitor power loads on Plasma Facing Components (PFCs). One of the fundamental issues for magnetic fusion devices is to secure controlled power exhaust within acceptable limits for the PFCs. A dominant fraction of the energy leaving the confined plasma region is guided by a so-called island divertor [3] towards ten divertor units, five upper and five lower divertors. Each divertor unit consists of a horizontal and a vertical target and it is designed to sustain a maximum heat flux of up to 10 MW/m^2 [4]. The structure of the power deposition pattern in W7-X depends on magnetic configuration, plasma parameters and chosen scenario.

During operation phase OP1.2, ten infrared (IR) thermographic systems with wide-angle optics were installed to monitor the surface temperature on the fine-grain graphite plates of the ten inertially cooled Test Divertor Units (TDUs) [2]. This allows to investigate heat fluxes [5] as well as local effects of leading edges [6], error fields [7] and particle drifts [8]. The identification and control of the regions of highest thermal load (strike lines) on the targets is one of the most important and challenging tasks. A complex interplay of magnetic topology in the island boundary, local shaping of the PFCs, ratio between parallel- and cross-field transport results in the strike lines observed. Actuators such as ECCD (Electron Cyclotron Current Drive) and control coils were tested in the first divertor campaign for an active control of power distributions on the divertor. In standard magnetic configuration, the distribution of energy deposited among the divertors is to some degree asymmetric due to $n = 1$ and $n = 2$ error fields [9]. Nevertheless, this asymmetry can be minimized with a little effort, thanks to existing correction coils at W7-X [7]. The set of correction coils is composed by five trim coils and ten control coils, as shown in Figure 1a.

Figure 1b shows a cross-section of the boundary magnetic islands at toroidal angle $\varphi = 0°$ for standard magnetic configuration in one of the five modules of the stellarator. The position and size of these magnetic islands depend on several

factors, e.g. the magnetic configuration, plasma beta, heating power, the toroidal current and the control coil currents. The position and size of the magnetic islands influence the position of the strike-line onto the divertor plates. Each divertor module is equipped with two control coils, as previously introduced, that create additional magnetic fields allowing to vary the position and the shape of the strike line, especially for the purpose of controlling the strike-line shape and position in the divertor targets.

In [10] the effect of control coils operating with DC or AC currents was analyzed. The authors confirmed that control coils are a useful mean for strike line control and highlighted that, due to error fields and divertor misalignments, the reaction of strike lines formed on different divertors may be different, i.e. control coils would need to be operated independently, what is technically possible. Control coils have two basic functions:

- to change the island size, leading to the shift of strike lines towards or away from the pumping gap.
- to rotate the boundary island chains and, as a result, to sweep the strike lines poloidally on the divertor plates.

The first effect can be obtained in stellarator symmetric mode, where currents in upper and lower control coils have the same polarity. The second effect can be achieved in stellarator antisymmetric mode, where currents in upper and lower control coils have opposite polarity.

In [11] the variation of the strike-line position and the distribution of the heat-flux footprints are analyzed and correlated with the toroidal current variation. The toroidal current depends on the experimental parameters, e.g. magnetic configurations, heating power and plasma density. Moreover, the toroidal current is observed to be modified by ECCD. This suggests a possible candidate for the strike-line control with ECCD in the future.

The work in this paper should be seen in the broader context of a camera-based diagnostic system being able to recover information from obtained images. For W7-X it is required to detect and classify thermal events such as strike-lines, hot spots, etc., to assess the related risk and to take the proper actions. Presently, the termination of the experiment is the only foreseen action in the existing concept of the monitoring system. Regarding strike-line control, the actions will include sweeping them away from the critical position, i.e. a leading edge or the edge of the divertor target plates, via suitably varying the currents of the control coils. Thus, to face the strike-line control problem, we need to learn which control coil currents configuration is needed in order to obtain a desired heat-flux distribution and a corresponding strike-line pattern. In this paper, this relationship between heat-flux distribution on the divertor surface and control coil currents is investigated. The main objective is to assess if it is possible, in a fixed experimental condition, to infer the value of the applied control coil currents directly from the heat-flux distribution. Data from experiments in standard magnetic configuration performed during OP1.2b campaign for strike line control studies were used for this study.

The proposed methods come from the world of machine learning and, in particular, of deep learning [12][13], which

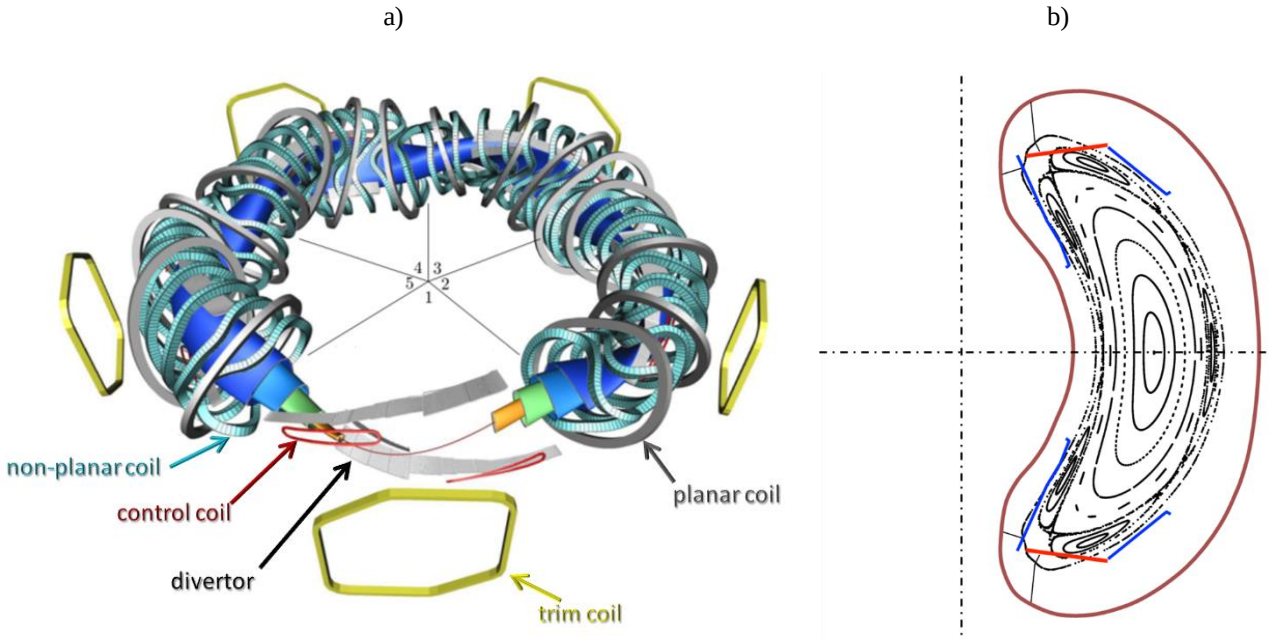

Figure 1 a) Coil system at W7-X; b) Poincaré plot showing the magnetic field lines and the boundary magnetic islands at toroidal angle $\varphi = 0°$ crossing the divertor targets in one module of the stellarator for standard magnetic configuration. Divertor targets in red, baffles in blue.

since 2010 is revolutionizing artificial intelligence. The keys of this revolution are the accessibility to large data sets and the use of massive GPU-based servers that allow to accelerate the learning process in case of very large models. Deep learning architectures make use of Deep Neural Networks (DNNs) which are composed of multiple processing layers able to learn complex representations of data with multiple levels of abstraction. These kinds of architectures are used in many applications like computer vision, speech recognition, bioinformatics and robotics with results comparable, and in some cases superior, to human experts. Among the different architectures, Convolutional Neural Networks (CNNs) are particularly successful for applications such as image classification, object detection and image registration.

