Chapter 5

SRF Diagnosis using TS imaging with FRP and Classifier Fusion

5.1 Introduction

Much attempt hasn't been made to utilize the TS property of input vibration signals in the literature of SRF diagnosis. Therefore, there have been essentially zero attempts to disclose or quantify nontrivial dynamical features in the vibration signals. It is true that a range of techniques has been used for the representation of the input data in order to make it compactable to the applying models, the majority of them are 1D, or 2D structures of the extracted features [10]. TS imaging enables the learning models like CNN to visually identify and classify the vibration signals. But most of the existing image representations from the vibration signals fail to utilize the TS property efficiently, in particular, when 2D data models are made by converting raw data. For example, the thermal images [28], the 2D grayscale texture of vibration [29], orbital images [118] etc., disregards TS property totally. The position-based polar image [122] and the infrared (IR) imaging [120, 136] are generally run using CNN, do not succeed in the maintenance of the data's temporal properties, whereas symmetrized dot pattern images are used for preserving the input data's sequential property to a

level. Recently a few new imaging techniques have been used for vibration data such as shaft orbit images [119], 2D Kurtograms [215], and the omnidirectional regeneration method [45]. It is noteworthy that there are broadly two approaches to generate images from vibration data. The first one is the way of stacking raw data directly in the 2D form, and the other way is the 2D arrangement of the data in a transformed domain, both of which fail to keep the TS property of the original signal. In other words, the TS imaging technique in input data representation lacks in the literature.

The recurrent plot, Gramian angular field, and Markov transition field are the popular image encodings that integrate TS properties of the original data. RP extracts trajectories from TS and computes the pairwise distances between these trajectories. The GAF images are characterized as a Gramian matrix by using a polar coordinate system such that every element is the trigonometric sum between time points. In MTF images, the first-order Markov transition probabilities are represented in one dimension and the temporal dependencies in the other dimension. The RPs are established to be a definitive representation used to quantify and reveal nontrivial dynamical features present in TS data when a comparison is drawn between the conventional imaging schemes discussed in the literature. Hence, RP has proven to be one of the precursors of TS imaging. The preliminary experiments with the three imaging schemes revealed that RP is the top performer with the SRF dataset, and hence we adopted RP as the basic imaging scheme in this regard. A fuzzy RP exhibits all the properties of RP along with some extra advantages with the incorporation of fuzzy logic. Furthermore, it is can possible to derive from the literature that fault diagnosis relies on a single information source and a distinct, intelligent approach showcases certain drawbacks.

Therefore, we build up a framework that uses two decision-making strategies parallel to each other, i.e., the LSTM network taking DFC stream as input and the CNN with taking FRP as input. Furthermore, the most informative sensor signals from the segmented raw vibration signals are proposed to be selected by a DFC-based ranking.

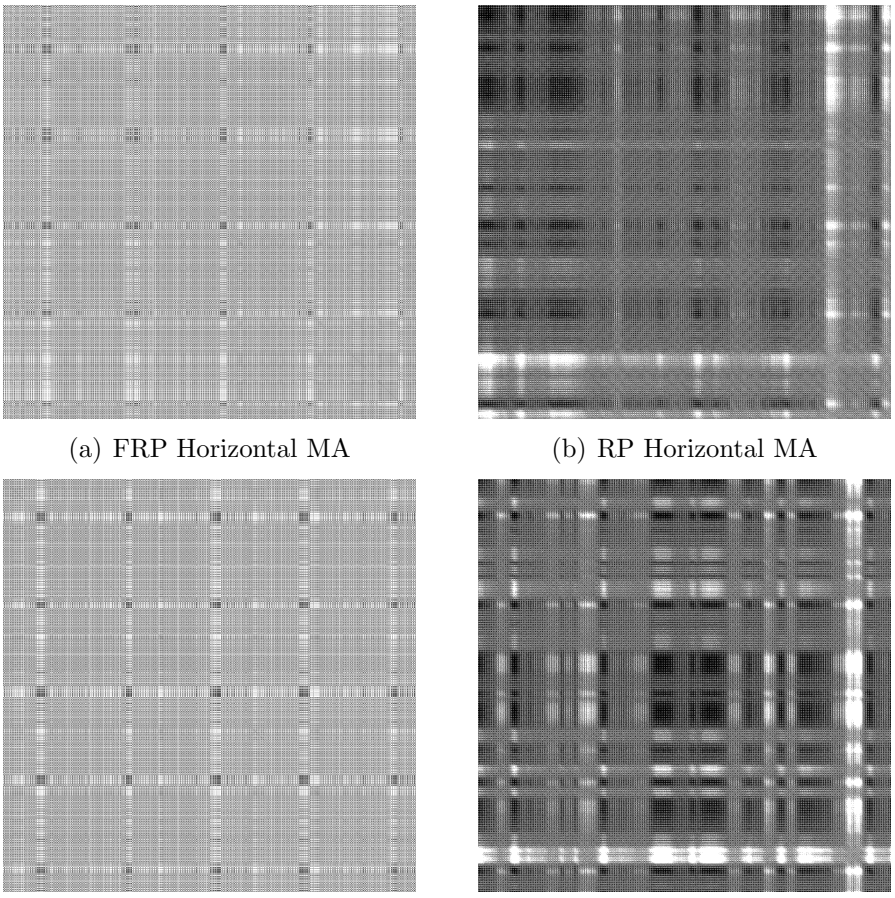
The image combining scheme generates a 3D-FRP and 2D flattened FRP (F-FRP). The system dynamics are characterized by these multi-sensor FRPs which also conserve the vibration input's TS properties. Symptom parameters-based decision-making is facilitated by the DFC stream, which is simultaneously catered to LSTM. In the end, we made use of classifier fusion in SRF diagnosis. The parallel decision scores here are fused by fuzzy integral (FI) grounded on fusion depending on the confidence of the sources.

5.2 Theoretical Background

5.2.1 Recurrence plot

The concept of RP was first put forward by Eckmann et al. [216]. It provides the service of a visualization tool that displays a dynamical systems' time constancy which has been characterized by large-scale topology and small-scale texture. A set of all possible states in a dynamic system where each point stands for one state is called phase space. A visual image of the system's dynamics is provided by the reconstruction of phase-space in different dimensions. The instances when a phase-space trajectory visits the same area in the phase-space are shown by RP. Embedding is a process that is used to transform a TS into an object in space. Embedding dimension is the number of state parameters required to embed an object or the phase space dimension. According to Takens' embedding theorem [217] a single TS can be used to yield the observed state when an appropriate value of time delay and embedding dimensions are provided. We make use of this property to obtain a visual image of the dynamic property of the system with different sensor values in an independent manner. After the work of Casdagli [218], RP's application for turning to the main properties of nonstationary TS became popular.

