2021 18th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)
Mexico City, Mexico November 10-12, 2021

Automotive engine fault detection and isolation using LSTM for model-based residual sequence classification

1st Mohammed Youssef
Mechatronics Department
German University in Cairo
Cairo, Egypt
mohamed.solimanyosef@student.guc.edu.eg

2nd Hesham Ibrahim
Associate Professor
German University in Cairo
Cairo, Egypt
hesham.hamed-ahmed@guc.edu.eg

***Abstract*—This paper proposed a machine learning methodology for model-based fault detection and isolation of automotive engines. A hybrid approach is implemented by utilizing both model-based residuals generation alongside a long short-term memory neural network (LSTM NN), which is a structure that can adaptively learn the dynamics of the residuals and perform the fault detection and isolation of automotive engines accordingly. Mathematical models for the combustion engine and its corresponding faults were constructed to synthesize the data required for neural network training. The network was able to fairly classify between different operating conditions with high accuracy, showing the great potential of utilizing LSTM as a robust technique in engine fault diagnosis. This approach proved efficient in terms of eliminating the need for constructing structured residual generators or feature extractors, while giving higher flexibility and less demand for expertise in modelling the system as it is adaptable to both physical-based and data-driven models.**

***Index Terms*—Automotive engine fault diagnosis, hybrid approach for fault diagnosis, Long short-term memory neural networks, recurrent neural network, Mean value engine model, model-based residual generation, sequence classification**

I. Introduction

Development in the automotive industry in the last three decades has been possible as a result of the increasing number of mechatronics components in engines and power-train. Along with this increased complexity, controlling functions and troubleshooting have become more difficult tasks, and so, advanced engine diagnostics have to support the quick recognition of failure causes and reduce downtime. While the supervision of internal combustion engines was originally done through limit checking of a few variables such as oil pressure, coolant temperature, and board-net voltage, the development of analog and digital control over the years opened up new ways for on-board check.

Analytical methods for fault diagnosis have been thoroughly discussed in the literature. Generally, in analytical model-based approaches, either signal models or a mathematical model of the process is utilized. However, they both rely on special models used for residual generation. For signal analysis, appropriate models (i.e., the Autoregressive–moving-average model, frequency spectra, correlation functions, etc.) are chosen from which time or frequency domain features are extracted. This approach works very well when it comes to machine vibrations and can be used to detect imbalances or bearing faults, as well as knocking and surging [1]. While process model-based approaches rely on representing the dependencies between the measured signals using mathematical relations and then residuals, parameter or state estimates are generated as extracted features. Different approaches for fault detection using mathematical models have been developed in the last decades as discussed in Iserman [1]. Features generated from either of the two approaches are then compared against those of the normal process and consequently, analytical symptoms are generated based upon which residual evaluation is done. However, conventional methods results rely directly on the manually designed feature extractors or structured residual generators and their ability to obtain features or residuals vectors that provide isolability and detectability of the faults. This requires great field expertise, and as the process complexity increases, it becomes more difficult and time-consuming.

That's where computational intelligence is found to be powerful. For example, neural network-based adaptive methods have been shown to significantly improve the symptom interpretation in case of malfunctions of mathematically difficult to describe systems and processes. Due to the universal nonlinear function approximation property of neural networks, it makes them a great fit in many health monitoring-related applications such as system identification, feature extraction, pattern recognition, and classification . The ability to recognize patterns of NN is mainly used for feature classification where time or frequency domain properties of the system are mapped to a corresponding fault. For example, In [2] a neural network was used to classify vibration and sound signature signals of worm gears for fault diagnosis. Ahmed, *et al.* [3] used the vibration data in the crank angle domain in conjunction with artificial neural networks for fault detection and classification of an experimental gasoline engine.

However, the abovementioned examples use neural net-

978-1-6654-0029-9/21/$31.00 ©2021 IEEE

works as classifiers or static non-linear approximators, which puts more emphasis on manual feature extraction and constrains the deep learning capabilities of the networks to extract hidden features, and does not exploit their full potential. Furthermore, most research on the fault diagnosis of automotive engines relies on vibrational data extracted from different engine domains, which limits the proposed methods to off-board diagnostics only.

