

Received January 3, 2021, accepted February 23, 2021, date of publication February 26, 2021, date of current version March 5, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3062496

An Intelligent Hybrid Scheme for Identification of Faults in Industrial Ball Screw Linear Motion Systems

NAVEED RIAZ¹, SYED IRTIZA ALI SHAH², FAISAL REHMAN³,
AND MUHAMMAD JAWAD KHAN³

¹Department of Mechanical Engineering, School of Mechanical and Manufacturing Engineering, National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan

²Department of Aerospace Engineering, College of Aeronautical Engineering, National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan

³Department of Robotics & Intelligent Machine Engineering, School of Mechanical and Manufacturing Engineering, National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan

Corresponding author: Naveed Riaz (engrnaveedriaz@gmail.com)

ABSTRACT Reliability of high precision linear motion system is one of the main concerns in industrial and military systems. The performance and repeatability of these systems are influenced by their respective linear drives and load bearings. A fault in these members severely affects the safe working of overall system. This paper gives a reliable intelligent approach to detect and classify faults for linear motion systems based on deep learning methods. Accuracy in faults identification is highly dependent on improved features extraction. For this purpose, a novel Residual Twin CNN (ResT-CNN) is proposed that uses combination of 1-D and 2-D CNN in parallel learning which improves features extraction performance; followed by knowledge base-Remnant-PCA (Kb-Rem-PCA) architecture in combination with multi-class support vector machine (Mc-SVM). This novel hybrid combination proved very effective in accurate faults identification. The performance of proposed methodology was also validated by IMS-UC (Intelligent Maintenance Systems – University of Cincinnati) public bearing dataset. The results confirm the effectiveness of proposed scheme in comparison to existing state of the art techniques.

INDEX TERMS Linear motion system, intelligent hybrid scheme, reliability, feature extraction, classification accuracy.

I. INTRODUCTION

Linear motion system (LMS) is the most common choice in precision motion applications, especially where high speed repeatability is desired under load. LMS covers wide range of applications including precision machining centers, industrial robots, automatic guided vehicles as well as surface actuation requirements in aerospace industries. LMS incorporates various subsystems, including power and control drive, actuation drive and linear motion structure. The actuation drive provides linear motion while carrying thrust loading. Ball screw (BS) is the most suitable linear actuation drive currently in practice due to high transmission efficiency, low friction and less backlash problems. Any degradation in ball screw drive leads to failure of entire system. These linear drives

carry big research potential for automation and aerospace industries and have gained popularity by researchers working with precision motion control systems. The utilization of BS linear drives for high precision applications also needs additional knowledge of possible failure modes and their effective investigation, since they have to accomplish critical tasks.

BS linear drives include servo motor with required speed reduction gear set and necessary bearings along with ball screw arrangement as shown in Figure 1. For high speed actuation drives, ball-screw and load bearings need to be considered as key components and therefore typical fault modes for these elements should be examined. Critical fault modes for these systems available in literature [1] include friction due to inadequate lubrication in bearings and ball screw, backlash and channel jamming in ball screw, wear or spall at bearing and ball screw surface and other structural faults.

The associate editor coordinating the review of this manuscript and approving it for publication was Xinyu Du.



FIGURE 1. Schematic of BS linear actuation drive.

Accurate identification of these faults is of great industrial concern to ensure system availability and has been considered by many researchers in the past [1], [2], [4], and [5]. For this purpose, dedicated monitoring setups are required which essentially need expertise in electro-mechanics, controls and intelligent systems. This eventually becomes high resource solution.

Fault data monitoring for BS linear systems have been the area of primary concern by various researchers in the past [2], [3]. For this purpose, system input data was monitored after installing linear motion drive on actual system to ensure monitoring of reliable data [4]. Data monitoring was also performed by collecting fault signals using some suitable experimentation [5]. More specifically, BS linear drives were explored in many analytical, simulation and experimental research studies using different approaches.

