A MACHINE LEARNING BASED TOOL FOR VOLTAGE DIP CLASSIFICATION

Houriyeh SHADMEHR
RSE S.p.A – Italy
houriyeh.shadmehr@rse-web.it

Riccardo CHIUMEO
RSE S.p.A – Italy
riccardo.chiumeo@rse-web.it

Liliana TENTI
RSE S.p.A – Italy
liliana.tenti@rse-web.it

ABSTRACT

A Machine Learning based tool is presented in order to make voltage dips (VD) ex-post analysis more automatic and effortless. The tool takes as input the full waveforms associated to voltage dips occurring in the Italian MV networks and recorded by QuEEN monitoring system implemented by RSE. The first tool has been developed to classify events on the base of their HV/MV origin since the utilities will be responsible only for the events due to faults occurred in their networks; it uses the self-tuning Kalman Filter and Support Vector Machine (SVM) for extracting the VD's features and classifying the events, respectively.
Instead, the second tool, based on end-to-end Deep Learning techniques, has been developed to distinguish between "true" and "false" VD; it utilizes a Convolutional Neural Network (CNN) whose first layers undertake the task of the features extraction while the last layers carry out the events classification.

INTRODUCTION

The Italian power quality monitoring system QuEEN has been performing Power Quality (PQ) measurements in the MV distribution network since 2006 [1]. Besides, monitoring the power quality indices (e.g. flicker, short voltage interruptions, voltage dips and swells, etc.), it is able to distinguish "true" from "false" voltage dips, these last ones occurring in the presence of voltage transformer (VT) saturation phenomena, for reliable statistical result reporting [2,3]. Moreover by ex-post analysis, starting from the recorded data, it's possible to detect the upstream and downstream origin (HV/MV) of the events, a useful as the MV events are the only disturbance in the DSO responsibility.

The monitoring system records automatically the events as upstream voltage dips by signals coming from HV line distance protections, while ex-post analysis allow the origin identification by comparing the occurrence time of events measured at primary substations underlying a common HV grid ("coincident" instants of voltage dips in the MV network are a proof of the HV origin of the event) (Fig. 2) ([1], [4]).
In the measurement units an advanced criterion, based on calculation of a 2nd harmonic component in the measured voltages [2, 3], is implemented in order to identify "false" events. In Fig. 2 the typical voltage waveform associated to "false" voltage dip is presented.

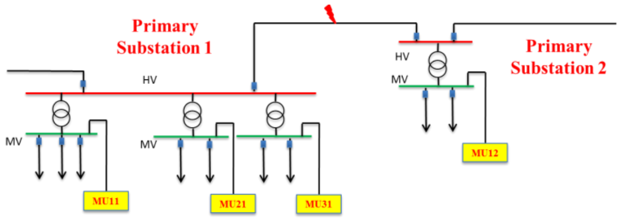

Fig. 1 – Event HV origin assessment by the "coincidence" of the records performed at primary substations 1and 2underlying a common HV grid

Although, the aforementioned methods adopted now for accomplishing those concerns have shown a good level of performance in the years, are not: i) fully automated, ii) able to give the right judgment for some difficult cases, relevant to the VT saturation effect consequent to a grounded polyphase fault clearance (e.g. the "true+false" voltage dips sequence in Fig. 3).

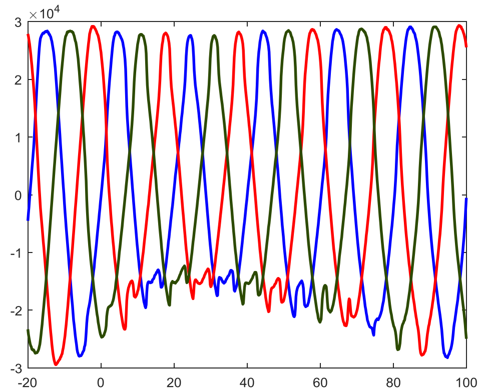

Fig. 2 – Line to line voltage waveform with "camel's humps", characterizing "false" voltage dips, due to VT saturation.

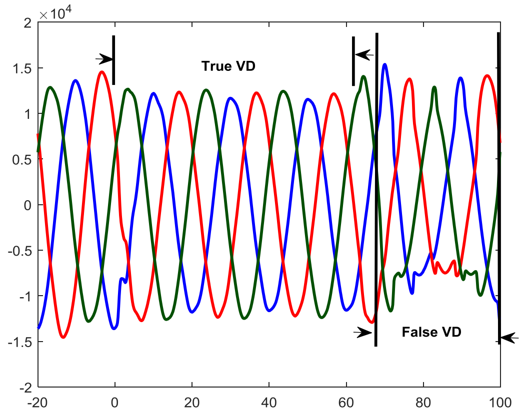

Fig. 3 – Line to line voltage waveform associated to a "true + false" voltage dips sequence.