Dealing with this problem by means of deep convolutional neural networks gives the possibility to avoid a feature extraction procedure for strike-line description. In fact, the first convolutional layers directly produce a set of features encoded by the same neural network which are then processed by the final fully connected layers. These features are implicit, since no physical meaning can be attributed to them, and encode the spatial and thermal properties of the strike-line. The CNN ability, to extract complex features expressing the image in much more detail, could be exploited in the future to select a much more compressed substitute of heat flux images to command the strike line control. A different set of strike-line descriptors for the limiter configuration was proposed in [14]. This set of features, generated by image processing from Fourier Descriptors, was extracted from IR images of the limiter.

Literature reports some applications of neural networks at W7-X whose aim is to develop and ultimately implement a feedback controller in the device operation system to ensure the safety of the first wall, in particular the divertor. In [15] NNs are trained to reconstruct the sum of the modular planar coil currents I_A and I_B of an ι-scan. The input to the NN is the heat flux image obtained through THEODOR. In [16] proxies for two edge magnetic field properties have been reconstructed given simulated heat load images on the W7-X divertor target plates. Six different artificial neural network architectures from shallow to complex Deep learning NN have been tested. The authors indentify deep convolutional neural networks as favorable algorithms for the stated problem. In [17] it is shown that the NNs show significantly improved performance on experimental data when the training set contains both experimental and simulated data. This is important since a very limited set of experimental training data is presently available at W7-X.

The paper is organized as follows: Section 2 gives an overview of the imaging system under development at W7-X. Section 3 introduces the CNNs architecture. Section 4 gives information about the case study and the database used. Section 5 shows the obtained results. Finally, in Section 6, the conclusions and the future work outlook are given.

2. Imaging System

Thermography is an essential tool for measuring the divertor surface temperature at W7-X. During the operation phase OP1.2b in 2018, ten high-resolution infrared systems were installed to monitor the heat fluxes over the whole divertor surface consisting of 5 upper and 5 lower divertors, each one equipped with a horizontal and vertical target. An endoscope was used to monitor the lower divertor in module 5. Nine immersion tubes, each one consisting of two visible cameras and one infrared micro-bolometric camera, were used to monitor the remaining nine divertors [2].

The Planck–law–based calibration of the infrared cameras has been carried out in the laboratory before the campaign, by providing a set of look-up tables which describes the relationship between the received digital levels and the surface temperatures. A software platform for the real time temperature observation incorporates the camera control and temperature conversion algorithms [18].

For each IR system, a scene model has been implemented to simulate the camera view towards the in-vessel components; this model contains a pixel-wise information about the observed PFC and its emissivity properties, the distance of the target material from the camera eye and the angle of the line of sight with respect to the surface normal, and the 3D coordinates of the observed target. This information is generated after a spatial calibration process, in which a pinhole camera model, reproducing the camera

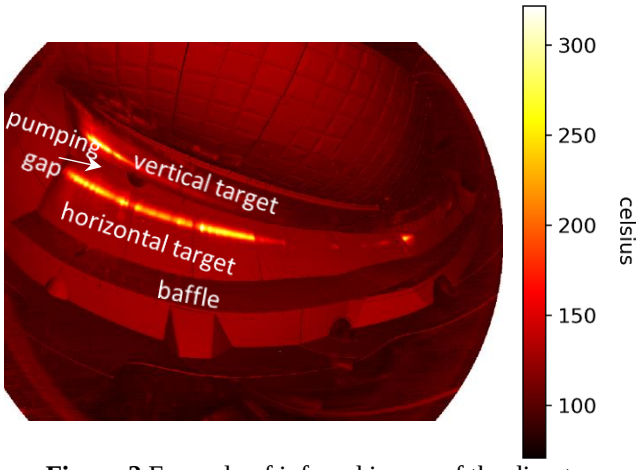

Figure 2 Example of infrared image of the divertor

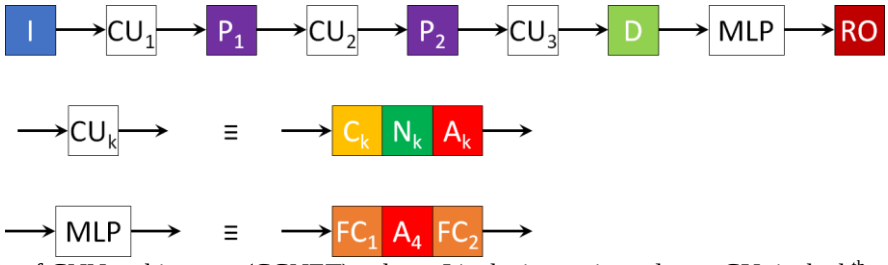

Figure 3 Example of CNN architecture (CCNET), where: I is the image input layer; CU_k is the k^{th} convolutional unit, composed by the cascade of a convolutional layer (C_k), a batch-normalization layer (N_k) and a nonlinear activation layer with ReLU functions (A_k); P_k is the k^{th} pooling layer; D is a dropout layer; MLP is a multi-layer perceptron, composed by the cascade of two fully-connected layers (FC_1 and FC_2) interconnected through a nonlinear activation layer (A_4); RO is the regression output layer.

perspective and the lens distortion effect, is derived [19]. Figure 2 shows an example of infrared image of the divertor. The regions with high temperature define the two strike-lines, one on horizontal target (on the top) and one on vertical target (on the bottom).

The heat flux distribution on the divertor elements has been calculated [5] by means of the so-called THEODOR (THermal Energy Onto DivertOR) code [20], which solves the 2D diffusion equation using a numerical discretization of forward-time central-space method.

3. Convolutional Neural Networks

In recent years, machine learning has become popular in research. Among various machine-learning algorithms, "deep learning" has established itself in solving problems of computer vision given the remarkable advancement in hardware technologies which allows to mimic the human ability in interpreting the visual information coming from images and videos. Among the architectures in deep learning in the field of computer vision, convolutional neural networks (CNNs) are the most used [12][13].

The first stage of a CNN is normally composed of several convolutional units CU_k followed by pooling layers P_k. Each convolutional unit is composed by a convolutional layer C_k usually followed by a batch normalization layer N_k, and by a ReLU (Rectified Linear Unit) layer A_k. Convolutional layers perform convolutional filtering of the input coming from the previous layers, normalization layer normalizes the output of the convolutional layer and the nonlinear ReLU layer keeps all input values above zero and sets all negative values to zero. Pooling layers produce a low-dimensional set of features by subsampling operations, such as maximum (max pooling layer) or average (average pooling layer). This set of features is then processed by few fully connected layers (identical to a MultiLayer Perceptron, MLP) which produces the output of the network. Also, the fully connected layers are normally connected through nonlinear activation layers, with ReLU or hyperbolic tangent function, improving the nonlinear regression capability of the network. A dropout layer D is usually inserted before the multi-layer perceptron in order to reduce overfitting on the training set and improve generalization. An example of CNN architecture is shown in Figure 3. More details on this architecture, called CCNET in the following, are reported in section 5 and in the Appendix.

4. Case Study

Data for this study comes from six experiments in standard configuration performed during OP1.2b campaign, in August 2018, for strike line control studies. The experimental settings of the discharges are: standard magnetic configuration (see e.g. Figure 1b), ECRH (Electron Cyclotron Resonance Heating) heating power P_{ECRH} = 3MW, plasma density n_e = $3.5 \cdot 10^{19} m^{-2}$. The experiments include the reference discharge in Figure 4 (named experiment XP0 in the following) where control coils are off, and five discharges characterized by the same experimental settings of the reference discharge, but with varying control coil currents in different ways:

- In experiment XP1, all control coil currents in the lower and upper divertors are switched on after 6s, in symmetric mode, ramping up to 900A with a slope of about 150A/s.