Mechanical faults identification can be treated as pattern recognition problem; therefore extraction of appropriate fault signal features with non-stationary behavior represents multiple failure pattern characteristics. This signal pattern must include fault sensitive features; as the performance of intelligent monitoring system is greatly influenced by these features. Accurate features extraction, being sensitive to dynamics of mechanical systems, requires high degree of domain expertise. Due to these concerns, design of appropriate feature extraction technique is vital for intelligent fault diagnosis system.

II. BRIEF DETAILS OF PREVIOUS WORK

A number of techniques have been developed that give model-based engineering solutions and numerical approaches to solve system diagnosis problems [6]. These include analytical and experimental approaches [7] to solve system level problems [8] and diagnostics [9]. These solutions although give reliable results, however, undergo application limitation due to customized modeler for each system. Other methodologies utilize vibration and temperature signal data measurement by using suitable sensors and applying different feature extraction techniques for faults identification in critical elements [10], [11]. Health assessment of BS drives based on experimental and systematic studies was also performed for early fault diagnosis [12], [13]. Recently, the applications of machine learning and deep learning based intelligent techniques have rapidly increased for mechanical

systems fault diagnosis [14], [15] and different approaches have been successfully applied for rotating machinery fault detection [16], bearings fault diagnosis [17] and gears fault classification [18]. These intelligent techniques work automatically by collecting and processing fault signals to identify condition of the system. An intelligent feature extraction technique for combination of faults in rotating machinery was proposed on the basis of Empirical mode decomposition (EMD) and Bayesian classifier was applied to diagnose faults [19]. Other contributions include faults classification using Support vector machine (SVM) for vibration signal data [20] and nearest neighbor classifier [21], [22]. In comparison to above mentioned machine learning techniques, deep learning techniques achieve better results, although its implementation in this area is still developing [23]. For BS drives data monitoring, a deep learning technique based on deep belief network (DBN) was performed using multi-sensor vibration signal data [23]. For motor faults, a lot of study has already been done previously [24] and will not be considered here. Among deep learning methods, Convolutional Neural Network (CNN) emerged as one of the most leading technique especially when dealing with more complex features [25], [26].

The reliable monitoring of BS drives is affected by insufficient and inappropriate data. The fault signal data from BS drive system is difficult to collect than from other systems. This implies that for normal working of linear motion systems, data collection can be performed by different methods; however, for different fault modes, sufficient data is not available. This unbalanced system data distribution affects the performance of algorithm leading to unsatisfactory results for minor data set [17]. Additionally, various fault detection techniques are in practice and have utilized successfully in different combinations [27]–[29], however, fault classifiers like Support Vector Machine (SVM) [30], [34] and Back Propagation Neural Network (BPNN) may sometimes misclassify samples because of low domain adaptability.

In order to deal with aforementioned concerns, this paper proposes data monitoring of BS linear drive using significant signal features from position error measurements. These remarkable features represent changes in system dynamics due to any upcoming failure. The focus of this work is to improve signal representation for improved feature extraction and classification. A novel combination of improved deep learning and knowledge base systems is proposed that gives better feature extraction and accurate faults classification for BS linear drive system. This hybrid arrangement gives superior performance over other techniques.

The summary of our contributions in this paper are as follows:

- a) Position measurement data was collected for fault-free and faulty conditions. Faults were induced in BS drive and load bearing to observe their combined effects on collected signals.
- b) Improved feature extraction with better domain adaptability was achieved using Residual Twin-CNN

(ResT-CNN) structure, that uses parallel learning of 1-D and 2-D CNN's and fusing the extracted features; followed by knowledge-based Deep PCA (Kb-Rem-PCA) architecture that ensures distinct features reduction resulting in high classification accuracy using Multi-class Support Vector Machine (Mc-SVM).

- c) This novel hybrid scheme was successfully utilized to detect different faults in BS LMS with improved results comparable with state of the art techniques.