In this paper, it will be presented two analytical tools, based on ML techniques, which have been developed for overcoming the above mentioned limits with the aim of increasing the monitoring system's performance.

FEXWAVES+SVMEC TOOL

The first tool which has been developed to detect the fault origin of the VDs (HV/MV) consists of two modules, as shown in Fig. 4: FEXWaveS (**F**eatures **EX**traction from **Wave**form **S**egmentation) and SVMEC (**S**upport **V**ector **M**achine **E**vent **C**lassifier). FEXWaveS module uses the self-tuning Kalman Filter for extracting the voltage dips features (such as: VD line-to-line voltages and shape, as shown in Fig. 5), while the SVMEC module utilizes the extracted features to classify events on the base of their HV/MV origin [5]. It's worth noting that SVMEC module takes other two features like: the dip duration and depth, both provided by the QuEEN system [5]. The tool accuracy is equal 0.91%. Nevertheless the FEXWaveS module relies completely on **manual features engineering**, hence being time consuming and requiring high human-efforts in the development stage.

Fig. 4 – Tool with manual feature engineering for the origin identification.

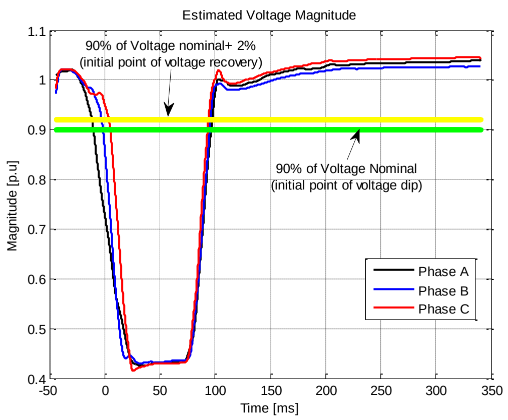

Fig. 5 – A rectangular "three voltages" event KF estimation.

Besides, it is not a module which could be generalized easily. For example, in Fig.6 FEXWaveS is applied to a "false" voltage dip with "camel's humps" characteristic in order to extract its particular feature with aim of individualizing it from a "true" one.

In fact comparing Fig. 6 and Fig.7, the estimated magnitude by KF doesn't show a specific feature useful to classify voltage dips as "false" or "true": in this case the criteria fails.

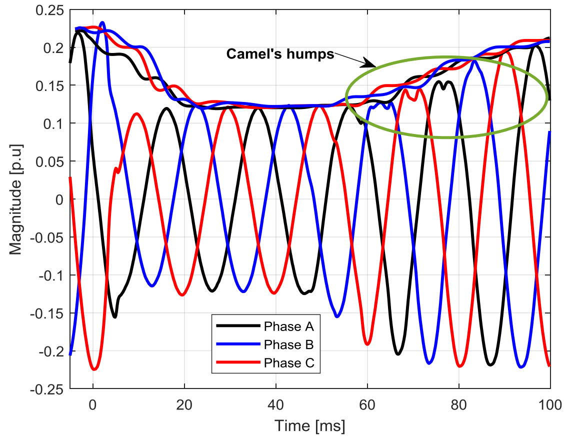

Fig. 6 – Estimate voltage magnitude from a "false" voltage dip line to line voltage waveform, characterized by typical "camel's humps" contour.

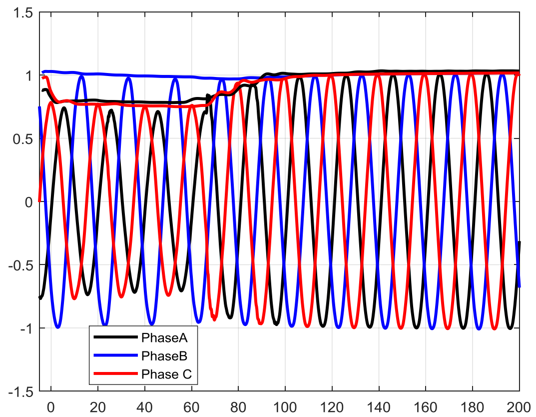

Fig. 7 – Estimate voltage magnitude from a "real" voltage dip line to line voltage waveform.