Suppose for set of state vectors, let $S = \{s_1, s_2, ...s_n\}$ is the phase-space reconstruc-



(a) FRP Horizontal MA (b) RP Horizontal MA

(c) FRP Unbalance (d) RP Unbalance

Figure 5.1: FRP and RP patterns

tion of the dynamic system with embedding dimension k, where each state $s_i \in R^k$, then RP is represented as an $n \times n$ matrix. A threshold ϵ defines the similarity of a state pair (s_i, s_j) with the following expression.

$$\mathbf{R}(i,j) = \theta\left(\epsilon - \|s_i - s_j\|\right) \tag{5.1}$$

Here $\theta(\cdot)$ defines a step function that assigns a 1 to $R(i,j)$, if $\epsilon - \|s_i - s_j\| \geq 0$ and 0 otherwise.

5.2.2 Fuzzy recurrence plot

The FRP is a variation of RP that incorporates the fuzziness in determining the similarity of state pairs in RP [219]. A fuzzy membership function $\mu \in [0, 1]$ that has reflexivity, symmetry, and transitivity properties characterizes the fuzzy grade of membership. For the purpose of determination of the state pair similarity, fuzzy clusters of phase-space states are formed with the popular fuzzy c-means clustering (FCM) algorithm [220] in FRP. The distance between the data point and cluster head is first determined, and then the membership assignment to each data point is performed by the algorithm. The fuzzy cluster set $C = \{c_1, c_2, ...c_v\}$ is generated by using the FCM algorithm after defining the number of cluster v. The cluster partition is determined by the degree of belongingness of a state to a particular cluster which is represented by the fuzzy membership grade. Hence the FRP is defined as:

$$\mathbf{R}_f(i, j) = \mu\left(s_i, s_j\right), \tag{5.2}$$

where $i, j = 1, ..., N$ and $\mu\left(s_i, s_j\right) \in [0, 1]$ shows the fuzzy membership similarity of state pairs s_i, and s_j. Here the fuzzy membership grades are updated iteratively by the following equation:

$$\mu_{ir} = 1/\sum_{p=1}^{v} \left(\partial_{ri}/\partial_{rp}\right)^{2/(m-1)} \tag{5.3}$$

where m is the fuzziness index such that $m \in [1, \infty]$, ∂_{ri} represents the Euclidean distance between the data point s_r and cluster center c_i, and μ_{ir} represents the membership of r^{th} data point s_r to the i^{th} cluster center c_i. Along with the updated membership grade μ_{ir}, the cluster center in each iteration is updated by:

$$c_i = \left(\sum_{r=1}^{N} \left(\mu_{ir}\right)^m s_r\right) / \left(\sum_{r=1}^{N} \left(\mu_{ir}\right)^m\right) \tag{5.4}$$

$\forall i = 1, 2, ..., v$, with an objective of minimizing

$$\sum_{r=1}^{N} \sum_{i=1}^{v} (\mu_{ir})^m \|s_r - c_i\|^2. \tag{5.5}$$

5.2.3 Fuzzy integral based classifier fusion

Successful application of FI has been achieved in machine fault diagnosis problem in the work of Liu et al. [221]. FI regards consequential parameters called fuzzy measures together with the decision scores of the classifier. It is a weightage given priorly for every confidence value for fusing the classifiers. Let $O = \{O_1, O_2, O_3, ...O_N\}$ be the set of output scores from the N classifiers. Then the fuzzy measures $g : 2^O \rightarrow [0, 1]$ are defined for all the possible combinations of O. It has the properties $g(\phi) = 0$ and $g(O) = 1$. For any $o \subseteq O$, the fuzzy measure $g(o)$ has got the monotonic property as follows. If $o_i \subseteq o_j \subseteq O$ then $g(o_i) \leq g(o_j) \leq 1$. The concept of Sugeno- λ measure was introduced in [222], which adds the property as if $o_i \subseteq O$, $o_j \subseteq O$ s.t. $o_i \cap o_j = \Phi$, then

$$g_\lambda(o_i \cup o_j) = g_\lambda(o_i) + g_\lambda(o_j) + \lambda . g_\lambda(o_i) . g_\lambda(o_j) \tag{5.6}$$

According to this, we can solve λ by

$$\lambda + 1 = \prod_{i=1}^{N} (1 + \lambda g^i) \tag{5.7}$$

where g^i means $g_\lambda(o_i)$. In 1953, the French mathematician Gustave Choquet introduced the Choquet integral [223]. Each and every linear combination of empirical fusion strategies can be realized on the basis of the choice of fuzzy measures [224]. By utilizing the fuzzy measures, the Choquet integral is able to assign importance to all possible groups of criteria, which offers far greater flexibility for aggregation. The normal integration operator is given by :

$$C_g(O) = \sum_{i=1}^{N} O_{\Pi_i}[g(o_{\Pi_i}) - g(o_{\Pi_{i-1}})] \tag{5.8}$$

where the output scores O is permuted to O_Π such that $O_{\Pi_1} \geq O_{\Pi_2} \geq, ..., \geq O_{\Pi_N}$. The Choquet integral finds the score by combining the fuzzy measure for a particular score along with the confidence of the score, called its support. As shown in Eq. 5.8, the intermediate decisions determine the ultimate decision-making process that enables sensitivity to each classifier.