Long short-term memory neural networks (LSTM NNs) are a type of recurrent neural network that can learn both short and long-term time dependencies of data sequences. The ability of LSTMs to process sequential data and their underlying dynamics rather than concurrent data processing in conventional neural network architectures gives the opportunity to feed the network raw sensory data without any feature extraction. LSTMs were used for end-to-end fault diagnosis of the Tennessee Eastman benchmark process in [4]. While this method could be beneficial when little or no field expertise is available, it gives no insight into the underlying process or the results. Xiao,*et al*. [5] proposed a hybrid approach by embedding manual feature extraction methods (e.g., RQA) into deep learning techniques for the fault diagnosis of induction motors. Zhao,*et al*. [6] also proposed a design (i.e., Convolutional Bi-directional Long Short-Term Memory network) and used it for processing raw sensory data to estimate cutting tool wear of a CNC milling machine.

This paper proposes a novel method for the fault diagnosis of automotive engines that utilizes a mathematical model of the engine to generate residual signals, then an LSTM NN is used to detect (identify when a fault has occurred) and isolate (pinpoint the type of fault and its location) faults by processing sequential sampled data points from those residuals.

The major contributions of this paper can be summarized as follows:

1) Taking full advantage of the network capabilities to extract the hidden features and process the underlying dynamics in the residual sequences. vspace1mm
2) Using existing analytical information to build a model for the engine has a number of advantages:
   a) In comparison to other end-to-end approaches, it provides more insight into the learning process and improves the interpretability and dependability of network outcomes.
   b) Residual sequences are more informative than raw sensory data for the network to process.
3) In contrast to conventional methods, where residual generators and feature extractors had to be precisely structured to ensure the detectability and isolability of the faults. This method is adaptable to both physical-based and data-driven modelling techniques, which makes it more flexible and requires less field expertise.
4) The method is suitable for on-board as well as off-board diagnostics as it utilizes the available engine sensors that are essential for controlling purposes.

II. Engine Modeling

The combustion process in reciprocating engines is highly transient (Otto cycle with large and rapid pressure and temperature variations). In addition, the thermodynamic boundary conditions governing the process (intake pressure, the composition of air/fuel mixture, etc.) are not constant. As a result, the phenomena that occur within these processes are usually not accessible for supervision and control purposes. Moreover, the models required to fully describe these phenomena are rather complex and computationally expensive. Therefore, in this work, we will use the Mean value engine model. This model is a control-oriented COM that models the input-output behaviour of the systems with reasonable precision but low computational complexity.

A. Mean value engine model MVEM

Mean value engine models (MVEMs) are dynamic models which describe dynamic engine variable (or state) responses as mean rather than instantaneous values on time scales slightly longer than an engine event. Fast dynamic relationships, i.e., relationships that reach equilibrium in a few engine cycles, are assumed static in this model, while time-developing processes are described by non-linear differential equations. A consequence of this simplification is that phenomena occurring during crank-shaft revolutions do not appear in the model.

B. Engine subsystems

Combustion engines are structures that generate mechanical power through the combustion of air and fuel mixture. Therefore, to model the engine, it's dissected into subsystems.

In this work, the dynamics are modelled as follows:

First, the air dynamics subsystem determines the air mass charge aspirated by the combustion cylinders.
Then, the fuel dynamics and λ-controller subsystems compute the amount of fuel to be injected and that is available for combustion.
Finally, the crank-shaft dynamics subsystem determines the generated power and the rotary speed available at the driving shaft.
Fig.1 shows the model with its subsystems configuration.

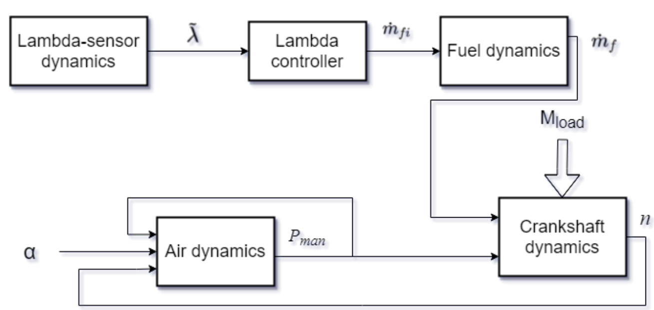

Fig. 1. Block representation of the engine model.