END-TO-END DEEP LEARNING TOOL

As previously mentioned, the primary tool relies on the manual features engineering in order to pull out features; this means it should observe many different cases and set a lot of rules to take account of all of them and that leads to raise the risk of error. As well as, the module could not be extended to other type of classification as that shown above. Hence, end-to end DL tool has been developed. This new module, based on a Deep Learning algorithm, replaces the FEXWaveS + SVMEC modules sequence and allows to extract features automatically (**automate features engineering**), as shown in Fig. 8.

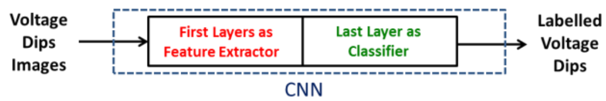

Fig. 8 – Developed tool with automated feature engineering.

In this solution the tool utilizes a Convolutional Neural Network (CNN) where: (i) input layer takes as input VDs waveforms transformed in JPG images; (ii) the intermediate layers extract the features by doing a convolution operation between kernel filters[1] and the inputs; (iii) the last layers (fully connected layer and classifiers "softmax") put all the extracted features together and calculate a probability score for all categories (**class labels**). The class with an higher probability score is the output of the classifier. For instance, if voltage dip probabilities of belonging to false/true voltage dip class are 0.9% and 0.1% respectively, then the event is classified as false.

The CNN network implemented in Matlab has indeed 3 layers:

1) the **input layer** which takes the transformed VD waveforms as JPG images with 875X656 pixels and 3 RGB channels (R: Red, G: Green B: Blue);
2) **one intermediate layer** performing the following operations:
   - convolution: 50 kernel filters with the 67X620 pixels whose dimensions correspond respectively to 10ms and peak-to-peak voltage magnitude over one cycle. These filters scan the input with an overlapping of 33 pixels. In such a way, 50 features of the VDs are extracted;
   - "rectification": in order to increase non-linearity to the model all the negative values obtained from the first operation are replaced by zero (the so called Rectified Linear Unit);
   - maxpooling: its aim is to downsample the results obtained from the previous operation with scanning the result by a 10X10 filter and select the maximum value placed in its dimension;
3) the **output layer** contains the fully connected layer and "softmax" classifier.

True/False voltage dip

As aforementioned, one of tasks carried out by the QuEEN monitoring system is to identify false voltage dip due to voltage transformer saturation phenomena. As it has been observed in Fig. 2, a false voltage dip has a typical deformation in its waveform, the so-called "**camel humps**" cited in [2, 3]; this false voltage dip characteristic allows to deal with the true/false VD identification problem as a typical "**pattern recognition**" problem in image processing.

A set of 840 true/false VD has been provided and divided randomly into two sets of VDs: a) **training set** for training the CNN network (80% of origin data set); b) **test set** for evaluating the CNN Network's performance (20% of origin data set).

[1] Every time a convolution operation is done, a feature is pulled out from inputs (e.g 20 convolution operations result in 20 features).

The following table (Tab. 1) presents the results of the end-to-end DL tool for 186 false/true voltage dips. The tool performance is around 86% with a success of 92.9% and 79.8% in identifying false and true VDs, respectively.

Tab. 1 – CNN network performance on a test set of 186 true/false VDs.

	Predicted Labels (Model)			
True Labels	False VD	True VD	Accuracy	
False VD	**78**	**6**	**92.9%**	**7.1%**
True VD	**17**	**67**	**79.8%**	**20.2%**
			Total Accuracy: 86%	

The tool capability has been also examined by inserting some difficult cases (Fig. 3). The accuracy in this case has decreased to 65%.

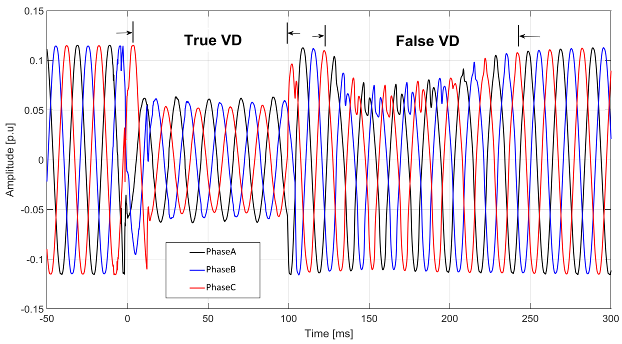

Fig. 9 – Line to line voltage waveform associated to a "true + false" voltage dips sequence.