5.3 Proposed Method

The proposed framework employs the classification ability of CNN and LSTM in parallel to yield the best classification result. Initially, the measured vibration signal has been segmented with sufficient length to ensure proper extraction of DFC as well as to create a TS image, i.e., FRP in our model. The RPs are established to be a definitive representation used to quantify and reveal nontrivial dynamical features present in TS data when a comparison is drawn between the conventional imaging schemes discussed in the literature. Hence, RP has proven to be one of the precursors of TS imaging. All the properties of RP are displayed by FRP with some extra advantages by integrating fuzzy logic. A temporal ordering to hasten sequential learning is preserved when the DFC is extracted from each segment by making use of overlapped bins. The two representations bring out two characteristics of input data, such as the temporal property being kept by the FRP and non-linear dynamical characteristics by the recurrence pattern. At the same time, DFC is proven to be the final parameter for decision-making by its symptomatic frequency components. We propose two image combining schemes for input generation and a DFC fault pattern-based FRP selection because we possess multiple sensor values for FRP generation. The top two most informative FRPs are flattened and stacked by our scheme to produce a single image. The other scheme

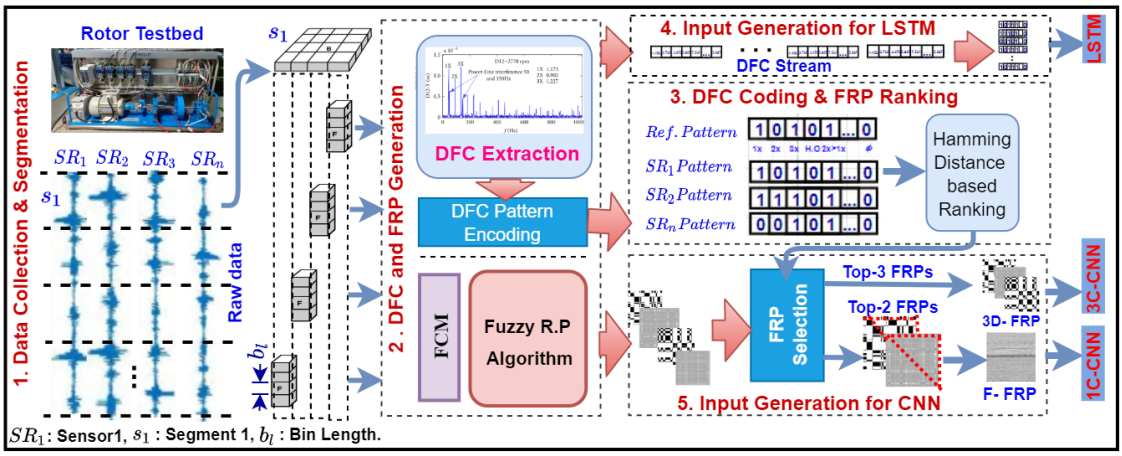


Figure 5.2: Input data preparation phase

forms a 3D-FRP by assembling the top three FRPs. The first scheme alters the temporal correlation of recurrence structure, but all the informative data points are included in it. In contrast, the second method is used to combine multiple recurrence structures simultaneously in a single model by keeping the temporal correlation. Therefore, two different aspects of recurrence plots in SRF decision-making are brought out by the two imaging schemes.

At the decision-making phase, a single-channel CNN (1C-CNN) and a three-channel CNN (3C-CNN) are fed with these two types of combined FRPs. The LSTM makes decisions corresponding to the DFC stream, drawn from the bins of the segment. Therefore, three sets of decision scores are yielded for the same segment of the input signal. Each of these sets carries with it different properties. In the end, a fuzzy integral-based fusion module carries out the classifier fusion based on the fuzzy measure weightage that is provided to these three decision values. Because Choquet fuzzy integral works as the generalization of the empirical schemes, it was followed in this process [225]. Based on practical expertise, the fuzzy measures are given and assigned to the classifiers. The FI fusion then modifies the classifier weightage based on previous decision scores.

The overall operation is described in three phases: input data preparation, decision score generation, and classifier fusion.

5.3.1 Input data preparation phase

A critical task of the framework is to create two representation categories for the same data segment for both channels simultaneously. To make sure that FRP and DFC are generated properly, it demands preserving enough data points in every segment. The overall operations in this phase can be seen in Fig. 5.2.

Table 5.1: Binary coded DFC

SRF	Amplitude					Phase		
	1x	2x	3x	H.O	2x>1x	0°	180°	Var.
Healthy/Norm	0	0	0	0	0	1	0	0
Static UB	1	0	0	0	0	1(R)	0	0
Couple UB	1	0	0	0	0	0	1(A)	0
Dynamic UB	1	0	0	0	0	0	0	1
Vertical MA	1	1	0	0	1	0	1(R)	0
Horizontal MA	1	1	1	0	1	0	1(A)	0
Looseness	1	1	0	1	0	0	0	1

H.O: higher Order, Var.: Varying, R: Radial, A: Axial

As shown in the first two parts of Fig. 5.2, the raw signals from the sensors are divided at first into equal-length segments to create the FRP and to extract the DFC stream. The segment length (S_l) is selected such that each segment contains the required number of bins to facilitate the necessary amount of DFC values. Likewise, sufficient data points must be considered to represent the dynamics of the system through FRP generation. Furthermore, a certain number of raw data points in each bin is required for the legitimate DFC extraction. The segmentation and DFC extraction follow the same process explained in section 3.4. Let b_l be the minimum number of data points in each bin, S_r be the number of sampling point in each rotation, and S_i be the number of interval points between the bins (in non-overlapping partitions, it is set to zero), then the condition: $S_i \leq S_r \leq b_l$ ensures proper DFC extraction from each bin. It ensures that every bin accommodates at least one rotation data. Likewise, every segment must be sufficiently lengthy ($S_l >> b_l$), as well as it considers the additional condition of

minimum points for FRP as 50 to 100 [226].

The initial five harmonics of the rotating frequency (1x to 5x) are utilized to derive the amplitude as well as phase after the DFC extraction, and they are utilized for DFC-based ranking (refer Section 5.3.1.2). The DFC stream is then created by using the first three of them (1x to 3x). Two main operations take place after this step, such as generation of DFC pattern and selection of FRP based on DFC pattern matching.

5.3.1.1 DFC pattern generation

A DFC encoding scheme is proposed that generates a binary pattern based on the SRF characteristics. The amplitude and phase characteristics of the symptomatic frequencies of SRF are have been referred from the Table 3.3. The resulting 8-bits binary codes of SRF are given in the Table 5.1. The encoding helps compare the current bin's DFC pattern with that of the referred fault with fast match score generation. The SRF is comparatively more sensitive to the lower band rotational frequencies, which is why whether the 1x to 3x amplitudes are high or low is represented by the first three bits of the pattern. The relativity between 1x and 2x has been represented by bit 5, whereas amplitudes related to all other higher-order harmonics are decoded to a single bit, i.e., bit 4. The three bits, at last, are designated for phase values where bit 6 and bit 7 represent the phase shift of 0° and 180°, respectively. Here, the final bit is used for representing the varying phase value between 0° and 180°. This encoding requires slight changes, suggests the sensor mounting position (radial, axial or tangential). At the time of encoding a new set of DFC, a threshold of the amplitude values decides the high and the low value, which is selected from the mean of all extracted amplitudes.