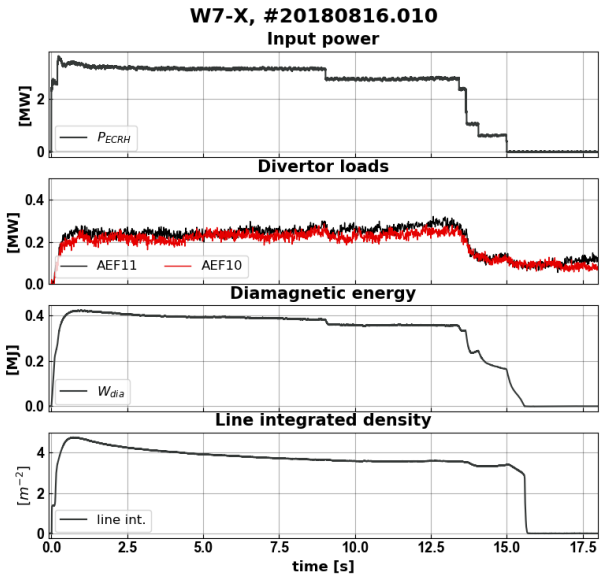

Figure 4 Overview plots for reference discharge

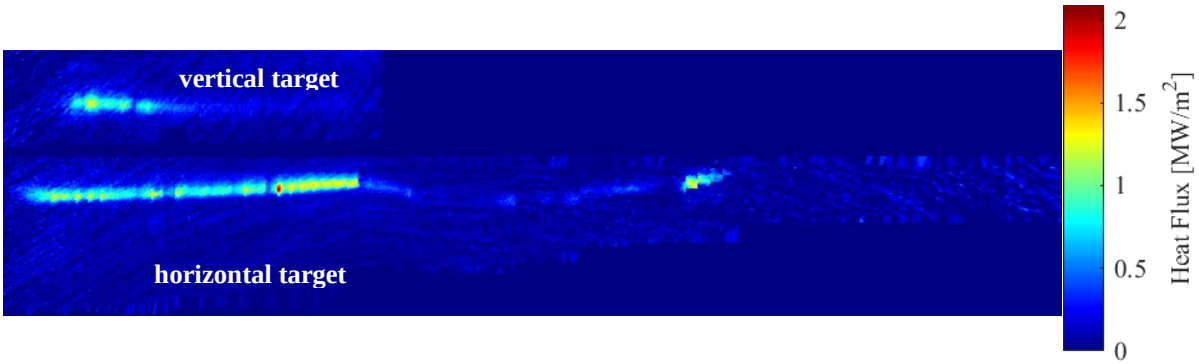

Figure 5 Example of a heat-flux image showing two strike-lines (one on vertical target, one on horizontal target).

- In experiment XP2, all control coil currents in the lower and upper divertors are switched on after 6s, in symmetric mode, ramping down to -1.35kA with a slope of about –150A/s.
- In experiment XP3 control coil currents in all the lower (upper) divertors are switched on after 6s, in antisymmetric mode, ramping down to –900A (up to 900A) with a slope of about –150A/s (150A/s).
- In experiment XP4 control coil currents in all the lower (upper) divertors are switched on after 6s, in antisymmetric mode, ramping up to 1.35kA (down to –1.35kA) with a slope of about 150A/s (–150A/s).
- In experiment XP5 control coil currents in all the lower and upper divertors are switched on after 10s, oscillating with frequency of 5Hz and gradually increasing the oscillation amplitude to 600A within 5s. The currents in the lower divertors have opposite polarity with respect to the currents in the upper divertors.

The behaviour of the toroidal current during all the investigated discharges was similar, starting from zero at the beginning of the discharge and approaching monotonically a limit value dependent on the experimental parameters [11]. The transient effect on strike-line position and shape correlated with the toroidal current evolution should be taken into account and combined with the effect of control coil currents.

For each experiment, IR videos of the surface temperatures in the ten TDUs are available. These videos have been produced by means of ten IR thermographic systems, each one looking at one out of ten divertors. The information about the surface temperature, together with the spatial and material properties of the observed surface, allows to calculate the heat-flux through the divertor target using the THEODOR code [20]. Further details regarding the heat flux calculation can be found in [5]. The obtained heat-flux values are projected on a 2D map of the corresponding W7-X divertor half module, by aligning the heat-flux profiles, defined as described in [5], next to each other in a picture. The single frame becomes a heat-flux image of size 1296×324 pixels, where the upper part represents the vertical target and the lower one the horizontal target (Figure 5). One image every 10ms is available.

4.1 Data preprocessing and selection

Time intervals in which the experimental parameters were stable have been used for the analysis. For example, the beginning of the discharge or phases with large variation of input power were not taken into account.

A data reduction process has been performed in order to discard frames characterized by similar currents configuration. In this way, the computational load during the training phase was reduced while maintaining a uniform density along the full operational space. In the following, a sample is considered as a set of simultaneous frames (coming from different cameras) and currents (toroidal and/or control coils) at a given time instant.

Table 1 Number of samples after/before the data selection process.

	Experiment						
Currents configuration	**XP0**	**XP1**	**XP2**	**XP3**	**XP4**	**XP5**	Total
null I_{CC}, low I_{TOR}	20/540	8/408	10/455	17/444	10/419	16/469	81/2735
null I_{CC}, high I_{TOR}	131/211	/	/	/	/	239/399	370/610
I_{CC} DC ramp, symmetric mode	/	264/643	409/895	/	/	/	673/1538
I_{CC} DC ramp, antisymmetric mode	/	/	/	262/607	411/931	/	673/1538
I_{CC} AC, antisymmetric mode	/	/	/	/	/	482/482	482/482
Total	151/751	272/1051	419/1350	279/1051	421/1350	737/1350	2279/6903

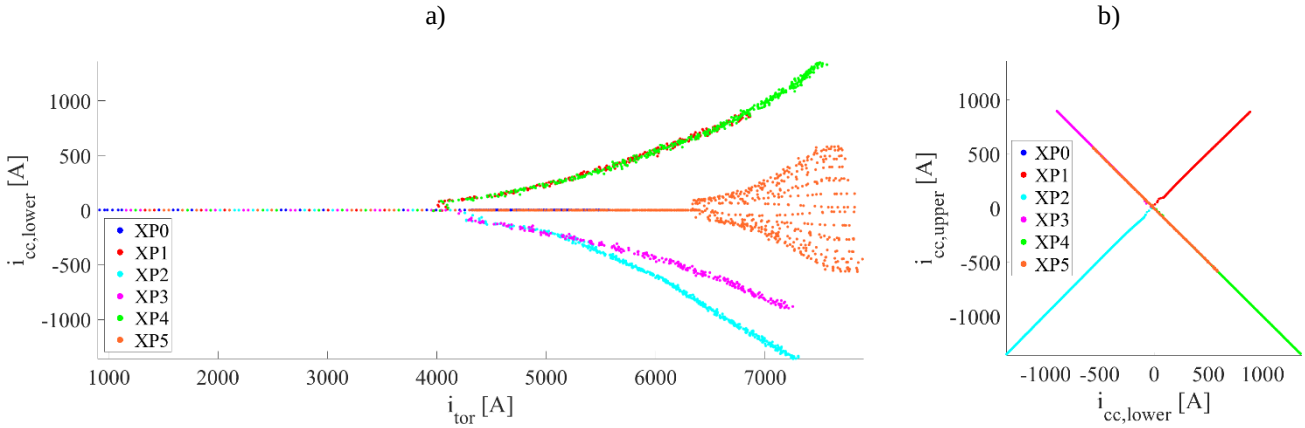

Figure 6 Operational space of the reduced dataset: a) lower divertor control coil current *vs* toroidal current; b) upper divertor control coil current *vs* lower divertor control coil current.