The remaining sections in this paper are composed as follows: Section III describes details of our proposed hybrid methodology. Section IV gives experimentation details that include testing setup description and details of fault cases. Section V elaborates the detailed results of our technique followed by comprehensive discussion in section VI.

III. PROPOSED HYBRID SYSTEM

Mechanical faults in BS LMS, like those stated in Section I, disturb entire system process and degrade its performance to great extent. For high precision systems, LMS continuously undergoes acceleration and deceleration for each travel cycle between two extremes. This gives variation in position measurements which is affected by system dynamics, axial and thrust loading and behavior of control system. These remarkable deviations in position behavior can be well analyzed for different operation modes. Variation of these signal features gives an indication of changed mechanical behavior of actuation drive due to any possible fault. Observed measurements for different fault modes including ball screw (BS) friction, BS backlash, load bearing (LB) friction and LB roller damage are shown in Figure 2.

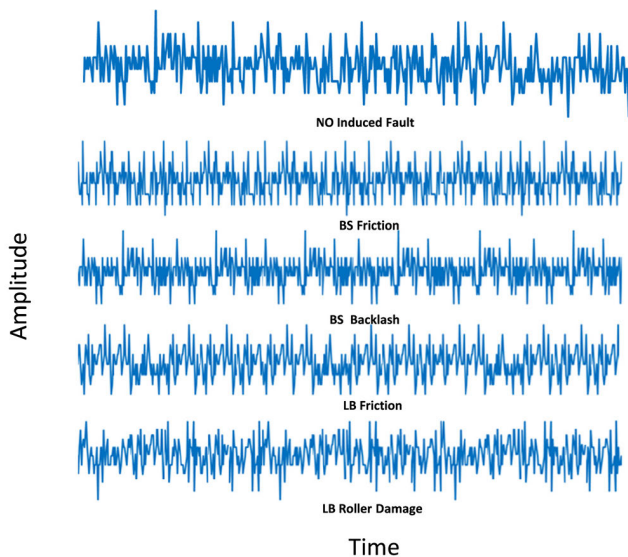


FIGURE 2. Position measurements for different faults of BS linear drive.

Position error measurement was performed for different load conditions using sinusoidal motion profile to assess the capability of developed technique. Sinusoidal profile gives smooth motion transition and was tested with 20mm linear travel which was completed in 4sec with 1sec waiting time at

position change. Three load scenarios were tested i.e. 600N, 750N & -750N. The tests were performed for both fault and no-fault (normal) conditions. To get significant number of observations, the sequence of motion profiles was repeated 3 times for each test. The dataset was generated for each test with 12 times repetition for each case under study.

Position measurement was conducted for normal function as well as for different faults as described in Section IV. The acquired dataset was transformed into 2-D gray image using CWT technique. This gives an opportunity to utilize two-dimensional features from collected signal data, and therefore time-signal can be considered as image feature set [31].

A. 1-D TO 2-D SIGNAL TRANSFORMATION USING CWT

Different techniques have been utilized for signal to image transformation including Hilbert-Huang transform [44], Recurrence plots to transform 1-D signal to 2-D texture image [45], Wavelet Transformation [49], [50], signal to gray image conversion using energy values [51], etc. Wavelet Transform gives time-frequency analysis by decomposing input signal into a family of wavelet components; each one with a resolution according to corresponding scale. This gives high frequency (low time) resolution in low frequency region and similarly low frequency in high region. Continuous wavelet transform utilizes wavelet functions generated by translating wavelet function $\psi_{q,p}(t)$ with translation parameter q and scaling with factor p . The scaled translated and normalized wavelet is given by

$$\psi_{q,p}(t) = \frac{1}{\sqrt{p}} \psi \left(\frac{t-q}{p} \right), \quad q \in \mathbb{R} \quad (1)$$

Continuous wavelet transform (CWT) of observed function $f(q,p)$, where $f \in R$ is given by [52]

$$CWF(q,p) = \int_{-\infty}^{+\infty} f(t) \dot{\psi}_{q,p}(t) dt \quad (2)$$

where, $\dot{\psi}_{q,p}(t)$ represents wavelet complex conjugate for mother function $\psi_{q,p}(t)$. It should be noted that CWT is sensitive to transitions in local occurrence which means time distant CWT will remain unaffected from local transitions.