5.3.1.2 FRP selection based on DFC pattern matching

The complexity of the framework is reduced by selecting the most informative sensor segments for FRP generation. This purpose is served by the proposed SRF pattern

Algorithm 5.1: Input_Data_Generation

Input: X^f is a set of N^f segment of TS of fault type f, each with g sensor values, and b is the bin from X^f, where $S_r \leq |b| \leq S_l$ with m such bins. P^f_{ref} is the reference pattern of fault f.

Output: $FRP_{3D}, FRP_F, DFC_{LSTM}$: the inputs for classifiers

Initialization :

The vectors D_{ij}, D_i are initialized to 0, and set $S = 0$.

1: **for** $i \leftarrow 1$ to N^f **do**
2: **for** $j \leftarrow 1$ to g **do**
3: **for** $k \leftarrow 1$ to m **do**
4: Let D_{ijk} be the DFC of X^f_{ijk}(k^{th} bin of j^{th}sensor of i^{th} sample)
5: $D_{ij}.append(D_{ijk})$
6: $P_{ijk} = EncodePattern(D_{ijk})$
7: $R_{ijk} = Hamming\_Distance\_Inv(P_{ijk}, P^f_{ref})$
8: $S \leftarrow S + R_{ijk}$
9: **end for**
10: $R_g = S/m$ /*Rank of X^f_{ij}
11: $D_i.extend(D_{ij})$
12: **end for**
13: Let $FRP_{R1}, FRP_{R2}, FRP_{R3}$ be the FRPs generated from top three ranked X^f_{ij}'s
14: FRP_{3D} ←Arrange three FRPs to 3-channel
15: FRP_F ← Flatten & Stack FRP_{R1}, FRP_{R2}
16: $DFC_{LSTM} \leftarrow D_i$
17: **end for**
18: **return** $FRP_{3D}, FRP_F, DFC_{LSTM}$

Procedure *Hamming_Distance_Inv(X,Y)*

19: **if** $|X| \neq |Y|$ **then**
20: print "Length mismatch"
21: **else**
22: $Z = X \oplus Y$
23: $H_{id} = |X| - |Z|_1$
24: **end if**
25: **return** H_{id}

matching, by choosing the sensor segments that create the closest DFC pattern of a particular fault. Algorithm 5.1 and part 3 and part 4 of Fig. 5.2 provides a description of this process. DFC extraction is performed in step 4 of the algorithm for every bin in a segment for each sensor vector. According to the description in section 5.3.1.1, the encoding of the corresponding pattern for each DFC is performed. An elementary inverse operation of the hamming distance then determines the matching score between the prescribed fault's reference pattern and the current fault pattern. A segment's closest fault pattern is set as the reference fault pattern in the testing phase. The average derived from the matching scores of all the bins is the ultimate matching score

of each sensor segment. This matching score is used to rank the top three sensor segments and for the selection for FRP generation. Steps 6 to 13 in the algorithm describes these operations.

5.3.1.3 Fuzzy recurrence plot generation

The significance of FRP was revealed by the comparative study of FRP and RP, which has been provided with an x-component of the Lorenz system, EMG signals, and pseudo periodic synthetic TS [219]. This shows that FRP has precedence over RP concerning parameter specification of dynamical systems and visual effects. Between RP and FRP, the latter is given preference when it comes to the representation of the non-linear TS signals due to the following properties.

1. The similarity of phase-space of the system is represented on the basis of a continuum of grades of membership that alleviates the issue of the threshold determination.
2. A much-enhanced system dynamics visualization is produced in FRP with significant information of texture.

The visual verification of the same is achieved from Fig. 5.1, which presents the FRPs and RPs of unbalance and horizontal misalignment faults from the MaFaulDa dataset. The experiments and results part bring forth the comparison of the same in section 5.4.1.2. The following parameters were selected for FRP generation from the experimentation performed with various FRP parameters. With an embedding dimension of 3, and time delay (τ) set to 1, the rebuilding of the phase space is performed. The number of clusters for FCM is selected as 3, whereas the fuzzy membership threshold (T) is set to 0.5 to change grayscale to black and white. The final FRP image generation scheme for CNNs is shown in part 5 of Fig. 5.2. The 3D-FRP is generated for 3C-CNN by arranging the top three FRPs in the z-axis, which is mentioned in step 14 of Algorithm 5.1. Likewise, to yield multi-sensor F-FRP, the symmetry property of the

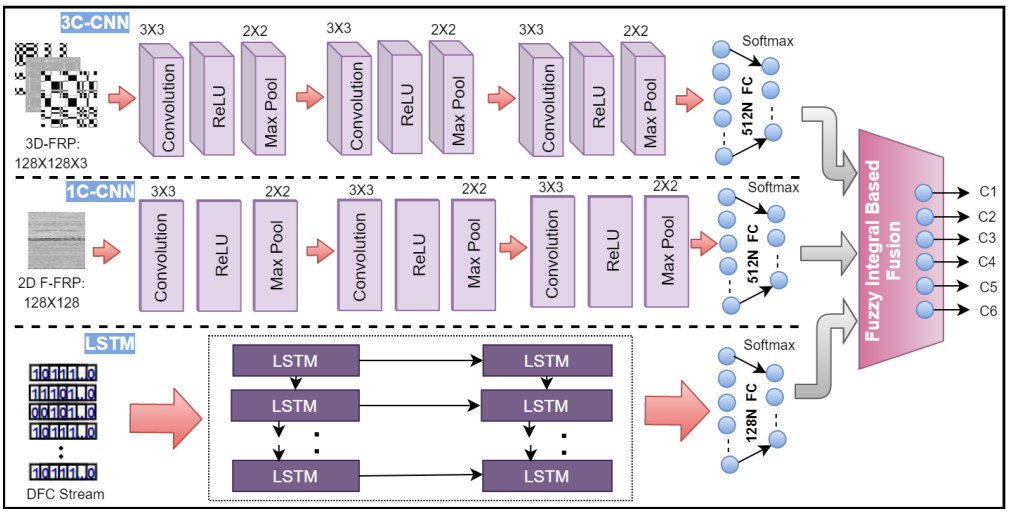


Figure 5.3: Decision-making phase

RP is utilized for combining the top two DFC-ranked FRPs. Step 15 of Algorithm 5.1 presents this process. The detailed procedure of the same is given below:

1. Extract the oblique triangles from the two selected FRPs.
2. Convert the triangles into flattened arrays, and arrange them to matrices of half the size of the original FRP.
3. Combine the two half-sized FRPs to generate the flatted 2D FRP.