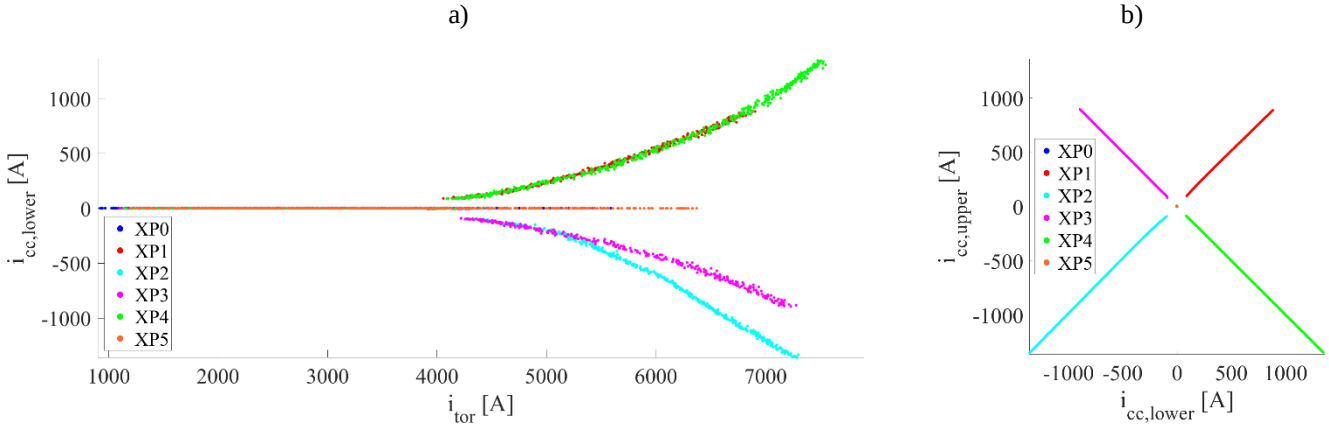

Figure 7 Operational space of the discarded dataset: a) lower divertor control coil current *vs* toroidal current; b) upper divertor control coil current *vs* lower divertor control coil current.

In this data reduction stage, the database has been reduced by taking into account five different combinations of control coil currents and toroidal current (see Table 1). The resulting reduced dataset is shown in Figure 6, where the corresponding control coil currents and toroidal current are plotted for each of the six experiments under study. Given the fact that the monotonic increasing evolution of the toroidal current was the same in all the experiments under consideration and that, during experiments, the onset time of the control coils was different (control coils were always off in XP0, were switched on after 6s in XP1, XP2, XP3 and XP4, and switched on after 10s in XP5), it was necessary to distinguish between samples with low toroidal current (t < 6s) and samples with high toroidal current (t > 6s). In fact, as it can be noticed in Figure 6a,samples with low toroidal current were always characterized by null control coil currents. On the contrary, samples with high toroidal currents were sometimes characterized by null control coil currents (in XP0 and XP5), some other times by not-null control coil currents (in XP1, XP2, XP3, XP4). From Figure 6a it is also possible to notice that, given that in XP5 the control coils were switched on only after 10s, a high value of the control coil currents in XP5 was linked to a higher value of the toroidal current, when compared to the other experiments.

Among the 6903 samples of the cleaned dataset, 81 samples with null control coil currents and low toroidal current, 370 samples with null control coil currents and high toroidal current, 673 samples with ramp in symmetric mode (from XP1 and XP2), 673 samples with ramp in antisymmetric mode (from XP3 and XP4) and all the 482 samples in AC (from XP5) have been selected, resulting in 2279 samples. The remaining 4624 samples, discarded during the data reduction process, are shown in Figure 7. As it can be noticed, since the points lie in the same regions of the reduced dataset, the reduced dataset results to be representative of the starting full dataset. Both in Figure 6a and Figure 7a, the slightly different slopes noticeable for XP2 and XP3 was due to a slightly different evolution in the toroidal current. This difference was not intentional, but only a consequence of the different configuration of the control coil currents.

The data reduction has been performed, in the cases with control coils switched off, by sorting the values of toroidal current from the lowest to the highest value, quantizing them and selecting just one sample for each quantization level. In

Table 2 Network architecture, inputs and outputs, datasets composition for each test.

Test #	Network Architecture	Inputs			Outputs		Datasets Composition	
		Lower divertor image (AEF10)	Upper divertor image (AEF11)	I_{TOR}	Lower I_{CC}	Upper I_{CC}	Null I_{CC} or I_{CC} DC	I_{CC} AC
1	$CCNET_0$	*			*	*	Tr/V/Ts	Tr/V/Ts
2	$CCNET_0^{mod}$	*		*	*	*	Tr/V/Ts	Tr/V/Ts
3	$CCNET_0$	*			*	*	Tr/V	Ts
4	$CCNET_0$		*		*	*	Tr/V	Ts
5	$CCNET_0$	*	*		*	*	Tr/V	Ts
6	$CCNET_0^{mod}$	*		*	*	*	Tr/V	Ts
7	$CCNET_0^{mod}$		*	*	*	*	Tr/V	Ts
8	$CCNET_0^{mod}$	*	*	*	*	*	Tr/V	Ts
9	$CCNET_0$		*			*	Tr/V	Ts
10	$CCNET_0$	*			*		Tr/V	Ts
11	$CCNET_0^{mod}$		*	*		*	Tr/V	Ts
12	$CCNET_0^{mod}$	*		*	*		Tr/V	Ts
13	$CCNET_1$		*			*	Tr/V	Ts
14	$CCNET_1$	*			*		Tr/V	Ts
15	$CCNET_1^{mod}$		*	*		*	Tr/V	Ts
16	$CCNET_1^{mod}$	*		*	*		Tr/V	Ts

the ramp up/down cases, the same procedure was applied to the values of the control coil current.

5. Results

Several tests have been performed in order to investigate the role of different parameters as network inputs, the prediction capability of the model in experiments never seen during training, and the impact of the network architecture. Table 2 reports for the performed tests, the network architecture used, the network inputs and outputs and the datasets composition.

The network inputs are chosen among the heat flux images of the lower and/or upper divertor for a given time instant, and for a corresponding value of the toroidal current. The network outputs are chosen among the control coil currents (I_{CC}) in the DC and/or AC experiments, behind the lower and/or upper divertor. Regarding the datasets composition, in the first two tests (*Tests 1–2*), data from the AC experiments have been included in the Training (Tr), Validation (V) and Test (Ts) sets, whereas in the others the networks have been trained with examples coming from null I_{CC} and I_{CC} ramp phases and tested on AC phases.

Two main CNN architectures have been presented in this paper for the different tests: CCNET (acronym of Control Coil NETwork), used in the tests without the toroidal current as input to the network, and the modified CCNET ($CCNET^{mod}$) used in the tests with the toroidal current as network input. The size of the input and output layers in each test depends on the size and number of inputs and outputs.

The details about the CCNET architecture (see Figure 3) are reported in Appendix A. In each training phase of CCNET, the network was trained for 1500 epochs with the stochastic gradient descent method with momentum [21], and then the network with the best results on the validation set was saved and used for the performance evaluation. The loss function used during the training was the Mean Squared Error (MSE).

The architecture of $CCNET^{mod}$ is shown in Figure 8, and obtained by adding the information given by the toroidal current to the trained CCNET. The toroidal current has been

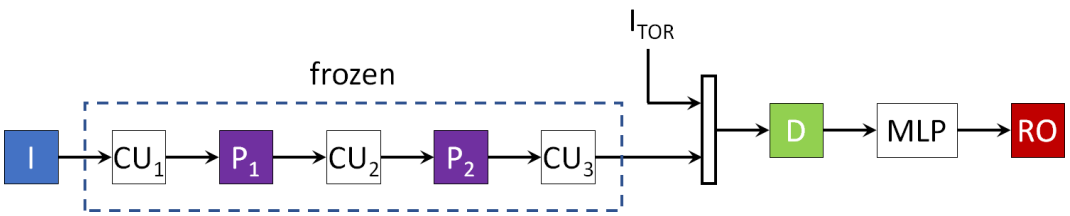

Figure 8 CCNET modified architecture ($CCNET^{mod}$) after addition of toroidal current as new network input. The new input (I_{TOR}) is concatenated with the feature vector output of CU_3, and fed to the MLP through the dropout layer D.

