

Article

Not peer-reviewed version

Hybrid LSTM+1DCNN Approach to Forecast Torque Internal Combustion Engines

Federico Ricci, [Luca Petrucci](#), [Francesco Mariani](#) *

Posted Date: 2 August 2023

doi: [10.20944/preprints202308.0149.v1](https://doi.org/10.20944/preprints202308.0149.v1)

Keywords: machine learning; LSTM+1DCNN architecture; neural network; torque



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Hybrid LSTM+1DCNN Approach to Forecast Torque Internal Combustion Engines

Federico Ricci, Luca Petrucci and Francesco Mariani *

Engineering Department, University of Perugia, Via Goffredo Duranti, 93, 06125 Perugia, Italy

* Correspondence: francesco.mariani@unipg.it

Abstract: Innovative solutions are now being researched to manage the ever-increasing amount of data required to optimize the performance of internal combustion engines. Machine Learning approaches have shown to be a valuable tool for signal prediction due to their real-time and cost-effective deployment. Among them, the architecture consisting of Long Short-Term Memory (LSTM) and one-dimensional Convolutional Neural Networks (1DCNN) has emerged as a highly promising and effective option for replacing the role of physical sensors. The architecture combines the capacity of LSTM to detect patterns and relationships in smaller segments of the signal with the ability of 1DCNN to detect patterns and relationships in larger segments of the signal. The purpose of this work is to assess the feasibility of substituting a physical device dedicated to calculating the torque supplied by a spark-ignition engine. The suggested architecture was trained and tested using signals from the field during a test campaign conducted under transient operating conditions. The results reveal that LSTM+1DCNN is particularly well suited for signal prediction with considerable variability. It constantly outperforms other architectures used for comparison, with average error percentages of less than 2%, proving the architecture's ability to replace physical sensors.

Keywords: machine learning; LSTM+1DCNN architecture; neural network; torque

1. Introduction

The more rigorous regulations regarding pollutant emissions from Internal Combustion Engines (ICEs), along with customer demands for increased performance, made vehicle control increasingly challenging [1–3]. The calibration and run-time operation of engines in different areas, such as automotive or aerospace, where huge amounts of data are required [4,6]. To successfully handle and process such produced information, significant computing efforts are necessary [7]. The many measurements obtained by sensors and monitoring systems during engine calibration operations are critical for fine-tuning and maximizing performance while also assuring efficient and dependable operation [8]. Furthermore, during run-time operations, the engines' real-time outputs are critical for monitoring engine health and detecting possible anomalies [9]. Advanced approaches are being researched to improve engine performance while lowering consumption, pollutant emissions, and operating longevity [10]. In the automobile industry, machine learning (ML) techniques are rapidly being employed to enhance computing performance and minimize costs [6,11]. Because of their small setup and low-cost hardware implementation, as well as their capacity to forecast operational parameters, they can reduce the number of operating points to be examined, resulting in significant memory and computational speed advantages [12]. LSTM+1DCNN appears to be a promising method to perform signal analysis among the ML approaches [13]. The LSTM (Long short-term memory) approach is a sort of Recurrent Neural Network (RNN) that can reproduce the sequential nature of non-linear observations across time [14,15]. One-dimensional convolutional neural networks (1DCNNs) are commonly utilized because of their simplicity and compact design in comparison to other neural network architectures, as well as their ability to quickly combine feature extraction and classification into a single adaptive learning body [16].

Ren et al. [17] proposed an Auto-CNN-LSTM technique for predicting the remaining useful life of a Lithium-Ion battery. The architecture surpasses two existing data-driven models based on ADNN and SVM, reducing prediction errors by more than 50%. Quin et al. [18] assessed the performance of an LSTM-1DCNN for anti-noise diesel engine misfire detection. The design has an average accuracy of approximately 98%, which is approximately 10% higher than other studied methods such as CNN, random forest, and deep neural networks. Based on high-speed flame image sequences, Lyu et al. [19] proposed an LSTM-1DCNN model for detecting thermoacoustic instability. The method achieves high accuracy (98.72%), sensitivity (99.99%), and specificity (97.50%), with a short processing time of about 1.23 ms per frame on a commercial GPU adapter which can be considered for real-time identification. A recent work of the same research group [20] developed a combined LSTM+1DCNN structure to evaluate the possibility of replacing a real sensor with a virtual one. This goal's achievement could be crucial for cutting costs and, in particular, avoiding the destruction of test bench components due to the resonance phenomena. The structure accurately reproduces the recorded signal's natural frequency, and the absolute difference between recorded and predicted values is always less than the fixed acceptable threshold of 10%.

This work intends to evaluate the possibilities of an LSTM+1DCNN structure in forecasting the torque delivered by a three-cylinder spark ignition engine. Under transient running conditions, experimental data from physical sensors and engine control unit (ECU) were utilized. Preliminary evaluations were carried out with the goal of optimizing the internal structure of the model and identifying the variables with the lowest impact on torque prediction. The LSTM+1DCNN was then tested, and its performance was compared to that of other optimized systems.

When compared to alternative neural architectures, the results demonstrated the ability of the LSTM+1DCNN structure to reproduce the target trend with reduced percentual error and a lower number of predictions below crucial thresholds. The proposed model is able to reach average percentual errors below 2%, without ever exceeding 10% of error on the single prediction.