Among different mother wavelet functions, Morlet wavelet has proved its effectiveness for non-stationary signal data because of transient impulse similarity [53], [55]. Time domain Morlet wavelet is defined by

$$\psi_{q,p}(t) = \cos \pi t \cdot e^{-\frac{\beta^2 t^2}{2}} \quad (3)$$

where, β is the shape factor for mother wavelet. Increase in β value increases time domain resolution. Morlet CWT represents improved signal characteristics and gives better time-frequency resolution. For different scaling parameters, CWT generates different coefficients for signal segments. Using these coefficients, signal can be expressed as 2-D image. The gray image matrix transformation P_{trans} ,

by putting wavelet coefficients, can be represented as

$$P_{trans}(m, n) = \frac{K(m, n) - MinK}{MaxK - MinK} * 255 \quad (4)$$

where Max K and Min K are max and min elements of matrix P. P_{trans} gives gray values from 0 to 255. This gives CWT gray image for original signal as shown in Figure 3.

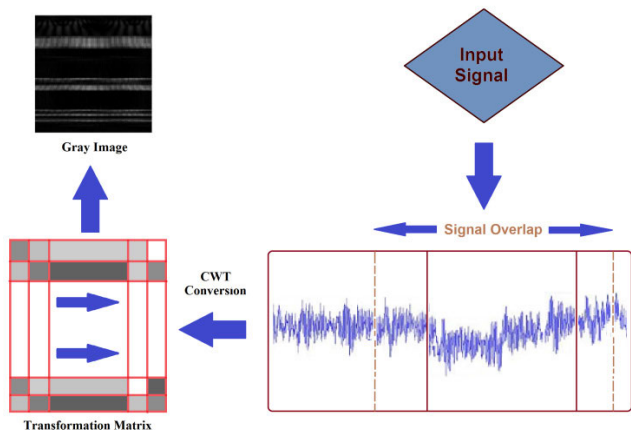


FIGURE 3. Signal Transformation from 1-D to 2-D using CWT.

B. PROPOSED ResT-CNN ARCHITECTURE

At present, convolutional neural networks (CNNs) are extensively used for fault identification and image recognition problems. CNN comprises combination of multiple filtering and classification layers that extract significant features from data. Filtering stage constitutes a number of convolutional and pooling layers. Convolutional layer contains multiple weights (kernels) that provide new feature maps when input feature is convolved with kernels which are sensitive to certain features. The function of pooling layer is to minimize size of feature map without any change in number of feature maps. The classification stage contains fully connected layer, a multi stage perceptron, in which high-level features are extracted and given as input to the final layer, where output of CNN model is generated. The convolution process can be expressed mathematically as shown in Eq. (5).

$$x^{l(i,j)} = K_i^l * m^{l(n)} = \sum_{j'=0}^W K_i^l(j') m^{l(j+j')} \quad (5)$$

where, K_i^l denotes the i^{th} kernel weights for l layer. $m^{l(n)}$ shows j^{th} region in l^{th} convolutional layer. W represents width of kernel and $K_i^l(j')$ gives j^{th} weight of the kernel.

The output of general model [46] can be given as in Eq. (6):

$$CNN_{out} = \begin{bmatrix} P(y = 1 | x; W_1, b_1) \\ P(y = 2 | x; W_2, b_2) \\ \dots \\ P(y = m | x; W_m, b_m) \end{bmatrix} = \frac{1}{\sum_{j=1}^k \exp(W_j x + b_j)} \begin{bmatrix} \exp(W_1 x + b_1) \\ \exp(W_2 x + b_2) \\ \dots \\ \exp(W_k x + b_k) \end{bmatrix} \quad (6)$$

where, b is the bias of each layer. More literature details are included in [46].