5.3.2 Decision score generation phase

The CNN and LSTM based schemes in the framework are used for the assessment of the decision scores as previously mentioned. The CNNs process the FRPs, and the DFC stream is processed by LSTM, which are finally fused by FI-based classifier fusion as displayed in Fig. 5.3.

5.3.2.1 CNN-based decision-making with FRP

For measuring the confidence of affiliating FRPs to a given class, a 3-Channel CNN and a 1C-CNN are used. The CNNs are trained on two different FRP representations, which perform different demonstrations of the same vibration signal segment. The softmax layer produces the decision scores for fusion by making use of the categorical-cross

entropy loss function. The 3D-FRP proposed as input to the 3C-CNN is subsampled to a size of $128 \times 128 \times 3$. Three convolution layers with $32, 64$, and 128; 3×3 filters are possessed by the 3C-CNN model. Each convolution layer has been added with the max-pooling layers of size 2×2. With the introduction of ReLU into each stage, non-linearity has been accomplished as the CNN made capable of learning complex structure FRPs. Column vectors are formed by ultimately flattening the feature maps generated by the convolution stages. A connected layer of 512 neurons is introduced as the last component of the network. To get faster convergence of parameter values, the training has been performed by the Adam optimizer, and the batch size has been set to 64. The 1C-CNN, which works on the multi-sensor F-FRP, follows the same layer structure and gives the next set of decision scores on the same segment of data.

5.3.2.2 Sequential learning-based decision-making

The common sequential learning models like RNN, LSTM, and GRU learn the temporal dynamic behavior of input with the help of their sequential connections. For the purpose of dealing with vanishing/exploding gradient problems, the LSTM and GRU regulate the information in each cell. The LSTM architecture has been utilized in the proposed model with the DFC stream derived from each segment of input vibration. The long-term relationships in the input signal and its temporal information in the network have been apprehended by several recurrent layers. After that, a fully connected layer provides the higher-level representation of data and combines the class discriminating capability. The categorical cross-entropy for loss function and Adam as an optimizer represent the other parameters of the model. Furthermore, $tanh$, $ReLU$, and $softmax$ activation functions are used by the RNN layer, dense layer, and the output layers, respectively.

The hyper parameter search was conducted over the recurrent layers $L \in \{1, 2, 3\}$, hidden nodes $HN \in \{16, 32, 64, 128\}$ and learning rate $\eta \in \{0.1, 0.01, 0.001\}$ for 300

epoch. At the end, the decision regarding the best performing parameters $L = 2$, $\eta = 0.001$, $HN = 64$ was made.

5.3.3 Classifier fusion phase

The CNN and LSTM models possess extensive classifier diversity with two different categories of inputs. Therefore, it is sought after to generate an aggregation scheme that deals with the issues related to classifier diversity and offers an improved classification result than the individual classification. In this respect, in many machine fault diagnosis applications, fuzzy integral fusion has been proven successful.

The Choquet integral finds the score by combining the fuzzy measure for a particular score along with the confidence of the score, called its support. It is evident from Eq. 5.8 that the intermediate decisions determine the final decision-making process that enables sensitivity to each classifier. We have N number of decision scores, and M is the number of classes in the process of fusion. Therefore, the fuzzy measure is altered according to Eq. 5.8 for N times for every class index in M. The prediction scores for each class are then found by making use of the modified fuzzy measure, according to Eq. 5.6. Thus compared to the normalized softmax probabilities previously obtained, a more robust modified fuzzy probability is generated. Finally, the final class prediction is obtained from the maximum of the predicted scores. An operation that checks the selected classifiers' appropriateness and ensures sufficient diversity among them has to front the classifiers' fusion. In the literature, there exist diverse measures that efficiently find classifier diversity in different domains. The Q statistics measure is one of the widely known and simple diversity measures [227]. It can be demonstrated by using the given equation:

$$Q = \frac{a.d - b.c}{a.d + b.c} \tag{5.9}$$

where the number of times both the classifiers come up with, the correct guesses and wrong guesses are given by a and d, respectively. The variable b represent the count

when the first classifier provides the wrong decision but the second displays it right. The value of c gives the count of the opposite condition. The classifier fusion is acceptable if $Q < 1$, no diversity exists if $Q = 1$, and complete diversity exists if $Q = -1$ [227].

To guarantee diversity as the first step of fusion, the three classifiers are checked according to pair with a confusion matrix for both classifiers' predicted value. From Table 5.4, the initial fuzzy measures for FI-based fusion are then calculated by using the following formula [225].

$$\frac{\text{(accuracy+F1-score)}}{2} \text{ - } \min(\text{accuracy}) \tag{5.10}$$

5.4 Results and Discussions

In this section, experimental results are presented and discussed. For the generation of FRPs, the raw data collected from each sensor of the rotor testbed is segmented into 2048 data points. A summary of the same is found in the subsampled space at 512 points with an overlapping of 64 points generate enough DFC stream to apply to LSTM. Since the dataset DS-1 is made up of healthy and five defect classes with five various speeds, the selection of DS-2 data is based on the DS-1 dataset with the closest matching speeds under all load conditions. The same subsampling and overlapping operations of DS-1 were followed by the DS-2 data. The train-test ratio is selected as 70.0%-30.0% to partition both the datasets initially. The evaluation of the effectiveness of the proposed framework is done with two CNNs, and an LSTM model that uses FRPs and DFC feature for operating, respectively. Moreover, the applicability of single sensor data has been checked with and without augmentation, which produced additional FRPs and DFCs for training. For the purpose of conducting the experiments, four sensor values (S1,. . . , S4) from both datasets were considered. The overall experiments have been performed to: i) check the suitability of FRPs as input for SRF diagnosis, ii) evaluate the performance of the proposed paradigm on multi-sensor as well as single-sensor data,

and iii) analyze the advantage of fusing CNNs and LSTM classifiers outputs with FI-based fusion along with other existing fusion techniques.