*1) **Air dynamics**:* The intake manifold is an engine part responsible for distributing the air charge that flows through the throttle valve among the engine cylinders. To model the air dynamics of the intake system, the intake manifold is considered as a single structure, called the Air receiver. The air receiver is viewed as a fixed volume for which the thermodynamic states are the same over the entire volume, with its inputs and outputs as the mass and energy flows.

$$\frac{dm_{manifold}}{dt} = \dot{m}_{in}(t) - \dot{m}_{out}(t) \tag{1}$$

Where $m_{manifold}$ is the air mass inside the manifold, $\dot{m}_{in}$ the air mass flow into the manifold and $\dot{m}_{out}$ air mass flow out of the manifold and into the cylinders.

The fluid flow into the air receiver (i.e., the intake manifold) is determined by the equation of fluid flow through orifices (i.e., the throttle valve) whose inputs are the upstream and downstream pressures.

$$\dot{m}_{in}(t) = C_d \cdot A(t) \cdot \frac{P_{amb}(t)}{\sqrt{R \cdot T_{amb}(t)}} \cdot \Psi\left(\frac{P_{amb}(t)}{P_{man}(t)}\right) \tag{2}$$

Where A(t) is the effective area of the throttle plate for a certain throttling angle α calculated based on the plate geometry as in [7], $\Psi(.)$ is the outflow function which depends on the ratio between both the ambient pressure P_{amb} and the pressure inside the intake manifold P_{man}, and the correcting factor C_d.

The flow out of the manifold and into the cylinders is determined by approximating engine cylinders as volumetric pump and then the relation follows.

$$\begin{aligned} \dot{m}_{out} &= \rho_{in}(t) \cdot \dot{V}(t) \\ &= \rho_{in}(t) \cdot \lambda_l\left(P_{man}, \omega_e\right) \cdot \frac{V_d}{N} \cdot \frac{\omega_e}{2\pi} \end{aligned} \tag{3}$$

Where ρ_{in} is the air density at the intake, λ_l is the engine volumetric efficiency, V_d is the cylinders volumetric displacement, N is the number of revolutions per cycle (2 for four-stroke engines).

Modeling the pressure inside the manifold is done using ideal gas law

$$pV = mRT \tag{4}$$

then by substituting from (1) into (4) we obtain

$$\dot{p}_{\text{man}} = \frac{RT_{\text{man}}}{V_{\text{man}}} \frac{1}{3600}\left(\dot{m}_{in} - \dot{m}_{out}\right) \tag{5}$$

Where R is the gas constant, V_{man} is the manifold volume T_{man} is the manifold temperature.

*2) **Crank-shaft dynamics**:* The crank-shaft dynamics considers the power generated through the combustion of the fuel and gas mixture delivered to the cylinders, and accounts for thermal efficiency η_{fc}, friction and inertial losses M_{fric} in the engine and the inertial load at the drive shaft M_{load}. The crank-shaft dynamics is modeled by Newton's 2nd law for rotating masses.

$$M = I\dot{\omega} \tag{6}$$

Where M is the momentum acting upon the rotating mass, ω is angular velocity and I is the moment of inertia. The crank-shaft equation then becomes

$$I_{\text{tot}} \frac{2\pi}{60} \dot{n} = M_{\text{gross}} - M_{\text{fric}} - M_{\text{load}} \tag{7}$$

$$M_{\text{gross}} = \frac{\eta_{fc} Q_{HV} \dot{m}_f \frac{1}{3600}}{n 2\pi \frac{1}{60}} \tag{8}$$

$$M_{\text{fric}} = a_0 + a_1 n + a_2 n^2 + \left(a_3 + a_4 n\right) p_{\text{man}} \tag{9}$$

Where M_{gross} is the torque produced by combustion inside the engine, Q_{hv} is the heating value of the fuel, m_f is the mass of the burned fuel. The coefficients a_0, a_1, a_2, a_3 and a_4 in (9) are determined experimentally.