Table 3 CNN performance for each dataset and for each experiment in terms of normalized RMSE (%) on lower divertor control coil current

Test	Dataset				Experiment					
#	Training	Validation	Test	Discarded	XP0	XP1	XP2	XP3	XP4	XP5
1	1.95	4.71	5.31	2.57	1.64	2.41	3.14	2.43	2.58	3.81
2	1.98	4.57	5.02	2.28	0.97	2.46	2.98	2.30	2.46	3.61
3	1.12	1.90	8.92	1.48	0.24	1.59	1.72	1.36	1.69	5.67
4	1.55	2.39	16.77	2.07	1.46	1.81	2.32	1.85	2.10	10.67
5	1.64	2.25	13.22	1.98	0.58	1.83	2.67	1.84	1.90	8.40
6	1.01	1.83	9.47	1.55	0.25	1.93	1.59	1.33	1.64	6.02
7	1.61	2.37	13.95	2.03	1.37	1.85	2.24	1.77	1.84	8.99
8	1.39	2.22	11.70	1.92	0.72	1.58	2.41	1.56	2.07	7.48
10	1.38	2.13	13.55	2.40	1.66	3.00	2.22	1.91	2.08	8.66
12	1.32	2.06	13.07	2.14	1.86	2.45	2.01	1.77	1.99	8.32
14	1.59	2.32	19.06	2.30	1.11	2.23	2.22	2.21	2.38	12.11
16	1.66	2.32	19.53	2.40	1.76	2.34	2.19	2.17	2.50	12.41

Table 4 CNN performance for each dataset and for each experiment in terms of normalized RMSE (%) on upper divertor control coil current

Test	Dataset				Experiment					
#	Training	Validation	Test	Discarded	XP0	XP1	XP2	XP3	XP4	XP5
1	1.69	4.80	5.34	2.48	1.54	2.33	3.12	2.49	2.49	3.89
2	2.02	4.72	5.11	2.45	1.04	2.96	2.88	3.01	2.22	3.66
3	1.01	1.87	22.86	1.52	0.41	1.79	1.76	1.35	1.58	14.37
4	1.63	2.57	29.38	2.16	1.58	2.13	2.44	2.00	2.03	18.52
5	1.34	2.32	28.00	2.09	0.78	2.01	2.56	2.03	1.88	17.63
6	1.07	1.91	24.03	1.75	0.31	2.42	1.79	1.39	1.66	15.11
7	1.72	2.52	29.50	2.20	1.53	1.84	2.47	2.17	1.91	18.64
8	1.35	2.31	23.52	2.10	0.68	1.91	2.65	1.98	1.93	14.83
9	1.66	2.53	19.74	1.96	0.81	1.85	2.41	2.14	2.04	12.46
11	1.44	2.38	17.62	2.05	1.28	2.20	2.35	1.96	1.97	11.12
13	1.10	3.02	12.95	2.89	2.60	2.04	2.98	2.80	2.74	8.36
15	1.10	2.82	10.84	2.73	2.58	2.08	2.92	2.51	2.34	7.07

added as input to the dropout layer, concatenated with the feature vector produced by the convolutional blocks. Thus, the layers of the trained CCNET prior to the MLP were frozen, while the layers of the MLP were kept on training and adapted for other 500 epochs. Also in this case, the network with the best results on the validation set was selected, saved and used for the performance evaluation. As for CCNET, two variants ($CCNET_0^{mod}$ and $CCNET_1^{mod}$) depending on the type of pooling layers used and on the type of activation function used in the MLP have been used.

The performances obtained in the different tests, in terms of root mean squared error (RMSE) on the values of the control coil currents, are reported in Table 3 (lower) and Table 4 (upper). Also simpler models like linear models and MLP have been tested, but they were found to perform significantly worse than the proposed CNN architecture. For this reason, these results are not reported in the paper. The errors are shown in percentage with respect to the maximum absolute current value in the dataset (1.36kA), for each dataset (training, validation and test set, and for the samples discarded during the data reduction process) and for all the time instants of each experiment (XP0÷XP5), including the discarded data.

5.1 Tests 1–2

In *Test 1* the input of the $CCNET_0$ was the heat-flux image related to the lower divertor of module 1, generated from the image of the lower divertor, while the outputs were the values of the control coil currents in both lower and upper divertors.

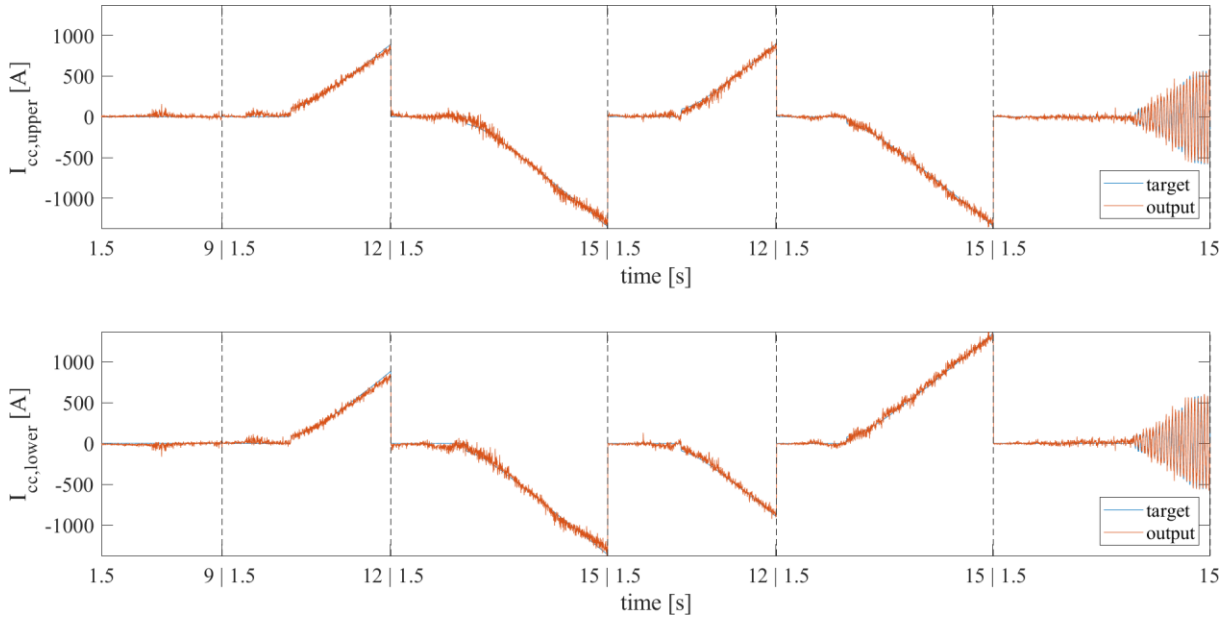

Figure 9 Output of CCNET[0] in *Test 1* superimposed on the actual time series of the control coil currents. The six time series related to the different experiments are stacked together and separated by dashed black vertical lines.