2. Materials and Methods

2.1. Experimental Setup

Tests were performed on a 1L 3-cylinder turbocharged engine of 84 CV maximum power at 5250 rpm and 120 Nm of maximum torque at 3250 rpm. The internal cylinder bore is 72 mm while the piston stroke is 81.8 mm. The compression ratio is equal to 10:1. The engine operates in Port Fuel Injection (PFI) with European market gasoline (E5, with RON = 95 and MON= 85) injected at 4.2 bar absolute. A Borghi&Saveri eddy current brake dynamometer of 600 CV ensures the engine speed in firing condition (Figure 1a). A Vascat electric motor of 66.2 kW allows controlling the engine speed both in motored and firing conditions. All the engine parameters are controlled using an EFI EURO-4 engine control unit. The signals coming from thermocouples TCK and pressure sensors PTX 1000 are acquired by data acquisition systems of National Instrument. The indicated analysis is performed through a Kistler Kibox combustion analysis system (maximum temporal resolution of 0.1 CAD) that acquires the pressure signals coming from the piezoresistive sensors (Kistler 4624A) placed in the intake and exhaust ports, the in-cylinder pressure of the piezoelectric sensor (Kistler 5018) placed on a side of the combustion chamber beside the flywheel, the ignition signal from ECU and the absolute crank angular position measured by an optical encoder (AVL 365C). Due to structural and mechanical constraints, only the combustion chamber adjacent to the flywheel has been equipped with a piezoelectric sensor, which is used to determine the Indicated Mean Effective Pressure (IMEP). The torque delivered by the engine is measured using a torquemeter positioned near the engine crankshaft. All of the above quantities are recorded by AdaMo Hyper software during engine operations, allowing simultaneous management of the engine's speed, torque, and valve throttle position in both firing and motored states. Figure 1b summarizes the experimental layout.

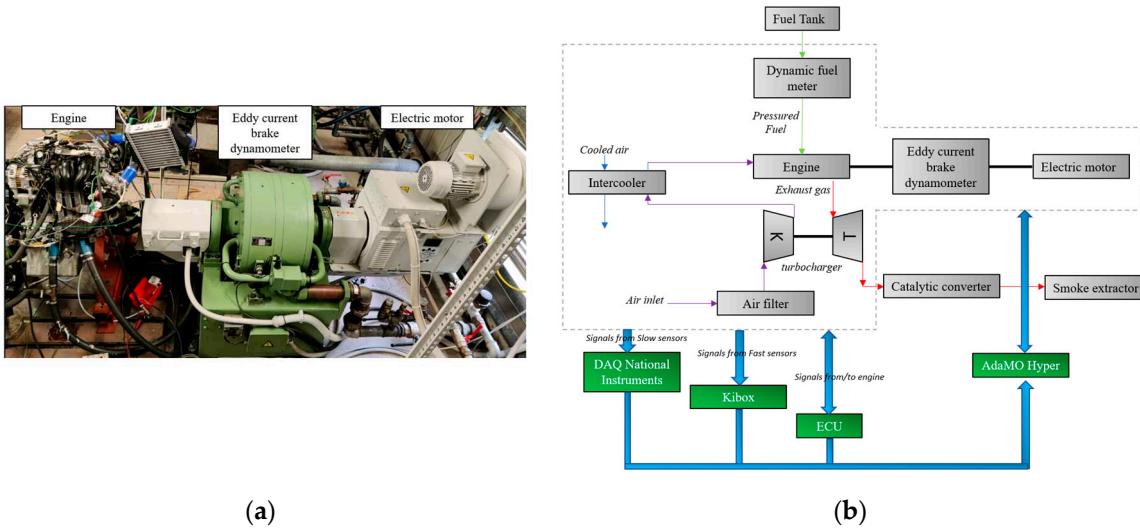


Figure 1. (a) SMART 1L three-cylinder turbo-charged. (b) configuration of the test bench.

2.2. Case Study

A transient cycle (Figure 2) has been chosen to preliminarily evaluate the performance of the LSTM+1DCNN proposed algorithm in predicting the torque delivered by the three-cylinder SI engine. A total of 12 variables acquired by AdaMo Hyper are initially selected among the most characteristics, as input parameters.

- Parameters coming from ECU: activation time of the injector (InjectionTime) and ignition timing of the spark (SparkAdvance) at the first cylinder beside the flywheel.
- Parameters coming from pressure sensors and thermocouples: temperature of the air before the filter (TC_Air_Intake), temperature and pressure of the air at the intake pipe (TC_ETB_OUT and MAP), pressure and temperature of the exhaust gas before (TC_Turbine IN, P_Turbine IN) and after the turbine (TC_Turbine OUT and P_Turbine OUT), temperature of the engine oil (TC_Engine Oil).
- Parameters related to the AdaMo actuation: throttle valve opening (Throttle Position) and engine speed (Engine speed).

The cycle is comprised of an input matrix of $[12 \times 28800]$ samples and of an output matrix of $[1 \times 28800]$ samples. 80% of the entire dataset is used for training sessions and the remaining 20% for the test sessions, i.e. Torque prediction (Figure 2).