*3) **Fuel dynamics**:* When the fuel is injected at the intake port it's only partially mixed with the air charge while the rest becomes a puddle at the engine port creating a fuel film. The fuel film is evaporated off the heated intake manifold with a time constant τ_{ff}. The fuel dynamics subsystem models this effect through the following equations :

$$\ddot{m}_{ff} = \frac{1}{\tau_{ff}}\left(-\dot{m}_{ff} + X\dot{m}_{fi}\right) \tag{10}$$

$$\dot{m}_f = (1 - X)\dot{m}_{fi} + \dot{m}_{ff} \tag{11}$$

Where m_{ff} is the evaporated fuel mass, m_{fi} is the mass of the injected fuel.

*4) **Lambda controller**:* The output of the lambda sensor is used by the lambda controller to determine the adequate amount of fuel to be injected to reach a stoichiometric ratio of λ=1.

$$\dot{m}_f = \frac{\dot{m}_{aircharge}}{14.67} \tag{12}$$

$$\begin{aligned} \dot{\tilde{\lambda}}(t) &= \frac{1}{\tau_\lambda}\left(-\tilde{\lambda}(t) + \lambda\left(t - \tau_d\right)\right) = \\ &= \frac{1}{\tau_\lambda}\left(-\tilde{\lambda}(t) + \frac{\frac{\dot{m}_{ac}(t-\tau_d)}{\dot{m}_f(t-\tau_d)}}{14.67}\right) \end{aligned} \tag{13}$$

The previous equations represent the lambda sensor dynamics where λ is the real air fuel ratio, $\tilde{\lambda}$ is the output of the sensor dynamics, τ_λ is the time delay of the oxygen sensor. The time delay τ_d in the previous equation exists due to the Significant time-delay between the inputs to the engine and the oxygen sensor output (e.g., transportation delay in manifold and pipes).

III. FAULT MODELING

A fault is defined as the state that might lead to system failure as a result of an unpermitted deviation of a process variable from an acceptable range. In this work, two blocks (i.e., fault generators) were designed to induce a faulty behavior in the pre-constructed engine model.

Two of the engine common faults are modeled.

1) Air mass leakage into the intake manifold.
2) Manifold pressure sensor failure.

A. Air mass leakage

Vacuum leakage through the manifold body or the connected hoses is a common problem that leads to a rich air mixture as air leaks into the manifold body causing an increase in the total air mass inside the combustion chambers. The air mass leakage into the manifold is modeled as an additive disturbance entering the system dynamics the same way as the actuator signal, as shown in Fig. 2 (a).

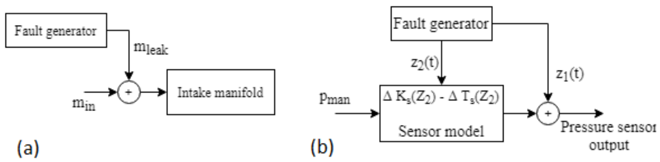

Fig. 2. Block representation of the fault models. (a) Air mass leak model, (b) Pressure sensor disturbance model.

The mass leak fault generator is fed the equivalent leakage diameter, based upon which an estimate of the leakage source area is computed and consequently the leakage air mass using the equations of fluid flow throw orifices. Additionally, the fault generators simulate different time-varying behavior for the fault. For example, the fault could be abrupt (sudden change) or a gradually increasing one (drift-like). Fig. 3 shows different behaviours of the mass leakage fault.

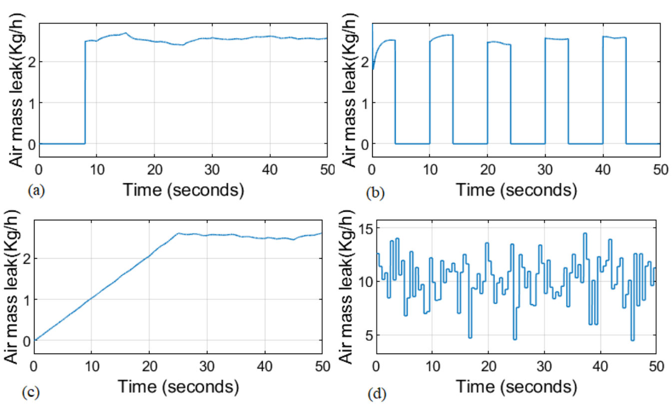

Fig. 3. Different forms of the mass leak model output. (a) abrupt, (b) intermittent, (c) drift-like, (d) stochastic signal.