The mentioned CNN model has been used in various combinations to solve engineering problems. For deeper networks the computation time increases, which creates over-fitting problems causing a decrease in network accuracy. These problems were solved by Kaiming He [26] using residual blocks in deep networks as shown in Figure 4.

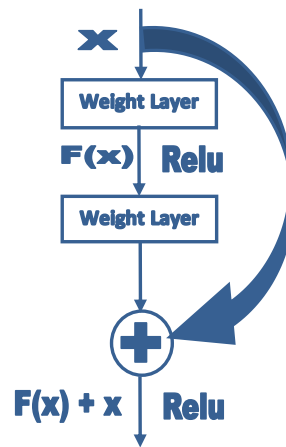


FIGURE 4. Original residual learning block.

This paper proposes a novel Residual Twin CNN (ResT-CNN) architecture inspired from residual learning scheme. The proposed network combines the benefits of two training networks 1-D with raw signal data and 2-D with transformed images in parallel learning. Both networks include residual connections that improve system performance and extract fault related strong features.

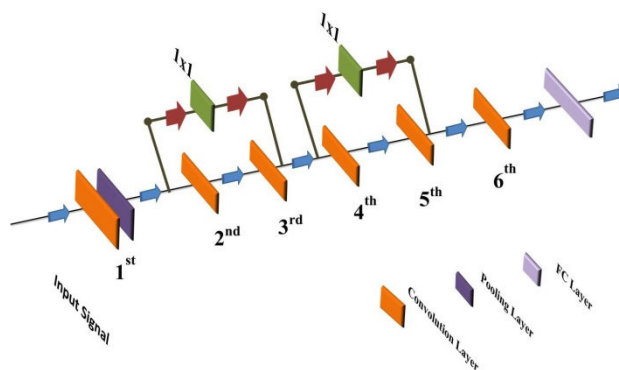


FIGURE 5. 1-D CNN architecture with residual connections.

1) 1-D CNN ARCHITECTURE WITH RESIDUAL CONNECTIONS
1-D CNN acts directly on raw signal data scaled linearly in the interval [0, 1]; and becomes input of 1-D CNN. 1*1 sized two residual connections are introduced in 1-D network as shown in Figure 5 to learn high level features by transferring multiple complex non-linear layers into split functions.

Additionally, these connections construct input-output identity mapping, which improves network speed by enabling information flow across stacked layers. Remaining structure comprises 1*12 size first convolutional kernel along with 1*3 size four successive kernels. The input signal mapping is required in the interval [0, 1], which was performed using max-min normalization as per following relation.

$$x_j^l |_{1}^{N-1} = \frac{x_j^l - \min x^l}{\max x^l - \min x^l} \quad (7)$$

where, x^l gives l^{th} sample in dataset with $N-1$ samples and max and min of x^l gives maximum and minimum sample values.

The forward propagation in 1-D network from any convolutional layer $k-1$ to the input in layer k is given by

$$x_i^k = b_i^k + \sum_{j=1}^{N_{k-1}} \{1DConv(n_{jk}^{k-1}, m_j^{k-1})\} \quad (8)$$

where, b_i^k shows bias of l^{th} layer, m_j^{k-1} gives the output of j^{th} neuron at layer $k-1$, n_{jk}^{k-1} is the kernel from j^{th} neuron at layer $k-1$ to the i^{th} neuron at layer k that computes k^{th} layer input x_i^k .