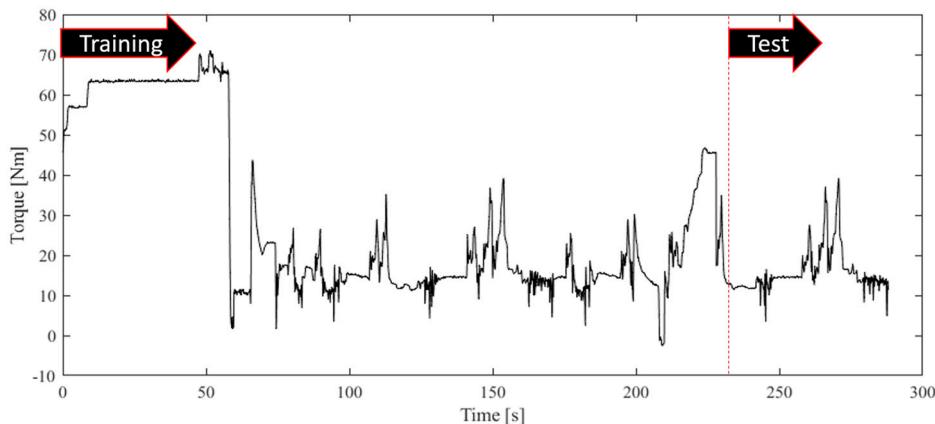


Figure 2. Delivered torque trend during transient operation.

It is feasible to efficiently reduce the dimensions of the model and improve its accuracy by removing parameters with low correlation. As a result, a preliminary analysis using the Shapley value

was done on the complete dataset. SHAP attempts to explain an instance's prediction by assessing the contribution of each attribute to the forecast. The authors were able to quantify the impact of the single measured quantities on the objective function using the average absolute Shapley values (ABSV) [21,22]. The less influential parameters, i.e. TC_Turbine OUT, SparkAdvance and Throttle Position (Figure 3) are excluded by the initial input dataset since presenting the lowest percentage of impact. In this way, the number of input parameters is reduced from 12 to 9.

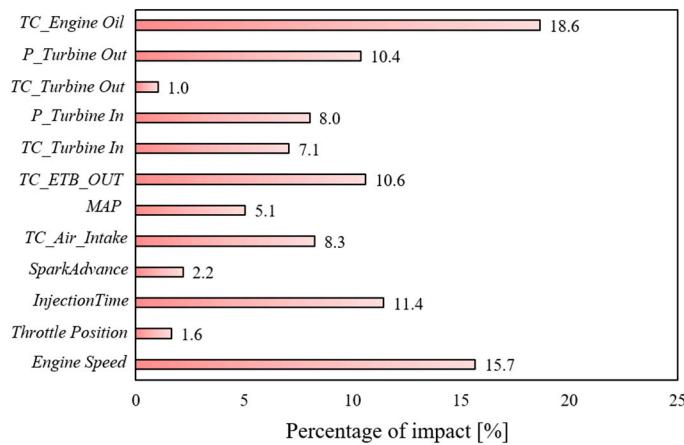


Figure 3. Shapley analysis: global interpretation of the feature's importance for the in-cylinder pressure prediction.

Previous work by the same research group [2,20] shown that when the architectures operate with the removal of the less relevant parameters, the performance improves. Based on this, the current work solely illustrates the architectural predicting performance with the previously established 9 input variables. After identifying the input parameters using the prior analysis, the data is normalized to reduce excessive prediction mistakes and to allow the architecture to converge faster. In this context, the normalization process allows for the avoidance of problems caused by differences in input and output parameters. The values supplied are mapped to the range [0, 1]. Following the prediction procedure, the predicted data is de-normalized to provide a direct comparison with the actual experimentally acquired target. Figure 4 describes the entire dataset used in this activity and the division between input and output parameters for each analyzed case.

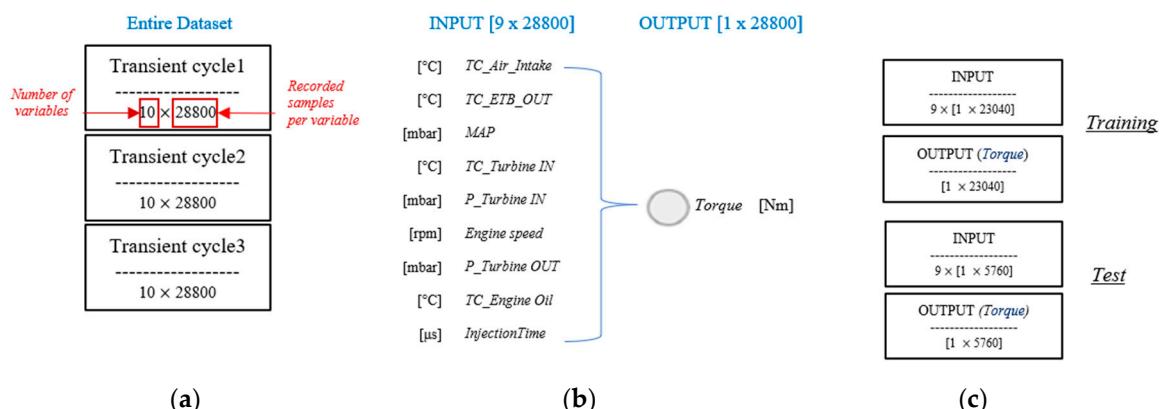


Figure 4. Description of the entire dataset used in this activity; (b) division between input and output parameters for each case analyzed and displayed in (a); (c) data set segmentation for the training and test session.