B. Sensor disturbances

The sensor readings can be influenced by External disturbances which can be in the form of a superimposing effect (e.g., electromagnetic influences) from the surrounding environment represented mathematically by the time varying variable z_1(t) affecting the sensor static behaviour as follows.

$$\begin{aligned} Y(t) &= c_1\left[y_0(t) + z_1(t)\right] + c_0 \\ \Delta Y(t) &= c_1 z_1(t) \end{aligned} \tag{14}$$

Additionally, deformations could occur in the sensor due to an increase in temperature, fluid flow velocity or contamination altering the sensor dynamic parameters. As the dynamic behavior of the sensor is frequently modeled as a first-order lag.

$$G_s(s) = \frac{K_s}{1 + T_1 s} \tag{15}$$

Where T_s and K_s are the sensor time constant and the sensor gain respectively.
These deformations are represented by the time varying variable z_2(t), which causes a deviation in the sensor dynamic parameters (i.e., K_s and T_s) as shown in (16).

$$\begin{aligned} \left[T_1 + \Delta T_1(z_2)\right]\dot{y}(t) + y(t) &= \left[K_s + \Delta K_s(z_2)\right] y_0(t) \\ \Delta y(t) &= -\Delta T_1(z_2)\,\dot{y}(t) + \Delta K_s(z_2)\, y_0(t) \end{aligned} \tag{16}$$

The pressure sensor fault generator was designed to manipulate the variables $z_1(t)$ and $z_2(t)$ also with different intensities and time dependencies to simulate the faulty behavior of the sensor.

C. Stochastic fault model

Additionally, each of the two previously discussed faults could take the form of a random disturbance using a stochastic signal of different parameters (i.e., mean, frequency, and variance) as shown in Fig. 3 (d).

Fig. 4 shows the overall structure of the fault generators.

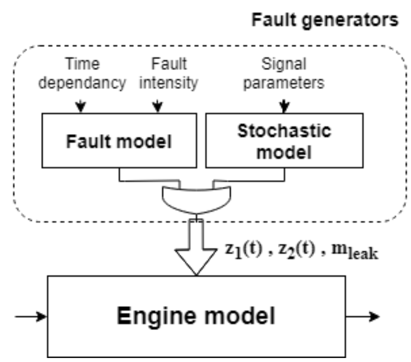

Fig. 4. Fault generators structure.

IV. TRAINING AND VALIDATION DATA

The objective is to train the neural network to detect and isolate faults based on a sequence of sampled data points from the residual signals.

To explore the limits of neural networks, their generalization capabilities, and robustness to fault-modeling uncertainties. The neural network was trained on a data set generated using the stochastic model of the faults and validated against a more structured data set where faults were generated using the corresponding fault model with a specific time dependency and intensity.

A. Training data set

The aim was to train the network to detect generic random disturbances influencing certain parts of the system dynamics. The stochastic fault model was used to produce training data by generating faults with no defined magnitude or time-related behaviour. The faulty variables (i.e., z_1(t), z_2(t) and the leaking air mass) were fed into the model as a stochastic signal, where the signal features (i.e., frequency, mean, and variance) were given sets of different values. The training data set was then generated by simulating each possible combination of these features and logging the relevant data.

B. Validation data set

The previously discussed fault models were used to induce more structured faulty data. Different leakage effective diameters were used with altering time dependencies to generate the validation faulty conditions of the mass leakage. Similarly, for the pressure sensor faulty conditions, different values for the sensor gain and time constants were used in addition to additive disturbances.

V. Results and Discussion

Two models were run in parallel, one representing the running engine where the faults are induced and the other representing our estimate of the engine's normal behavior. Residual signals are generated from 4 selected signals (rpm, manifold pressure, cylinder air mass flow, fuel mass flow) based on which the classification will occur.

Healthy simulation conditions are generated by varying the load torque on the engine and torque demand by the driver represented by varying throttle valve opening. The mean value of both load torque and throttle angle change to obtain different engine operating zones. Fig. 5 shows the load torque and throttle valve angles for a complete simulation run with a mean value of 60 N.m and 30 deg of the load torque and throttle valve angle respectively.