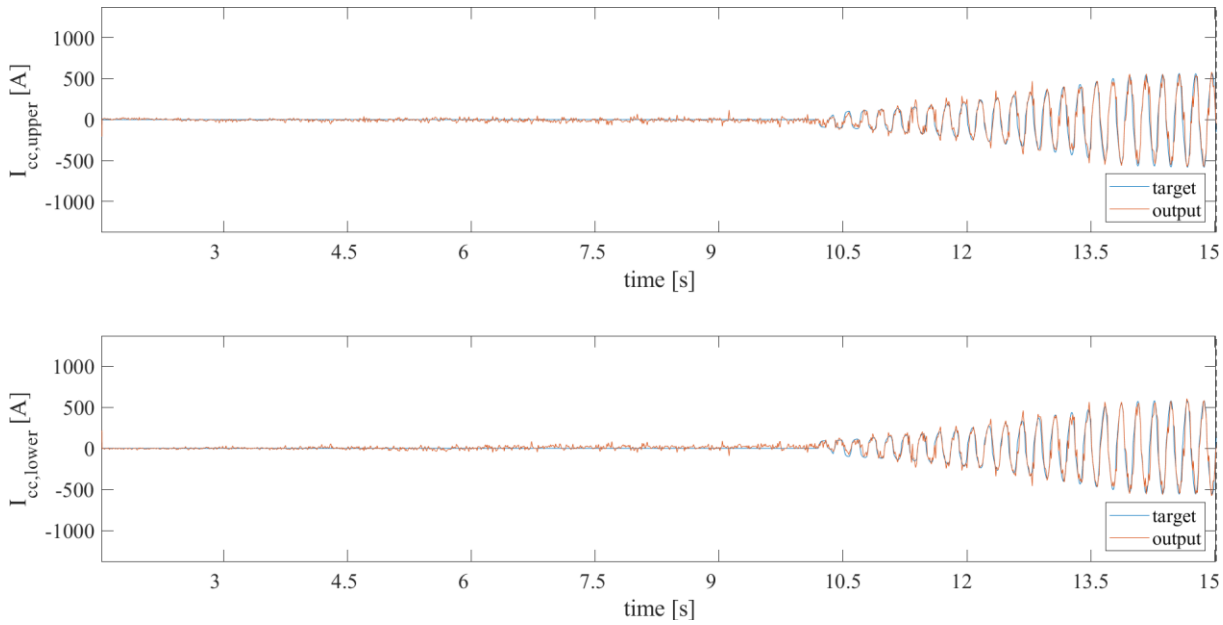

Figure 10 *Test 1*: zoom of Figure 9 for the AC experiment

All experiments have been used to feed the network during the training (1365 samples), validation (477 samples) and test (437 samples). As it can be noticed from the results in Table 3 and Table 4, despite a little overfitting on the training set, the CNN performance on the test set is quite good, with a good balance between the two outputs of the network. The performances on the discarded set, close to those on the training set, demonstrate that the data reduction policy was successful.

Figure 9 shows the output of the network superimposed on the actual time series of the upper and lower divertor control coil currents. Figure 10 shows a zoom for the AC experiment.

In *Test 2* the information given by the toroidal current has been added. As it can be noticed, the further input slightly improved the results.

5.2 Tests 3–8

In the previous tests, all experiments (null, DC and AC) are used to form the training, validation and test sets. This results in similar performances on all three data sets. The tests discussed below (*Tests 3–16*) aim to verify if the model learnt by the network from DC experiments only can predict control

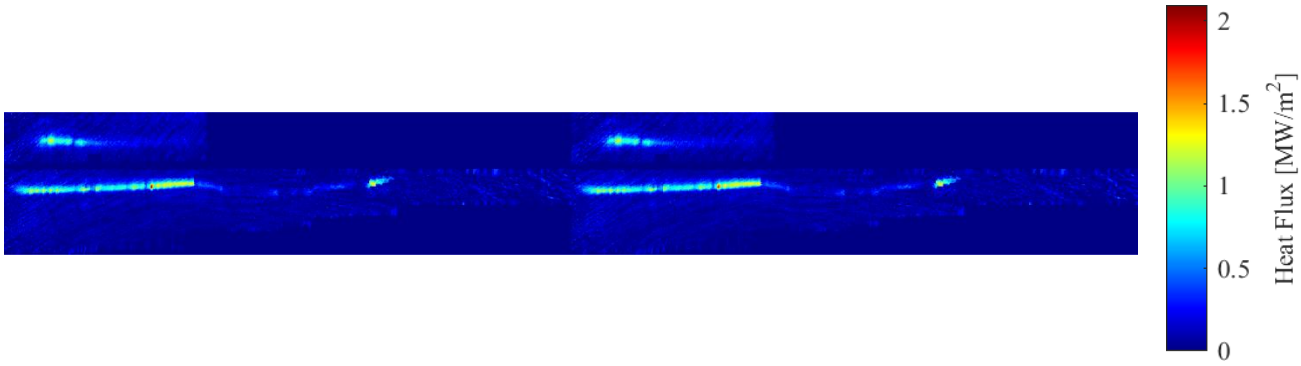

Figure 11 Example of concatenated heat-flux images from lower (left) and upper (right) divertor.

coil currents during AC experiments. A total of 1235 samples have been used to feed the network during the training, 512 samples to validate the network and 532 samples to test the network.

As it can be seen from the example in Figure 11, where the two heat-flux images from the two divertors in module 1 are stacked together horizontally (lower on the left), the strike-line patterns in the two divertors may differ considerably. This difference is mainly caused by the uncompensated error fields [7][9] and small misalignment of the divertor targets. In particular, the strike-line patterns in the divertor targets are different in terms of thickness and size. In order to evaluate the influence of the different heat-flux images, *Tests 3–5* have been performed.

In *Test 3*, only the heat-flux image of the lower divertor has been used as network input, in *Test 4* only the heat-flux image of the upper divertor has been used as network input, and in *Test 5* both heat-flux images of the upper and lower divertor have been used as network inputs and stacked together horizontally, obtaining a single input image of size 2592x324 (as in Figure 11). In order to support the addition of this new image input in *Test 5*, the number of neurons in the first fully connected layer has been doubled to 128, so that 64 features are generated by the first fully connected layer for the lower divertor image and 64 features are generated for the upper one. In all cases, the outputs were the control coil currents of the lower and upper divertors (see Table 2).

As it can be noticed from Table 3 and Table 4, for *Tests 3–5*, there is a large difference between the network performance on the upper control coil prediction and the one on the lower control coil prediction. In fact, the RMSE on the upper control coil current is more than twice that on the lower control coil current. This result is evident in Figure 12, where the output of the network trained in *Test 3* (for which the best performance on both currents is obtained) is superimposed on the actual time series of the control coil currents, and in Figure 13, where a zoom for the AC experiment is shown. As it can be noticed, the network is able to follow pretty well the evolution of the lower control coil current, whereas the evolution of the second output shows a decreasing trend with respect to the upper control coil current.

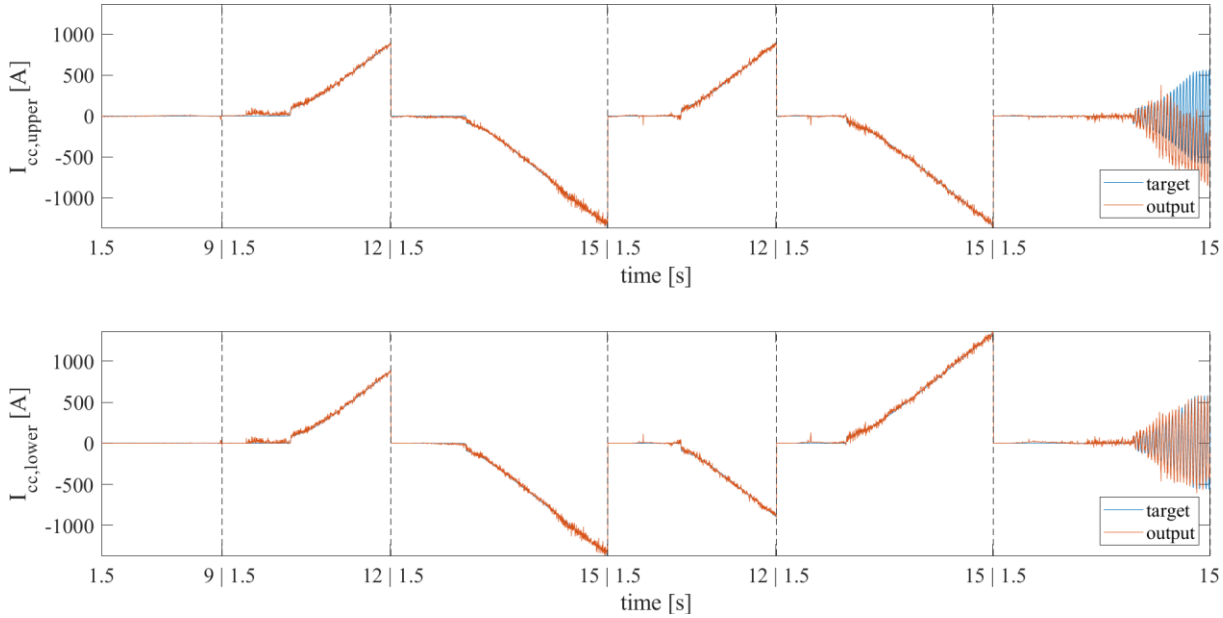

Figure 12 Output of $CCNET_0$ in *Test 3* superimposed on the actual time series of the control coil currents. The six time series related to the different experiments are stacked together and separated by dashed black vertical lines.