2.3. Designing of the Neural Architecture to Predict Engine Torque

2.3.1. Structure of the Proposed Model

The predictive scheme of the LSTM+1DCNN structure used in this work for the Torque prediction is reported in Figure 5a.

- A SequenceInputLayer is used to pass the dataset to the network. Such a layer enters the sequence data into the network by setting its size and building the related structures.
- A one-dimensional CNN layer applies a 1-D convolutional filter, comprised of 100 neurons, to each input frame. To perform convolution operations on time series data, the 1D-CNN employs matrix multiplication. It maps the data variables to a higher-dimensional space and finds local features based on spatial and temporal correlations. The convolution kernel of the 1D-CNN moves horizontally or vertically along the data in this process, depending on the nature of the data. The kernel for time series data moves along the time axis, making it excellent for examining sensor data over time [23]. This approach is especially beneficial for analyzing signal data in a short period of time. Because Torque data includes times series recorded by sensors, the CNN can efficiently extract characteristics from such variable data and improve the prediction accuracy of the model.
- ReLu activation function was chosen to improve CNN fitting and sparsity because of its ability to address difficulties such as delayed convergence and gradient disappearance. [24].
- The AveragePoolingLayer calculates the average value for feature map patches and allows for map downsampling by utilizing the mean value in the 2x2 cells square. It uses downsampling to improve computation speed and the durability of derived characteristics. [25].

In the following stage, another 1D convolutional level, similar to the previous one, is used. LSTM is used to process the feature maps at this point. The internal architecture of the LSTM network is made up of components known as gates (Figure 5b) [20]. The LSTM network is a sort of recurrent neural network (RNN) that handles the issue of gradients receding or exploding during long-term information propagation. The LSTM network, unlike typical RNNs, has a more sophisticated neuron structure in the hidden layer. It uses three control mechanisms: the input gate, the forgetting gate, and the output gate to retain long-term knowledge. [26].

LSTMs provide a distinct additive gradient structure with direct access to forget gate activations, allowing the network to encourage desired behavior from the error gradient by employing frequent port updates at each stage of the learning process [20]. Following LSTM, the feature map is distributed in a temporal vectorial sequence by TimeDistributedLayer, and the loss of mean square error for the specified regression issue is computed by RegressionOutputLevel.

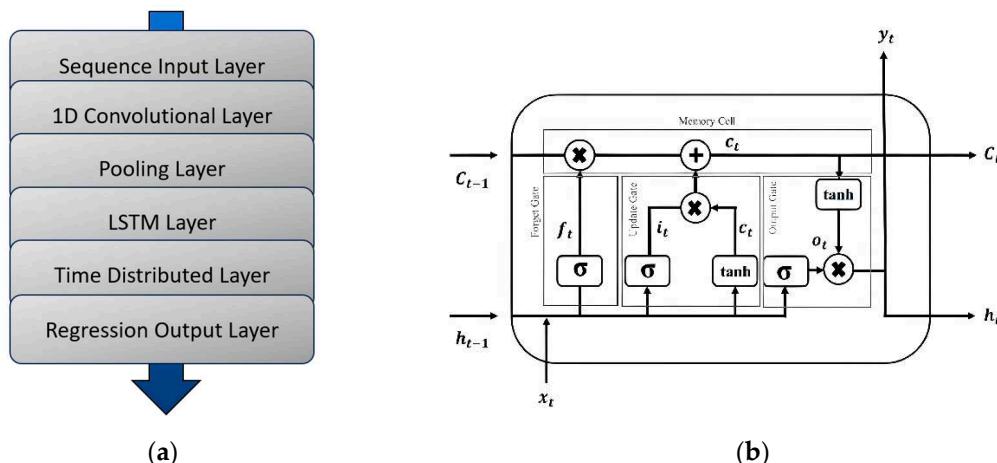


Figure 5. (a) Predictive scheme and (b) LSTM internal structure; subdivision of the structure into gates.

2.3.2. Optimization of the Proposed Model

The definition of the neural structure is determined through preliminary analysis by considering the training sessions' performance:

- The number of neurons in hidden layers N_h varies from 50 to 200.
- The batch size B_s is varied from 10 to 100.
- The model depth M_d is varied from 1 to 10.

A total of 81 combinations are evaluated. The Adam optimizer is used to streamline the updating of the LSTM network model's weight matrix and bias, as well as to adjust the learning rate adaptively during the training process. To evaluate the performance of model parameters, the loss function is created, and the mean square error (MSE) is chosen as the loss function [27]. Set the number of network epoch iterations, for example, to 100, to calculate the final value of the loss function for each prediction model once the network training reaches the maximum learning iteration. To predict the Torque signal, the optimal solution with the lowest loss function value is chosen. For the sake of clarity, Figure 6 shows the val_loss and training_loss of the structure which performed best, i.e., $N_h = 150$, $B_s = 70$, $M_d = 1$. The training results highlight how the model converges without overfitting.