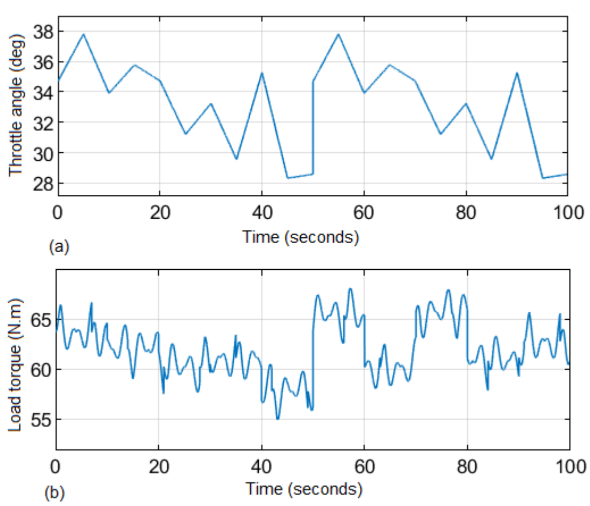

Fig. 5. (a) Throttle valve angles (b) The torque load exerted on the engine during a simulation run.

For the training mass leak and sensor failure data, the stochastic signal parameters (i.e., mean, variance, and frequency) were given sets of different values. Table I contains the values of the signal parameters used for mass leakage, z_1 & z_2 signals, respectively. Finally, simulation conditions were generated for each combination of the parameters, resulting in 24 simulations total. A similar process was done for validation

TABLE I
PARAMETERS OF THE STOCHASTIC MASS LEAK, z_1 & z_2 SIGNALS USED FOR TRAINING DATA

Mean	2.5	10	25
Frequency	0.5	1	-
Variance	0.01	0.05	-

Mean	0.01	0.03	0.05
Frequency	0.5	1	-
Variance	0.01	0.05	-

data by specifying the effective leak diameter, z_1 and z_2 as shown in Table II with different time dependencies generating a total of 26 simulation runs.

TABLE II
SIMULATION CONDITIONS FOR VALIDATION DATA.

effective leak diameter	5 mm	10 mm	15 mmm
z_1 & z_2	0.01	0.03	0.05

Table III contains the total number of simulation runs generated for each operating condition. The simulation conditions

TABLE III
NUMBER OF SIMULATION RUNS FOR EACH OPERATING CONDITION.

Conditions	Healthy	Mass leak	Sensor failure
Training	5	12	12
Validation	2	13	13

are collected and then run for 100 seconds each with a sampling time of 0.005. That yielded a total of 20000 data points for each logged signal for a single simulation run. The logged data was normalized, pre-processed, and labeled. A total of 630 four-dimensional input residual vectors were generated for training and their corresponding labels, fig. 6 shows the pressure sensor residuals generated from the validation simulation runs.

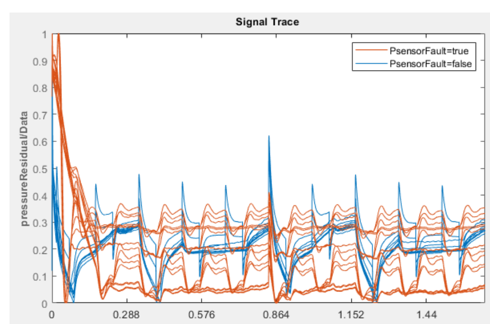

Fig. 6. The generated pressure sensor residuals used for validation.

Based on the duration of the induced disturbances and the engine response, a 1000-point sequence, which corresponds to 5 seconds of operation, was found to be optimal for training and classification. The sequence length was chosen to be of minimal (i.e., faster detection) length while ensuring that they carry an informative amount of data for the network to process and perform proper classifications.

Because the validation data was not chosen to be a subset of the training data, but rather both datasets were generated with different approaches. It was critical to choose the network structure such that it would not overfit the training data and be able to generalize to both the validation and training datasets.