The output feature for weight layer K_j is obtained by sliding convolutional kernel with size 1×1 on input feature signal K_i . The output feature signal y at j^{th} node is given by

$$y_j = K_j x_i + b_2 \quad (9)$$

The output of residual learning block will be given by

$$Out(res) = F(x_j, W_{ij}) + y = F(x_j, W_j) + K_j x_i + b_2 \quad (10)$$

2) 2-D CNN ARCHITECTURE WITH RESIDUAL CONNECTIONS

An improved 2-D CNN structure is proposed that includes two residual connections; each comprises a single 3*3 convolutional layer in between coupled 1*1 convolutional layers. To minimize high frequency noise effects, a wide size 9*9 first convolution layer is applied. To make deeper network and improve input data representation, 3*3 sized remaining convolution kernels are selected. At the end of residual connections, max pooling layer is used to avoid image data over-smoothing, therefore minimizing the errors in desired features extraction. Moreover, two Dropout and FC layers are utilized for better structure adaptability. This composition makes CNN structure deeper than previously implemented CNN models.

The output $O(res)$ of residual blocks can be expressed mathematically as;

$$\begin{aligned} O(res) &= f_{r1}(W_1.x + W_2.x) + O_{res1} + m(x) \\ &= f_r(W_1.x + W_1.O'_{res1} + b_1) + m(x) \end{aligned} \quad (11)$$

where, f_r being ReLU non-linear activation function. O_{res1} and O_{res2} shows feature output at first & second residual structure, obtained by sliding kernel with signal stride. Identity mapping $m(x)$ was added to the first residual block and the

output was passed by ReLU activation function. The second filter weight W_2 was found by the output of first weight layer O'_{res1} and gives final output of residual block structure. Figure 6 shows the basic architecture of proposed residual structure.

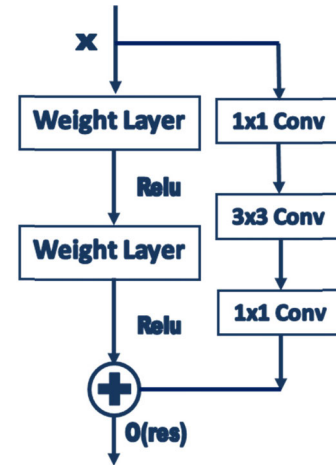


FIGURE 6. 2-D CNN residual block structure.

For x data series where $x \in \{x_1, x_2, \dots, x_m\}$; m being the signal length, x is the output of signal feature data that learns the residual function $Res(x)$ in the neural network more easily without any deterioration thus providing better CNN model learning and improves the performance of system. The architecture of developed 2D CNN is shown in Figure 7.

1-D and 2-D CNNs work independently for features extraction. These features are concatenated before classification processing for faults identification. For 1-D CNN structure, the raw signal data was used to train the model. For 2-D CNN, signal data was transformed to gray image using CWT conversion. If we assume 1D-CNN features as C_{1f} and 2D-CNN features as C_{2f} , then the features of concatenation layer (C_f) can be expressed as;

$$C_f = C_{1f}; C_{2f} \quad (12)$$

These final features were given to the proposed deep PCA for dimensionality reduction. The fused residual twin CNN (ResT-CNN) architecture is shown in Figure 8.

C. PROPOSED DEEP KNOWLEDGE-BASE REMNANT PCA (KB-REM-PCA) ARCHITECTURE

To reduce high dimensionality of features extracted by ResT-CNN, a deep knowledge-based Remnant PCA (Kb-Rem-PCA) is proposed that takes the advantage of deep residual connections in PCA followed by knowledge-based feature selection technique based on improvements in existing Rough set theory [36]. This scheme improves classification accuracy along with reduction in computation cost.