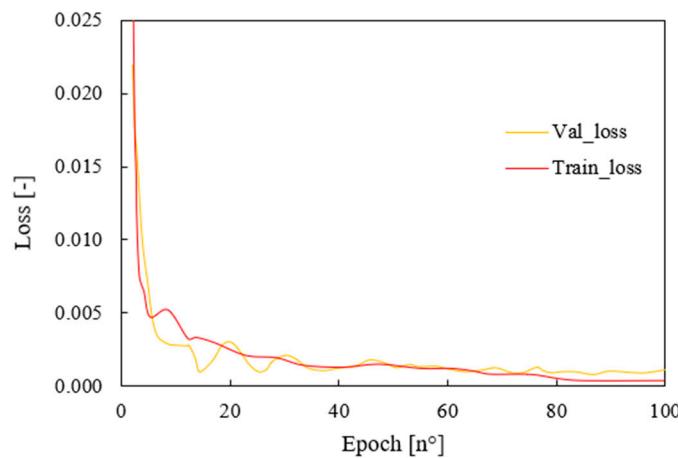


Figure 6. Trend of loss value of the LSTM+1DCNN architecture which performed best during the training session.

To sum up, The LSTM+1DCNN structure is composed of a one-dimensional convolutional layer with 100 neurons, kernel size equal to 3, and ReLu activation function; a max pooling 1D layer which uses a pool size of 2 and a stride of 2; a LSTM layer composed by 150 neurons, batch size equal to 70 and model depth is 1; a time distributed layer and a dense layer composed by 1 unit to perform regression task. The performance of the proposed structures is compared with those deriving from the utilization of other two different architectures, whose optimizations have been performed through extensive preliminary analysis:

1. The structure of a Back Propagation (BP) algorithm [28] is composed of one input layer, three hidden layers, each of which is comprised of 55, 180 and 110 neurons, respectively, and one output layer.
2. A LSTM [20] network prediction model composed of one input layer, one hidden layer containing 150 neurons, respectively, one output layer, and one fully connected layer.

3. Results and Discussion

The first comparison between the proposed algorithm is performed via training_loss function (Figure 7) as described in the previous paragraph. All the structures show a decrease trend as epochs increase and they tend to stabilize around 50 epochs up to reaching a training_loss value lower than 0.001 around the 100th epoch. This certifies that the models converge without overfitting. In particular, LSTM+1DCNN shows the fastest converge speed since showing training loss below 0.005 already at

about 10 epochs. Moreover, once stabilized, it presents very low oscillations suggesting that the model could be more robust than the others.

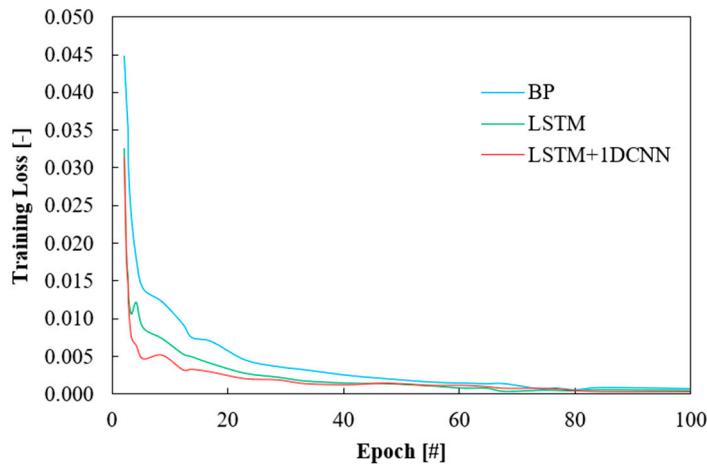


Figure 7. Training loss for the tested architecture during the training session.

Figure 8 displays the prediction of the Torque traces performed by each tested structure. To make a comparison over the entire predicted range, for each forecast, the average deviation of the prediction from the target throughout the range is computed (1):

$$Err = \frac{1}{N} \sum_{i=1}^N \left[\frac{\left| Y_{target}^i - Y_{predicted}^i \right|}{Y_{target}^i} \right] * 100 \quad (1)$$

where N is the number of samples considered for the test case and i the i th sample. The average percentual error, i.e. Err_{avg} , is computed as well to draw attention to the global prediction quality. For this kind of application, a maximum critical threshold of 10 is established for the abovementioned errors.

All the tested structures (Figure 8) are able to reproduce the trend of the Torque over time. Starting from BP, the structure performs an average error of about 2.7%, less than the critical threshold of 10. The number of predictions exceeding such a threshold is 320 samples, corresponding to about 5.55% of the total predictions. The structure is capable of following the low fluctuations of the target signal, while when the torque becomes higher in the range between 250 to 280 seconds, the model underestimates the maximum peaks even if the percentual errors stay below critical thresholds. In such a zone, the architecture predicts in advance the target peaks showing underestimations of about 2 Nm, corresponding to Err higher than 6%. However, it is worth highlighting the structure's capability of following the fluctuation of the signals in such a large range from 22 to 34 Nm. Concerning LSTM, such a model outperforms the BP performance since presenting $Err_{avg} = 1.70\%$ with 67 samples exceeding $Err = 10\%$, corresponding to 1.16% of the total predictions. With respect to BP, LSTM never exceeds Err by 20%. LSTM is capable of following the low fluctuations of the target and, in the range of highest Torque values, i.e. between 250 to 280 seconds, well-reproduces the peaks without any advances or delays. At around 265 seconds (second highest peak zone) the structure underestimates the target value of about 1.5 Nm. According to the LSTM+1DCNN structure, the architecture is capable of well-reproducing the target trend with average percentage errors below 1.5%, i.e. $Err_{avg} = 1.19\%$. The prediction never exceeds 10% of error throughout the entire torque signal and better follow the rapid oscillations of the signals if compared to the other architectures.