Different network structures were tested, choosing the number of LSTM layers and neurons between one, two, or three layers with 50 to 10 neurons each. It was observed that 2-layered networks show faster convergence (learn in less number of iterations) when it comes to the training data-set while showing poor results in classifying the validation data. On the other hand, 3-layered networks have slower learning rates but they show better generalization properties when it comes to learning both the training and validation data. The final network was chosen to be a 3-layer network, with a single 20-neuron layer and two 10-neuron hidden layers. The classification results were outstanding, with an accuracy of 96% for both the training data set and the validation data. Fig.7 shows the classification results of the network for the unseen validation data, which shows an ideal performance for isolating pressure sensor and mass leakage faults. However, it shows a great tendency to classify healthy operations as a mass leak. Modeling the mass leak as an additive disturbance, especially, with small intensities (e.g., leak diameter of 5 mm) leads to a great similarity between the healthy and mass leak residuals as shown in Fig.8 especially after normalization. This could be avoided either by neglecting the mass leak of small intensities as it results in an insignificant air mass (i.e., less than 1% of the normal air mass flow at normal operating conditions) or by using a more accurate dynamic model of the mass leak could result in better accuracy.

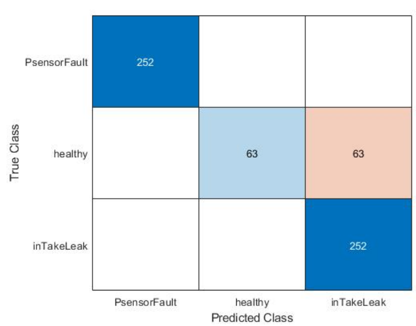

Fig. 7. Confusion matrix of the network classification for validation data.

VI. Conclusion

In this paper, a new hybrid approach was taken in an attempt to detect and isolate the faults of an automotive engine model where a model-based residual generation along with sequence classification using Long short-term memory neural network for residual evaluation is used.

A mean value engine model of the relevant subsystems was constructed to synthesize the training and validation data for the neural network to train on. The neural network was trained to classify between two different faults and a healthy operation based on a four-dimensional residual vector. Two different approaches were taken for generating training and validation data to examine the network's generalization capabilities and how the network would deal with new unseen data of a different nature. The long short-term memory neural network showed great results in detecting and classifying the faults with an accuracy of 96%. The results indicate a great potential of LSTMs in fault diagnosis; it overcomes the need to construct structured residual generators or feature extractors. Also, it gives higher flexibility and less demand for expertise in modelling the system, while taking advantage of any prior knowledge about the system to facilitate and ensure reliable results from the network classification. Additionally, the ability of the network to classify unseen data could imply the ability to synthesize the training data without downgrading performance, which could be huge when dealing with valuable components where failure data could be costly.

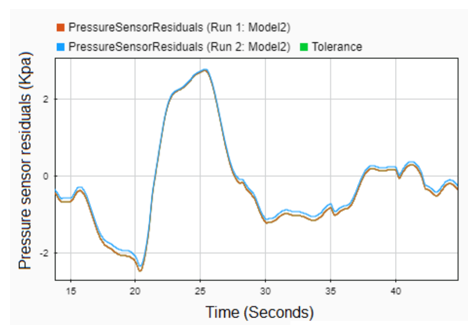

Fig. 8. Pressure sensor residuals for healthy and 5mm mass leak with abrupt behavior plotted over each other.

References

[1] Isermann R. Combustion engine diagnosis. Springer; 2017.

[2] Waqar T, Demetgul M. Thermal analysis MLP neural network based fault diagnosis on worm gears. Measurement. 2016 May 1;86:56-66.

[3] Ahmed R, El Sayed M, Gadsden SA, Tjong J, Habibi S. Automotive internal-combustion-engine fault detection and classification using artificial neural network techniques. IEEE Transactions on vehicular technology. 2014 Apr 21;64(1):21-33.

[4] H. Zhao, S. Sun and B. Jin, "Sequential Fault Diagnosis Based on LSTM Neural Network," in IEEE Access, vol. 6, pp. 12929-12939, 2018.

[5] Xiao D, Huang Y, Qin C, Shi H, Li Y. Fault diagnosis of induction motors using recurrence quantification analysis and LSTM with weighted BN. Shock and Vibration. 2019 Jan 27;2019.

[6] Zhao R, Yan R, Wang J, Mao K. Learning to monitor machine health with convolutional bi-directional LSTM networks. Sensors. 2017 Feb;17(2):273.

[7] Isermann R. Engine modeling and control. Berlin: Springers Berlin Heidelberg. 2014;1017.