The existing state of the arts PCA techniques are characterized by unsupervised dimensional reduction from extracted features and are employed in different applications with different combinations [37], [38]. However, these techniques

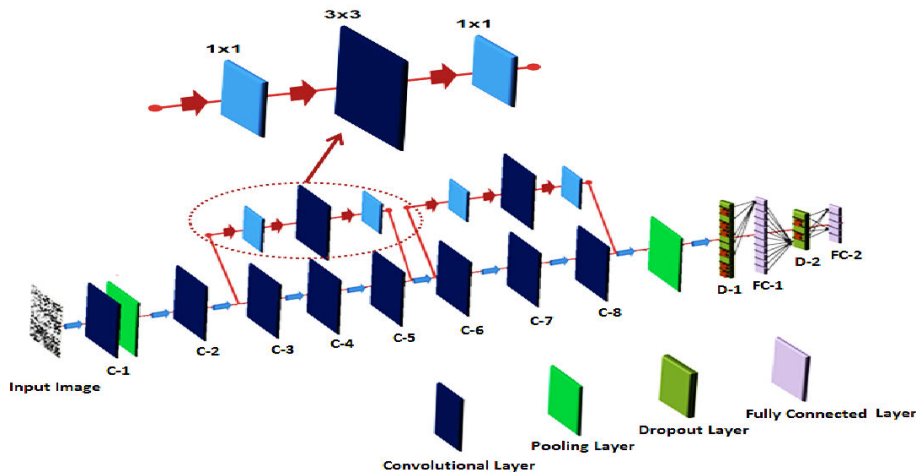


FIGURE 7. 2-D CNN architecture with Residual connections.

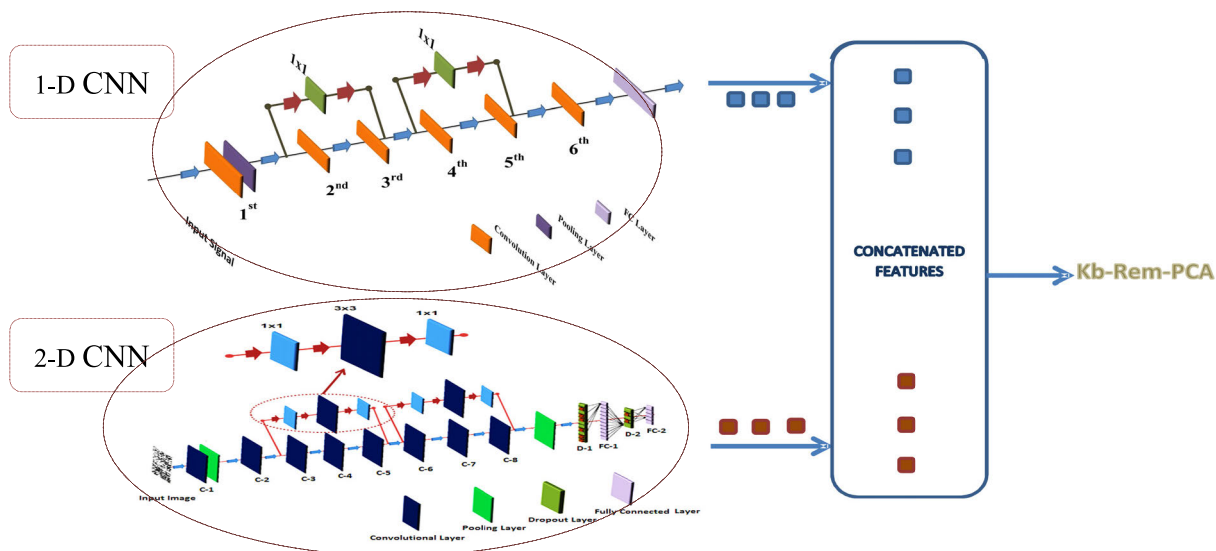


FIGURE 8. ResT-CNN architecture.

undergo some limitations that may lead to incompetency specially for extracting rich features.

The proposed PCA architecture overcomes these limitations easily by utilizing deep multiple feature extraction layers that transform high level features to low level effectively with improved algorithm speed as shown in Figure 9. Deep PCA, however, experiences information loss in each process layer which gives undesirable results. To cater these effects, residual connections are added between each successive layer of deep PCA structure.