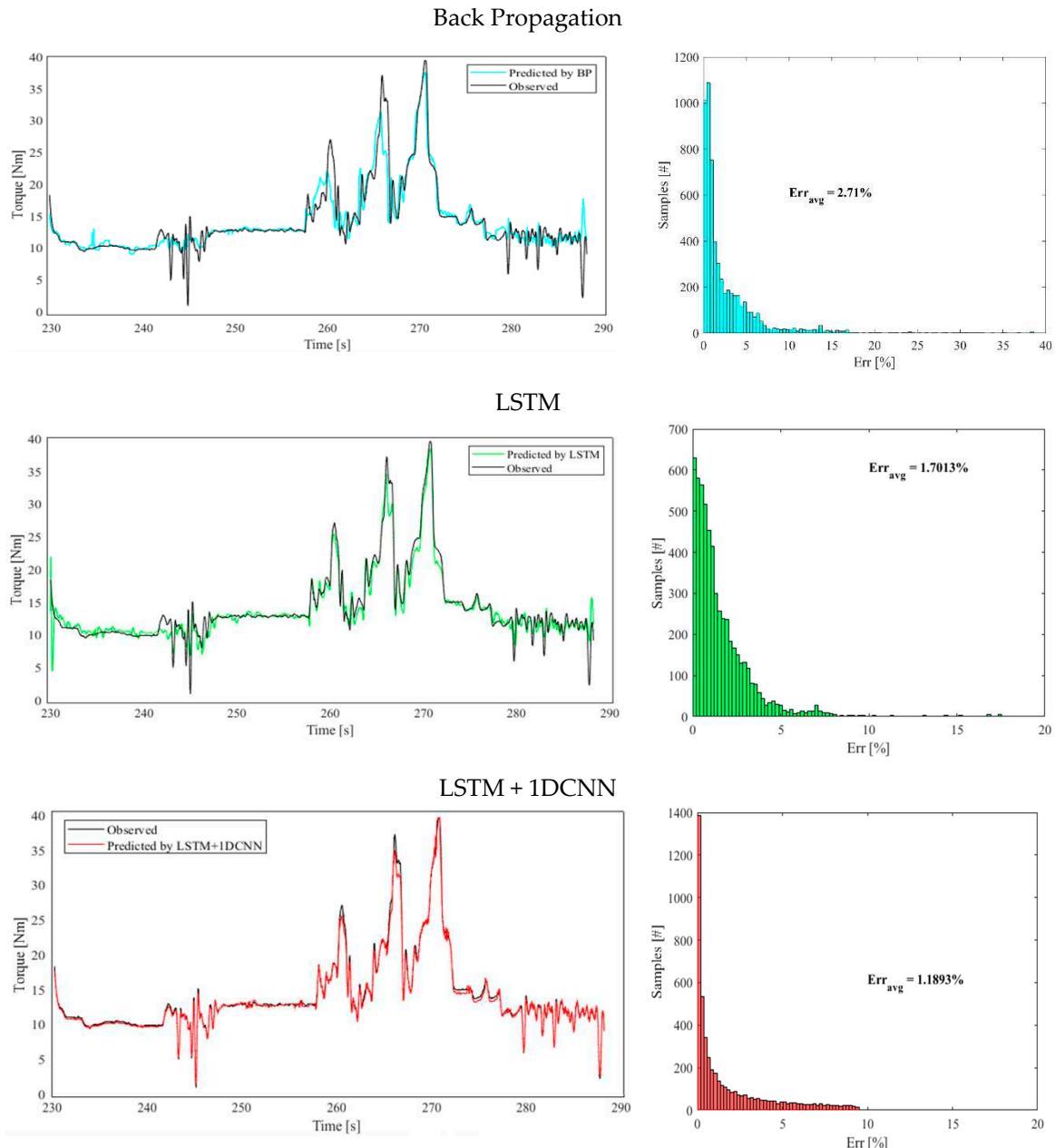


Figure 8. Comparison between the predictions performed by the tested structures and the target Torque with the corresponding percentual errors.

After comparing the prediction performance, and considering the obtained results, the proposed LSTM+1DCNN structure has been tested on another two transient cycles presenting different trends, compared to the case of Figure 2, but the same number of samples (Figure 9). As possible to observe, the structure confirms its capability to reproduce the Torque trend, showing average percentual errors below 1.7% and no predictions above 10% of Err, in both reported cases.