5.4.1 Evaluation of FRP imaging schemes with CNNs

The imaging scheme of combining FRPs generated from multiple sensors has been compared in two ways to ensure its appropriateness as an input representation. Firstly, a comparison has been made with the existing pre-trained CNN models along with the proposed CNNs. Secondly, results from the FRP are evaluated against its alternative imaging scheme, i.e., RP. The effectiveness of the proposed FRP image combining scheme with DFC-based ranking has been presented in this section.

5.4.1.1 Comparison with existing pre-trained models

Popular pre-trained CNN models have been used first to check the proposed image combining schemes for FRPs. The evaluation result has revealed the generalized usage of FRP inputs for vibration analysis. The derived results are displayed for sensor-wise FRP, 3D-FRP, and F-FRP images as input categories. Demonstration of the comparison of the proposed model is available in Table 5.2 with four pre-trained CNN models like VGG16, Inception V3, Resnet50, and EfficientB0. Single-sensor input images are duplicated to other channels to match the pre-trained CNN model's input shape. For comparison purposes, FRPs are generated from individual sensor values and applied to the pre-trained CNN model to identify their performance on the non-natural FRP image.

As shown in Table 5.2, apart from EfficientB0, the pre-trained models produced only around 75.0% accuracy. On the other hand, EfficientB0 can attain approximately 82.0% of accuracy. This shows that the performance beyond a specific level with FRP cannot be achieved with the natural image-based pre-trained models. In the meantime, it is seen that an improved accuracy compared to the individual sensor FRPs has been

Table 5.2: Comparison with various pre-trained CNN models

Model	S1 FRP		S2 FRP		S3 FRP		S4 FRP		3D-FRP		F-FRP	
	DS-1 (%)	DS-2 (%)	DS-1 (%)	DS-2 (%)	DS-1 (%)	DS-2 (%)	DS-1 (%)	DS-2 (%)	DS-1 (%)	DS-2 (%)	DS-1 (%)	DS-2 (%)
VGG16	72.09	74.50	73.00	73.14	69.23	72.40	70.10	74.60	73.11	77.49	75.50	76.01
Resnet50	76.00	76.56	74.3	75.00	72.69	74.22	73.73	75.85	77.34	78.90	78.89	80.01
Inception V3	76.89	77.01	74.31	76.56	73.99	75.45	74.56	74.04	80.11	81.56	80.55	82.00
EfficientB0	82.03	82.04	83.46	82.01	83.11	82.96	79.45	80.98	84.45	85.56	85.96	86.00
Proposed 3C-CNN	92.67	93.50	90.50	92.07	90.75	92.65	89.50	91.60	95.56	96.00	96.13	96.22
Proposed 1C-CNN	91.32	91.99	90.01	91.14	90.10	92.11	89.40	90.43	N.A	N.A	96.82	96.73

Table 5.3: Comparison of FRPs with RPs

Model	RP						FRP					
	Single Sensor (Avg.)		3D-RP		F-RP		Single Sensor (Avg.)		3D-FRP		F-FRP	
	DS-1 (%)	DS-2 (%)	DS-1 (%)	DS-2 (%)	DS-1 (%)	DS-2 (%)	DS 1 (%)	DS 2 (%)	DS-1 (%)	DS-2 (%)	DS-1 (%)	DS-2 (%)
VGG16	70.11	73.11	72.15	75.65	72.25	73.14	71.11	73.66	73.11	77.49	75.50	76.01
Resnet50	72.67	73.04	75.29	76.65	73.96	75.51	74.18	75.41	77.34	78.9	78.89	80.01
Inception V3	73.00	77.23	80.55	80.21	78.99	79.64	74.94	75.77	80.11	81.56	80.55	82.00
EfficientB0	79.87	80.45	82.43	82.96	81.41	82.85	82.01	81.99	84.45	85.56	85.96	86.00
Proposed CNN	88.11	91.35	92.19	94.04	93.76	93.46	90.86	92.46	95.56	96.00	96.82	96.73

observed in the two latter FRPs, i.e., 3D-FRP and F-FRP. DS-2 shows slightly enhanced accuracy compared to DS-1 with the majority of the pre-trained models among the two datasets. This tendency is more apparent with the 3D-FRP and F-FRP inputs also. It is due to more regular patterns produced in the DS-2 data collection setup than the DS-1 data collection setup.

Comparing the different sensor FRPs based on accuracy, the S1, and S2 sensor data can produce slightly better accuracy than the other sensors. Apart from that, approximately 20.0% variation in the performance is seen on various classifiers for each sensor-wise FRP, 3D-FRP, and F-FRP image. Certain facts are revealed with the higher accuracy achieved by the proposed model for all categories of input FRP images. In other words, training of the proposed CNN model is done based on the non-natural image FRPs in order for it to perform much better compared to the natural image pre-trained models. Therefore, this shows that the proposed CNN model can effectively detect the specific fault diagnosis pattern using FRP images. Subsequently, the performance of more than 90.0% for each sensor-wise FRP, 3D-FRP, and F-FRP image is manifested by it. Furthermore, the highest accuracy was achieved by the two proposed FRP image combining schemes. This signified that DFC is influential in selecting the most informative FRPs to present the different faults in SRF diagnosis. Apart from being perceptible in the proposed model, it is also visible with the results of all the compared pre-trained models.

5.4.1.2 Comparison with recurrence plots

In order to establish the dominance of the proposed FRPs in making use of the system's underlying dynamics, the experiments that were initially performed with FRP have been repeated with RP. The effectiveness of using FRP over RP is demonstrated in Table 5.3 along with both the datasets of SRF diagnosis. Virtually, the same trend of FRP on both pre-trained models and the proposed CNN models can be observed from the

performance of RP, i.e., DS-2 got slightly increased accuracy over DS-1. Also, compared to the pre-trained models, the proposed CNNs showed to be much better. That said, compared to RP's performance, the FRP has demonstrated slightly better. According to the critical observation, FRP produced around 2.0% increased accuracy than RP, which is more apparent in 3D-FRP and F-FRP input data. In any case, according to the result, it is evident that the error rate has been reduced at least by approximately 23.0% with the use of the FRP image scheme and 13.0% compared to RP for DS-1 and DS-2 datasets, respectively. However, a significant reduction in the error rate in F-FRP has been observed, and it is about 50.0% for both datasets. Likewise, in the case of 3D-FRP, it is close to 43.0% and 33.0% for DS-1 and DS-2 datasets, respectively. The effect becomes noticeable due to the soft thresholding that has been made use of in FRP gives rise to more finite patterns than the hard threshold RPs. When the latter two input representations are observed, the property increases the information content to present various faults. This shows that the proposed FRP combining scheme which is based on DFC, is efficient against FRP and RP on both datasets. Therefore, the proposed FRP selection and combining scheme can be adopted in the case of SRF diagnosis domains where multi-sensor data is present with any non-natural imaging schemes. According to what was discussed in the previous section, it created additional impact over natural image pre-trained models as the proposed CNNs are trained with non-natural RPs or FRPs. Based on the entire trend demonstrated in Table 5.2 and 5.3, the best performer is the presented CNNs with the proposed FRP image combining schemes.