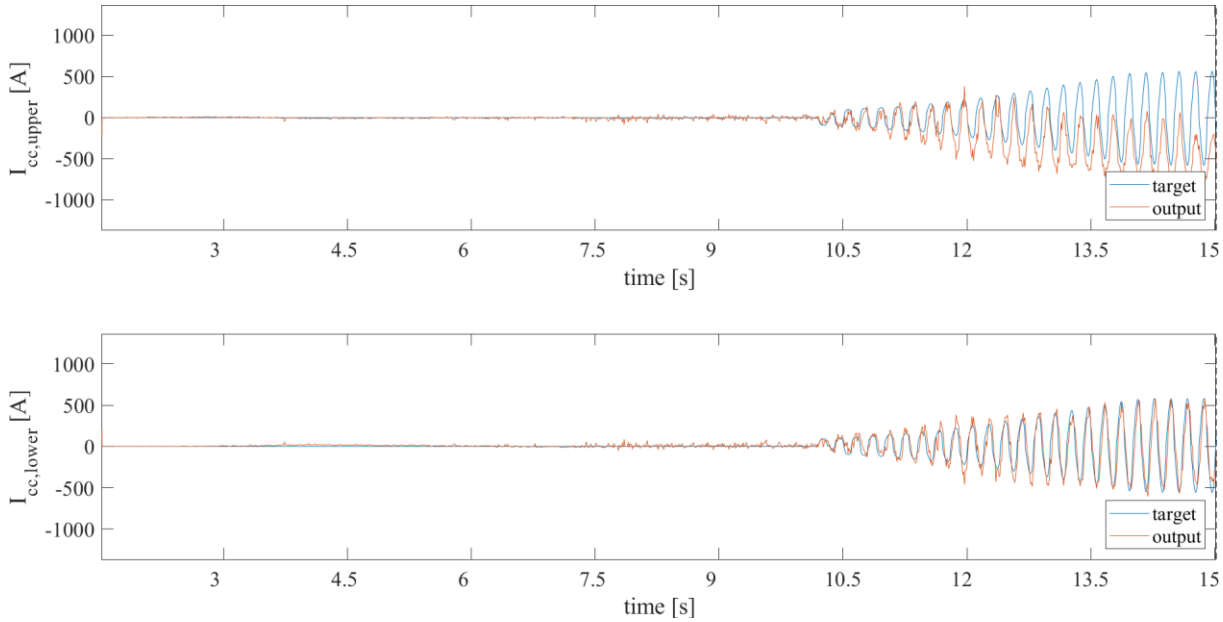

Figure 13 *Test 3* - zoom of Figure 12 for the AC experiment

The effect of adding the toroidal current (*Tests 6–8*) as an input is not substantial. Furthermore, adding the information from the upper divertor heat-flux image to the inputs of the CNN (*Tests 4–5*) did not improve the performance of the network. Indeed, the accuracy on the upper control coil current prediction is worse as compared to the previous case with the input consisting of the lower heat flux image only (*Test 3*).

5.3 Tests 9–12

Since addition of the upper image as input did not improve the performance, two separate networks have been trained: one (*Test 9*) infers the upper divertor control coil current from the upper divertor heat-flux image; the other (*Test 10*) infers the lower divertor control coil current from the lower divertor heat-flux image. The architecture of the two networks was that of $CCNET_0$ with a 1-dimensional output. Also, the addition of the toroidal current as a network input has been investigated in *Tests 11–12*.

As it can be noticed from Table 4, better performance on the test set for the upper divertor control coil have been obtained in *Test 11* when the network is fed with the upper divertor image and the toroidal current is added in input to the MLP block. Nevertheless, the upper control coil prediction is still a more difficult task with respect to the lower current prediction.

5.4 Tests 13–16

Some final tests have been performed in order to enhance the performance, by changing the network architecture. In the new architecture, called $CCNET_1$, a max pooling layer (P_2) has been used in order to emphasize more higher frequency components from the input image. Additionally, a hyperbolic tangent function in the activation layer A_4 of the MLP has been used to increase the symmetry in the last layers. In this case, the best performance on the test set for the upper divertor control coil have been obtained in *Test 15* when the network is fed with the upper divertor image and the toroidal current is added in input to the MLP block. Figure 14 and Figure 15 show the output of the network superimposed on the actual time series of the upper control coil current for *Test 15*.

Regarding the lower divertor control coil (*Test 14* and *Test 16*), the network tends to overfit on the training and validation set more than $CCNET_0$.

Summarizing, the change in the architecture (from $CCNET_0$ to $CCNET_1$) helped to highly improve the performance on the upper control coil current prediction, while worsening the performance on the lower control coil current prediction.

6. Conclusions

In this paper, the possibility of inferring the value of applied control coil currents directly from heat-flux images, in a fixed experimental condition, has been investigated by means of convolutional neural networks. Several tests have been performed in order to investigate the role of different input parameters and architectures on the network performance and the accuracy of predictions. Results showed that CNNs are able to supply a model between heat-flux images and control coil currents for the given experimental condition. Modelling this relationship is an exercise of inverse modelling [22][23]. Inverse models are held to allow determining the commands (control coil currents) necessary to achieve a desired state (heat-flux distribution).

In particular, when data from all experiments with monotonic and AC currents in the control coils were used to train the network, the prediction presented a good

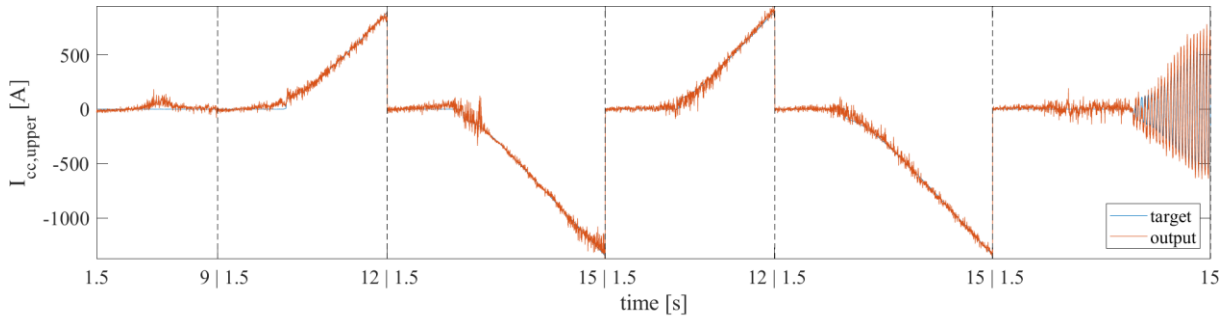

Figure 14 Output of $\mathrm{CCNET}_1^{\mathrm{mod}}$ in *Test 15* superimposed on the actual time series of the upper control coil current. The six time series related to the different experiments are stacked together and separated by dashed black vertical lines.