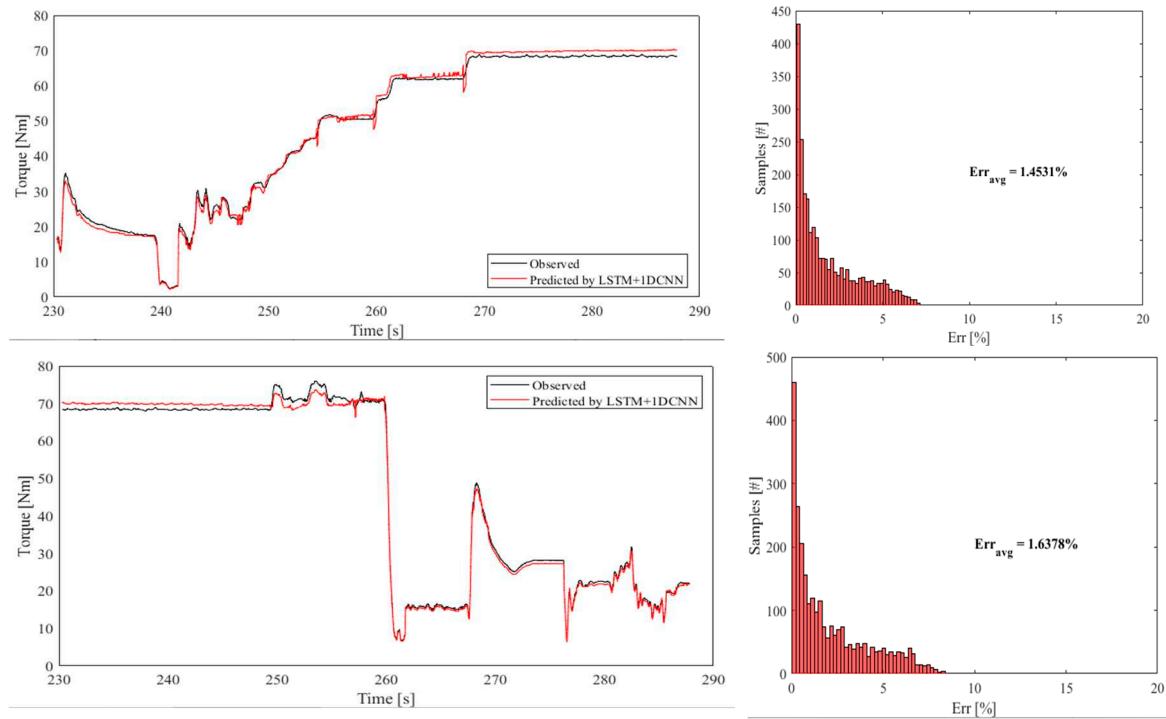


Figure 9. Comparison between the predictions performed by the LSTM+1DCNN structure and the target Torque experimentally obtained by performing two different transient cycles compared to the one used for the preliminary activity.

4. Conclusions

The present work evaluates the possibility of replacing a physical sensor dedicated to computing the torque delivered by an internal combustion engine, under transient condition, by using a LSTM+1DCNN approach. The optimized structure combines the capability of LSTM to capture long-term dependencies and temporal pattern with the ability of 1DCNN to detect patterns within smaller signal segments. The performance of the proposed architecture was compared with the ones of other optimized artificial neural structures, i.e. Back Propagation and LSTM, used for comparative purposes. All the structures proved to be able to reproduce the experimental trend of the engine delivered torque. Specifically, BP model achieved average error of about 3% with 6% of prediction exceeding the critical threshold set at 10%. It accurately reproduced low fluctuations of the signals but underestimated the maximum peaks of the torque. LSTM overperformed BP showing an average error of 1.7% with about 1.6% of predictions exceeding the critical threshold. Even in this case, underestimations of the local maximum peaks were shown. With average errors of less than 1.5% ($\text{Err}_{\text{avg}} = 1.19\%$), the LSTM+1DCNN structure outperforms dealing architectures. It follows torque trends properly, never going over a 10% inaccuracy, and captures fast signal oscillations better. The structure's performance was confirmed further in testing on two additional transient cycles, with average errors maintained at 1.7% and no forecasts above 10% inaccuracy. Overall, the study shows that LSTM+1DCNN can replace physical sensors in torque computation for spark-ignition engines.

Author Contributions: Conceptualization, L.P. and F.R.; methodology, F.R.; software, F.R.; validation, L.P., F.R.; formal analysis, L.P. and F.R.; investigation, F.R.; re-sources, F.M.; data curation, F.R.; writing—original draft preparation, F.R.; writing re-view and editing, L.P.; visualization, L.P.; supervision, F.M.; project administration, F.M.; funding acquisition, F.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available from the corresponding author. The data are not publicly available due to privacy-related choices.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

Err	Percentage Errors
Err _{avg}	Average Percentage Errors
ABSV	Absolute Shapley Values
CNN	Convolutional Neural Network
ECU	Engine Control Unit
ICE	Internal Combustion Engine
ML	Machine Learning
LSTM	Long Short-Term Memory
MON	Motor Octane Number
PFI	Port Fuel Injection
RON	Research Octane Number
SI	Spark Ignition