Table 5.4: Performance of individual classifiers

Model	**DS-1**		**DS-2**	
	Accuracy (%)	F1-Score (%)	Accuracy (%)	F1-Score (%)
LSTM	97.25	97.19	96.69	96.70
3C-CNN	95.56	95.60	96.00	96.06
1C-CNN	96.82	96.80	96.73	96.80

5.4.2 Performance of proposed framework

The performance assessment of the entire framework comprising of three independent classifiers has been given in this section. To begin with, the classifiers are checked one after the other for accuracy and F1-score with the proposed input representations. These values comprehend the level to which the fusion scheme has enhanced the performance, and serve the parameter setting of FI-based fusion. Apart from FI-based fusion, testing has also been performed on the other empirical decision fusion strategies like maximum, average, and weighted average schemes. The subsequent subsection carries the presented results.

5.4.2.1 Performance of individual classifiers

We have used three classifier models to handle different input representations generated out of segments from vibration signals of each sensor. To generate output from the feature-based scheme, from each bin, the DFC extracted values are combined in a sequential manner and fed to an LSTM. The performance of these three individual classifiers is shown in Table 5.4. Performance exceeding 95.0% on both datasets is demonstrated by all the classifiers. Also, the performance of the classifiers remains in an accuracy cluster of 95.0% to 97.0%. In any way, the highest accuracy for both datasets is provided by LSTM, and it is close to 97.0%. It is the usual course that DS-2 performs a little better than DS-1 in CNN models. However, in the case of LSTM, DS-1 features give better accuracy than DS-2. This happens as DS-2 has greater diverse working conditions, which generate more varying nature features compared to DS-1. Consequently, DS-1 gives more consistent accuracy with FRP as a result of the overall system dynamics represented by FRP not reflecting the diverse working conditions as different sets of patterns. According to what was seen in the previous section, slightly increased accuracy compared to 3C-CNNs is being provided by 1C-CNN with F-FRPs. The classification performance is tested with accuracy as well as F1-score

Table 5.5: Performance of classifier fusion

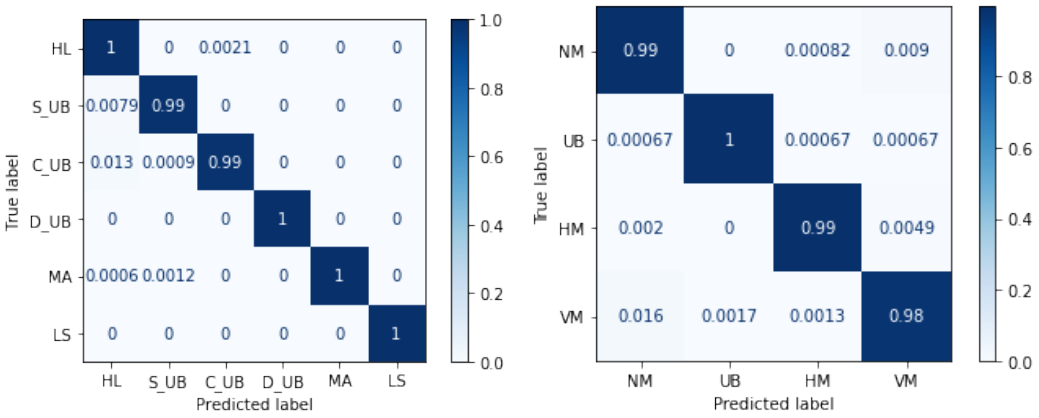
Method	**DS-1 (%)**	**DS-2 (%)**
Maximum	95.62	96.19
Average	96.37	96.38
Weighted Average	98.82	97.11
Fuzzy Integral	**99.60**	**99.06**

metrics because the real-world machinery fault datasets are comprised of imbalanced class distribution. The individual classifier performance is in a satisfactory range, as shown by the results expressed in Table 5.4. Moreover, the determination of the initial fuzzy measures in FI-based fusion is also facilitated by the F1-scores.

5.4.2.2 Framework performance with fuzzy integral fusion

Table 5.5 shows the comparison of the fusion results of FI with certain other decision fusion strategies. Amongst the empirical fusion strategies that were compared, the maximum accuracy is produced by the FI-based fusion. An accuracy that is slightly better than the minimum accuracy of individual classifiers is generated by the two initial fusion methods. In the meantime, an accuracy that is closer to the FI method is generated by the weighted average fusion. It is noteworthy that an accuracy better than the highest individual classifier accuracy can only be produced by the weighted average and FI-based fusion strategies, which causes fusion between the classifier relative scoring. Therefore, the error rate of an individual classifier has been reduced by the FI-based fusion strategy by around 85.0% and 71.0%, respectively, for DS-1 and DS-2 datasets. Thus, the importance of FI-based fusion strategy in the proposed method to better the overall fault diagnosis performance is evidently established.

Furthermore, to evaluate the final fusion performance with multi-sensor data, the confusion matrix is shown in Fig. 5.4. The outstanding performance for healthy and faulty components is manifested by the proposed FI fusion. At the same time, certain



(a) DS-1 (b) DS-2

Figure 5.4: Confusion matrix of FI fusion

samples of couple UB in DS-1 have been misclassified as healthy or static UB. This happens due to the properties of couple UB with equal weights, which at times can create the same effect of healthy samples. In other instances, its effect is similar to static UB, which creates some adversity in classification. Apart from the misclassified samples of vertical misalignment, for the DS-2 dataset, a decent performance was shown by the proposed fusion. All in all, by achieving decent results for SRF diagnosis, the worth of the proposed fusion method is showed.

5.4.3 Individual sensor performance with fusion

In order to evaluate the effectiveness of the proposed framework, the outcomes of a single sensor situation are also assessed with the FI fusion scheme on FRP and DFC. The benefit of the DFC pattern-based ranking scheme of FRPs with the multi-sensor data is evidently showed by analyzing the individual sensor performance in this experiment.