This accuracy drop is caused by the new inserted cases (true+false VD sequences). Despite the adequate number of true + false VD sequences, the tool was not able to predict them correctly. For instance the VD in Fig. 9 has been identified as a false VD. This confirms that more features such as VD duration/depth must be injected to the fully connected layer, in order to improve the tool performance. For example, true+false VD sequence cases, typically, have shown a duration greater than 100 ms and this condition could be chosen as a their "typical" characteristic to be added to their feature vectors.

HV/MV voltage dip

The performance of the end-to-end DL tool for VD HV/MV origin assessment has been compared to the FEXWavS+SVMEC one. For this purpose, a set of 846 VDs of HV/MV origin has been provided. This time, the data set has been divided into three sets: a) **training set** (80% of the origin data set); b) **development set** (10% of the origin data set); c) **test set** (10% of the origin data set). As the new data are given to the CNN network, its parameters tuning is required as: the number of layers, the number of filters, etc.. Thus, a development set is adopted for setting up the optimal value of the aforementioned parameters, by calculating the network performance on them.

The model performance for the test set, is shown in Tab.2.

Tab. 2 – CNN network performance on a test set of 84 HV/MV VDs.

	Predicted Labels (Model)			
True Labels	HV	MV	Accuracy	
HV	**35**	**9**	**79.5%**	**20.5%**
MV	**7**	**33**	**82.5%**	**17.5%**
			Total Accuracy: 80%	

The end-to-end DL tool accuracy drops to 10 % with respect to the accuracy of the FEXWavS+SVMEC tool, as the second tool adopts only one feature, associated to the voltage dip's shape whereas the first tool employ four features.

From studying the failed cases by tool, it figured out that: a) the training set does not contain enough of those cases which are essential in the training phase of a CNN network; b) some of those cases require more features (e.g: VDs duration/depth) for being identified correctly by the classifier.

CONCLUSION

Two tools based on ML techniques have been developed in order to automate voltage dips ex-post analysis in monitoring system as much as possible. The first tool (FEXWavS+SVMEC) is based on manual feature engineering and support vector machine algorithm for detecting voltage dips upstream/downstream origin, whereas the second tool (end-to-end Deep Learning) is a Convolutional Neural Network (automated feature engineering) that has been employed to both identify true/false VDs and classify them on the base of their HV/MV origin.

Although the FEXWavS+SVMEC tool shows a 91% accuracy in classifying task, with respect to the end-to-end DL tool (accuracy equal or more than 80%), it is a time-consuming and error-prone tool due to relying on the manual feature engineering. On the contrary, the end-to-end DL tool is more flexible and user-friendly due to the adoption of automated feature engineering, as confirmed by applying the tool to different types of dataset.

In order to increase the performance of the end-to-end DL tool, as a future work, it will be evaluated the effect, on the tool performance, of inserting more features to the last layer of the model (the fully connected layer).

Acknowledgments

This work has been financed by the Research Fund for the Italian Electrical System in compliance with the Decree of Minister of Economic Development April 16, 2018.

REFERENCES

[1] R. Chiumeo, C. Gandolfi, L. Garbero, F. Malegori, M. Volta, L. Tenti, 2013, "The origin of voltage dips monitored in MV network and its effect on the evaluation of MV voltage dips performance indices", 22nd International Conference & Exhibition on Electricity Distribution, CIRED, paper 0655.

[2] R. Chiumeo, M. de Nigris, C. Gandolfi, L. Garbero, L. Tenti, E. Carpaneto, 2010, "Implementation of a New Method for an Improved Voltage Dips Evaluation by the Italian Power Quality Monitoring System in Presence of VT Saturation Effects", 14th International Conference on Harmonics and Quality of Power, ICREPQ'10, paper 383.

[3] R. Chiumeo, C. Gandolfi, L. Garbero, L. Tenti, E. Carpaneto, 2016, Power Quality, Cambridge Scholar Publishing, Newcastle upon Tyne, UK, 102-114.

[4] R. Chiumeo, L. Garbero, F. Malegori, L. Tenti, 2015, "Feasible methods to evaluate voltage dips origin", International Conference on Harmonics and Quality of Power, ICREPQ'15, paper 325-15.

[5] H. Shadmehr, R. Chiumeo, L. Tenti, 2016, "Self-tuning Kalman filter and machine learning algorithms for upstream and downstream origin detection", International Conference on Harmonics and Quality of Power, ICREPQ'16, paper 246-16.