References

1. K.Y.H. Lim, P. Zheng, C.H. Chen, A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives, *J. Intell. Manuf.* 31 (6) (2020) 1313–1337.
2. Ricci, Federico, Luca Petrucci, and Francesco Mariani. "Using a Machine Learning Approach to Evaluate the NO_x Emissions in a Spark-Ignition Optical Engine." *Information* 14.4 (2023): 224.
3. Reitz, Rolf D., et al. "IJER editorial: The future of the internal combustion engine." *International Journal of Engine Research* 21.1 (2020): 3-10.
4. R. Fiifi, F. Yan, M. Kamal, A. Ali, and J. Hu, "Engineering Science and Technology, an International Journal Artificial neural network applications in the calibration of spark-ignition engines: An overview," *Eng. Sci. Technol. an Int. J.*, vol. 19, no. 3, pp. 1346–1359, 2016
5. L. Petrucci et al. "A Development of a New Image Analysis Technique for Detecting the Flame Front Evolution in Spark Ignition Engine under Lean Condition." *Vehicles* 4.1 (2022): 145-166.
6. L. Petrucci, et al. "Engine knock evaluation using a machine learning approach.", No. 2020-24-0005. SAE Technical Paper, 2020.
7. Srinivasu, Parvathaneni Naga, et al. "Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM." *Sensors* 21.8 (2021): 2852.
8. Ankobe-Ansah, King, and Carrie Michele Hall. "A hybrid physics-based and stochastic neural network model structure for diesel engine combustion events." *Vehicles* 4.1 (2022): 259-296.
9. Krepelka, Michal, and Jiri Vrany. "Synthesizing Vehicle Speed-Related Features with Neural Networks." *Vehicles* 5.3 (2023): 732-743.
10. Fadairo, Adebayo, and Weng Fai Ip. "A Study on Performance Evaluation of Biodiesel from Grape Seed Oil and Its Blends for Diesel Vehicles." *Vehicles* 3.4 (2021): 790-806.
11. C.A. Escobar, R. Morales-Menendez, Machine learning techniques for quality control in high conformance manufacturing environment, *Adv. Mech. Eng.* 10 (2) (2018) 1–16.
12. L. Petrucci et al. "Detecting the Flame Front Evolution in Spark-Ignition Engine under Lean Condition Using the Mask R-CNN Approach." *Vehicles* 4.4 (2022): 978-995.
13. Wang, Ke, et al. "A hybrid deep learning model with 1DCNN-LSTM-Attention networks for short-term traffic flow prediction." *Physica A: Statistical Mechanics and its Applications* 583 (2021): 126293.
14. Elmaz, F.; Eyckerman, R.; Casteels, W.; Latré, S.; Hellinckx, P. CNN-LSTM architecture for predictive indoor temperature mod-568 eling. *Building and Environment*, 2021, 206, 108327. doi:10.1016/j.buildenv.2021.108327
15. Zhao, J.; Deng, F.; Cai, Y.; Chen, J. Long short-term memory - Fully connected (LSTM-FC) neural network for PM2.5 concentration prediction. *Chemosphere*, 2019, 220, 486–492. doi:10.1016/j.chemosphere.2018.12.128
16. Wang, Xin, Dongxing Mao, and Xiaodong Li. "Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network." *Measurement* 173 (2021): 108518.
17. Ren, Lei, et al. "A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life." *IEEE Transactions on Industrial Informatics* 17.5 (2020): 3478-3487.
18. Qin, Chengjin, et al. "Anti-noise diesel engine misfire diagnosis using a multi-scale CNN-LSTM neural network with denoising module." *CAAI Transactions on Intelligence Technology* (2023).

19. Lyu, Zengyi, et al. "A comprehensive investigation of LSTM-CNN deep learning model for fast detection of combustion instability." *Fuel* 303 (2021): 121300.
20. Petrucci, L., Ricci, F., Mariani, F., & Mariani, A. (2022). From real to virtual sensors, an artificial intelligence approach for the industrial phase of end-of-line quality control of GDI pumps. *Measurement*, 199, 111583.
21. S. Tang, et al., Data valuation for medical imaging using Shapley value and application to a large-scale chest X-ray dataset, *Sci. Rep.* 11 343 (1) (2021) 1–9.
22. S. Hart, "Shapley value." *Game Theory*, Palgrave Macmillan, London, 1989, pp. 210–216.
23. Ozcanli, Asiye Kaymaz, and Mustafa Baysal. "Islanding detection in microgrid using deep learning based on 1D CNN and CNN-LSTM networks." *Sustainable Energy, Grids and Networks* 32 (2022): 100839.
24. Bai, Yuhan. "RELU-function and derived function review." *SHS Web of Conferences*. Vol. 144. EDP Sciences, 2022.
25. Zafar, Afia, et al. "A comparison of pooling methods for convolutional neural networks." *Applied Sciences* 12.17 (2022): 8643.
26. Hu, Haowen, et al. "Development and application of an evolutionary deep learning framework of LSTM based on improved grasshopper optimization algorithm for short-term load forecasting." *Journal of Building Engineering* 57 (2022): 104975.
27. Ćalasan, Martin, Shady HE Abdel Aleem, and Ahmed F. Zobaa. "On the root mean square error (RMSE) calculation for parameter estimation of photovoltaic models: A novel exact analytical solution based on Lambert W function." *Energy conversion and management* 210 (2020): 112716.
28. Cui, Yanqing, et al. "Investigation on the ignition delay prediction model of multi-component surrogates based on back propagation (BP) neural network." *Combustion and Flame* 237 (2022): 111852.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.