As shown in Fig. 5.5, the fusion of sub-sampled DFC patterns and FRP even with a single sensor value can produce an accuracy around 92.0% to 94.0%. The benefit of fusion and the improvement of the proposed multi-sensor FRPs combining scheme are signified by the results. What is noteworthy here is that a specific sensor data, whose increased effectiveness with FRP was showed, does not need to have the same impact

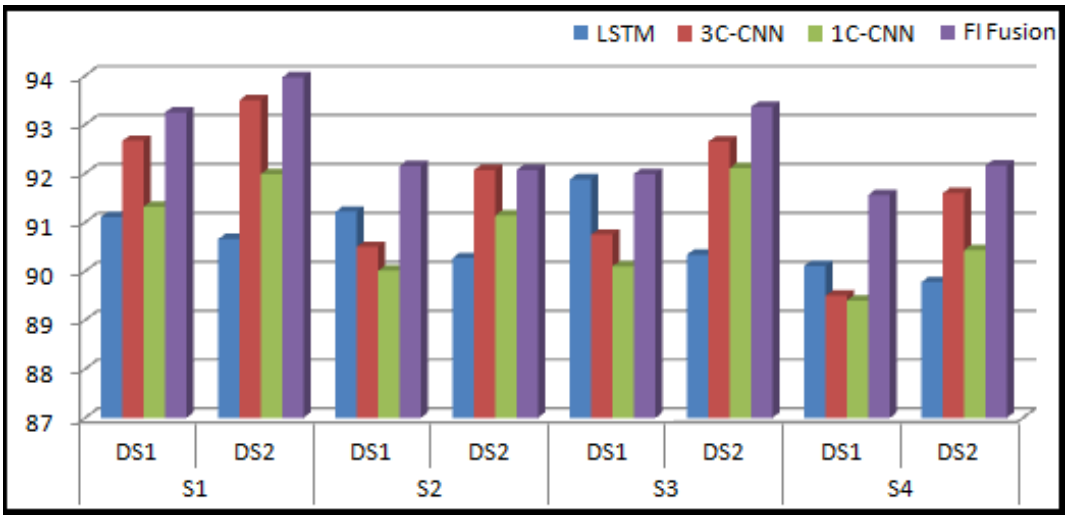


Figure 5.5: Individual sensor fusion

Table 5.6: Effect of augmentation on sensor data

Classifier	Sensor Avgerage Accuracy			
	Without Augmentation		With Augmentation	
	DS-1(%)	DS-2(%)	DS-1(%)	DS-2(%)
LSTM	91.08	90.26	92.15	91.09
3C-CNN	90.86	92.46	91.23	92.86
1C-CNN	90.21	91.42	90.67	91.98
FI Fusion	92.18	92.92	93.00	93.57

with DFC features and vis-a-vis.

A comparison of the results in Table 5.5, shows that an accuracy closer to 100.0% can be achieved by the fusion of the combined FRPs and multi-sensor DFCs. Apart from this, it was observed that when single sensor DFCs produced accuracy around 91.0% with LSTM, and it is raised around by 97.0% with multi-sensor DFCs (refer Table 5.4). To summarise, effective utilizing of multi-sensor data with DFC was proven by the proposed framework, producing satisfactory accuracy even with single sensor data.

5.4.4 Effect of data augmentation

In the situations where single sensor data is available, the training samples lack diversity both in FRPs and DFCs, resulting in reduced accuracy. To overcome such a situation we followed the soft-DTW-based augmentation scheme [183] that preserves the TS properties, which is showed efficient with the same SRF datasets (DS-1 and DS-2).

Table 5.6 shares an interesting observation that the augmentation preserves the TS property and increases the accuracy of LSTM, which uses DFC around 1.0%. Meanwhile, this impact is not much visible with the FRP data. In the case of the overall sensor data applied to the FI classifier fusion, the augmentation produced negligible accuracy enhancement.

5.5 Summary

This work presents a framework for SRF diagnosis with classifier fusion based on the fuzzy integral. It incorporated the TS imaging scheme to maintain the temporal property generally lost in other imaging schemes and explore hidden dynamical patterns and non-linearities in data. The preliminary experiments conducted with GAF, MTF, and RP on SRF datasets signified the dominance of RP in this domain. A more enhanced system dynamics visualization with comprehensive texture information is produced by using fuzzy membership in the FRP imaging. The proposed framework showed efficient in generating FRPs utilizing multi-sensor data with DFC-based ranking that effectively deals with the FRPs from multiple sensors. This mitigates the inordinate effort required in developing the less informative FRPs, reducing the complexity.

The model has the advantages of combining automatically extracted features from FRP by CNN as well as the manually extracted DFC features from the same input for decision-making. Thus the fusion of the CNN-based decision scores along with the sequential classifier like LSTM based decision scores improved the results across

the board to a great extent. The DFC stream extraction suppressing the industrial data issues makes the framework suitable for practical industrial environments. The experiments showed the suitability of FRPs as input for SRF diagnosis, and at the same time, it showed that the proposed CNN model trained on the non-natural image FRPs, is more suitable for SRF than the natural image pre-trained CNN models. Comparing FRPs and RPs, the critical observation is that FRP produced around 2.0% increased accuracy than RP, which is more evident in 3D-FRP and F-FRP input data. This is observed because the soft thresholding used in FRP creates more finite patterns than the hard threshold RPs. Moreover, all the individual classifiers produced more than 95.0% accuracy on both datasets. It is also noticed that the overall system dynamics represented by FRP do not reflect the varying working conditions as different sets of patterns, and therefore DS-1 gives more consistent accuracy with FRP than DS-2.

The experiments with classifier fusion show that the FI-based fusion produced the maximum accuracy among the compared empirical fusion strategies. It is also observed that the fusion strategies, which fuses the classifier relative scoring, can only produce an accuracy better than the highest individual classifier accuracy. As the properties of the couple UB fault sometimes create a similar effect of healthy samples, such misclassification is noticed from the results of DS-1, and similar adversity in classification is created for static UB in DS-2. In the case of the overall sensor data applied to the FI classifier fusion, the soft-DTW augmentation produced negligible accuracy enhancement as it preserves the TS properties in the synthesized samples. The experimental results demonstrated the efficacy of the proposed framework in satisfying objectives like TS characterization and system dynamics consideration of the input, effective DFC usage, and efficient classifier fusion.