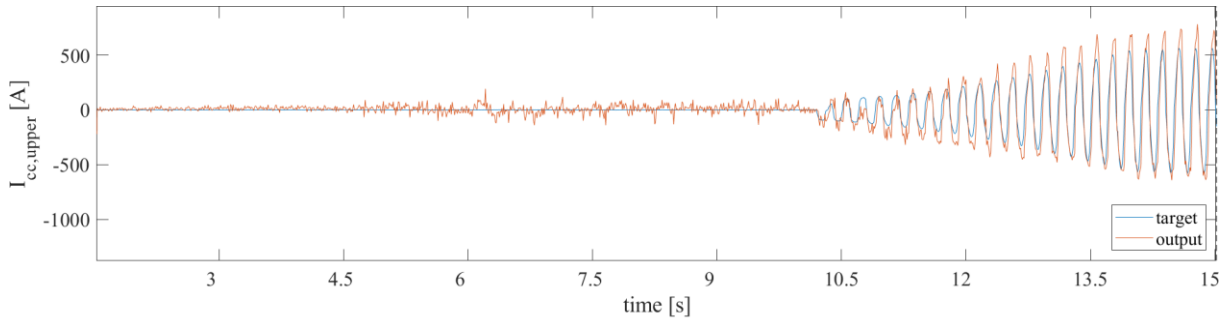

Figure 15 *Test 15* - Zoom of Figure 14 for the AC experiment

performance in the test data. In fact, the prediction root mean square error (RMSE) is about 70A, (about 5% of the maximum value of the current during the experiments).

Regarding the possibility to apply the network in experiments with AC currents never seen during the training, the results showed that also in this case the network is able to learn a static relationship between heat-flux images and control coil currents, and can be applied to predict the oscillating behavior of the control coils currents even when trained with data from experiments with monotonic currents. The analysis of the images of upper and lower divertors, in the experiments with opposite AC control coil currents, shows differences in strike-line size and positions. The possible causes could be tied to the error fields and particle drift effects, so that the control coil currents evolution doesn't produce the same effect on the heat-flux distribution of the two divertors, resulting in different strike-line patterns. Summarizing, a network trained with the heat-flux image of the lower divertor as input and the currents from both control coils as targets (*Test 3*), obtained the best performance in predicting the lower control coil current, with a prediction error of about 120A (~9% of accuracy) in terms of RMSE. Conversely, best results in predicting the upper control coil current have been obtained in *Test 15*, with a prediction error of about 150A (~11% of accuracy). In this test, the heat-flux image of the upper divertor and the toroidal current were used as inputs to the CNN, and the current from the upper divertor control coil was used as target. The two networks have different pooling layers and MLP activation function. The proposed architectures allow to differently extract and process the peculiar features from the lower and upper divertor heat-flux images.

Regarding the future prospects, the study will be extended to different magnetic configurations and experimental conditions. A general model will be developed able to infer control coil currents from strike-line properties. For this reason, the architecture of CCNET will be adapted and, in case, expanded in order to take into account a larger set of inputs parameters including magnetic configurations and experimental conditions. In particular, the influence of higher values of the toroidal current will be assessed.

Moreover, the possibility to build a forward model, including relevant plasma parameters, will be investigated. This forward model should be able to simulate the heat-flux distribution on the divertor target and the strike-line properties, when fed with the control coil currents, the toroidal current and all the relevant plasma parameters.

Furthermore, an autoencoder will be used to automatize the feature extraction procedure in a unsupervised way, exclusively depending on the heat-flux image and not on magnetic configuration, experimental conditions and control coil currents. It will be possible, in such a way, to reconstruct the heat flux images directly from a compressed set of features, describing the strike-line pattern, by means of a deconvolutional neural network. These features can be linked with the applied control coil currents and relevant plasma parameters and experimental conditions by means of simpler models, such as MLP networks, in both ways (forward and inverse model).

Acknowledgements

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014–2018 and 2019–2020 under Grant Agreement No. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Appendix A. Architecture of CCNET

The architecture of CCNET is shown in Figure 3, and has two variants ($CCNET_0$ and $CCNET_1$) depending on the type of pooling layers and MLP activation function. This difference is related to the possibility of filtering the input images in different ways, i.e., using low pass filtering as in average pooling layers or high-pass filtering as in max pooling layers.

Regarding the MLP activation function, two activation functions have been tested. One is the ReLU function, commonly used in deep learning for its convergence performance, better than sigmoid (not vanishing gradient and small computational load). The second is the hyperbolic tangent function, commonly used in classic MLP networks.

In CCNET, a first convolutional unit (CU_1), made up of three layers, followed by an average pooling layer (P_1), with pool size 9x9 and stride 9x9, transforms the input image into an 8-channels image. A second convolutional unit (CU_2), made up of three layers, followed by a second pooling layer (P_2), with pool size 9x9 and stride 9x9, transforms the 8-channels image into a 12-channels image. This second pooling layer is an average pooling layer in $CCNET_0$, while a max pooling layer in $CCNET_1$. A third convolutional unit (CU_3), made up of three layers, filters out the 12-channels image by means of 16 filters, obtaining a 16-channels image. Finally, a multi-layer perceptron (MLP), made up of two fully connected layers (FC_1 and FC_2) interconnected through a nonlinear activation layer (A_4) processes the features, output of the third convolutional unit, to predict the output in the regression output layer (RO). This last nonlinear activation layer has ReLU activation functions in $CCNET_0$, while hyperbolic tangent functions in $CCNET_1$. A dropout layer with dropout probability of 20% has been included between the third convolutional unit and the multi-layer perceptron in order to reduce overfitting on the training set and improve generalization.

References

[1] Kocsis G *et al.* 2015 *Fusion Eng. Des.* **96-97** 808
[2] Jakubowski M *et al.* 2018 *Rev. Sci. Instrum.* **89** 10E116
[3] Grigull P et al. 2001 Plasma Phys. Control. Fusion **43** A175
[4] Renner H *et al.* 2000 *Nucl. Fusion* **40** 1083
[5] Gao Y *et al.* 2019 *Nucl. Fusion* **59** 066007
[6] Endler M et al. 2018 Plasma Phys. Control. Fusion **61** 025004
[7] Lazerson S A et al. 2018 Plasma Phys. Control. Fusion **60** 124002
[8] Hammond K C et al. 2019 Plasma Phys. Control. Fusion **61** 125001
[9] Bozhenkov S A *et al.* 2017 *Nucl. Fusion* **57** 126030
[10] Ślęczka M. *et al.* 2018 Modulation of the strike line position using control coils in Wendelstein 7-X *45th EPS Conf. Plasma Physics* (Prague, Czech Republic, 2–6 July 2018) P2.1020 (http://ocs.ciemat.es/EPS2018PAP/pdf/P2.1020.pdf)
[11] Gao Y *et al.* 2019 *Nucl. Fusion* **59** 106015
[12] Pouyanfar S *et al.* 2018 *ACM Comput. Surv.* **51** 92
[13] Alom Z *et al.* 2019 Electronics **8** 292
[14] Pisano F *et al.* 2018 *Rev. Sci. Instrum.* **89** 123503
[15] Böckenhoff D *et al.* 2018 *Nucl. Fusion* **58** 056009
[16] Blatzheim M *et al.* 2019 *Nucl. Fusion* **59** 126029
[17] Blatzheim M *et al.* 2019 *Nucl. Fusion* **59** 016012
[18] Puig Sitjes A *et al.* 2018 *Fusion Sci. Technol.* **74** 116
[19] Pisano F *et al.* 2020 *Fusion Sci. Technol.* (accepted)
[20] Herrmann A *et al.* 1995 *Plasma Phys. Control. Fusion* **37** 17
[21] Qian N 1999 *Neural netw.* **12** 145
[22] Kawato M *et al.* 1999 *Curr. Opin. Neurobiol.* **9** 718
[23] Hussain M A et al. 2014 *Abstr. Appl. Anal.* **2014** 645982