Mathematically, for N training set features, $I \in \{i_1, i_2, \dots, i_x\}$ in \mathbb{R}^N , the dimensions reduction for extracted features occur in successive learning layers of PCA, $J \in \{j_1, j_2, \dots, j_y\}$ with projection matrix P and non-linear feature matrix N_t . The layer output can be given by $F \in [0, L-1]$ as:

$$J_y = I_x \cdot P \tag{13}$$

$$F_{0,L-1}^{Z=1} = J_y \cdot N_t \tag{14}$$

The non-linear feature matrix is introduced to improve PCA performance by providing additional non-linear function with

the traditional architecture. For this purpose, the nonlinear feature matrix N_t as mentioned is given as

$$N_t = (\varphi(j_y))(\varphi(j_y))^T \tag{15}$$

where $\varphi(\cdot)$ is the non-linear mapping function. The output F_0 for the first layer will be given by Eq. 16. For layers $[1, L-1]$, the output will be given as:

$$F_{1,L-1}^Z = F_{0,L-1}^{Z-1} + C_{rem} \tag{16}$$

where, C_{rem} gives the output of remnant connections given as

$$C_{rem} = G.(J_{y-1} \cdot N_{t-1}) \tag{17}$$

The function $G(\cdot)$ gives the activation function while stacking successive layers of PCA in rem connections.

In order to collect more compact set of features and to improve PCA performance, we have integrated a knowledge-based feature selection technique with the output of PCA layers. This technique is based on some improvements in existing Rough set theory which was originally

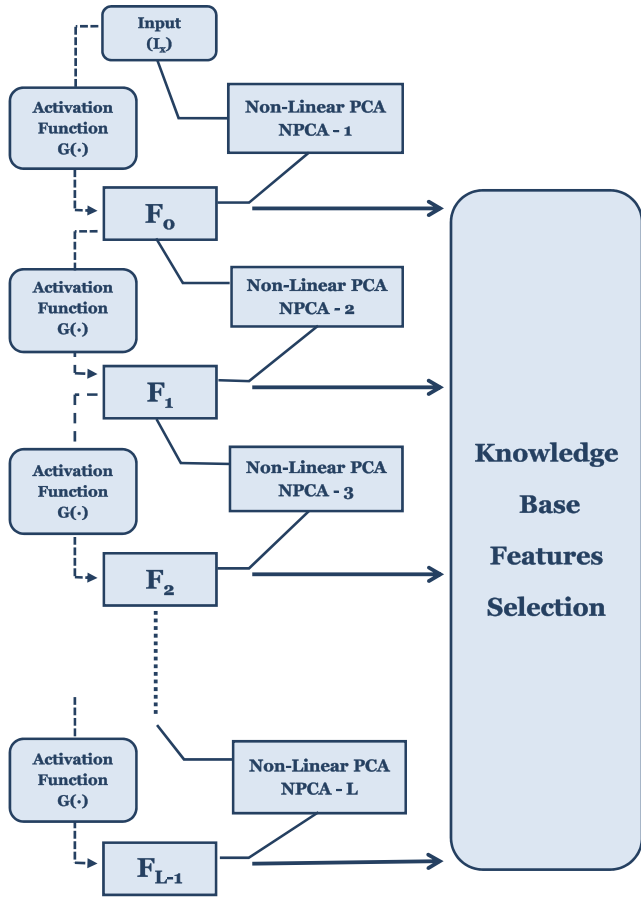


FIGURE 9. Proposed Kb-Rem-PCA architecture.

proposed by Z. Pawlak and has successfully been utilized to eliminate imprecise, redundant and uncertain features [36], [39], [40].

Rough set can be defined as a knowledge base $K_b [O, S]$ where O and S represent overall feature set and selected feature set respectively. The probabilistic based rough set model considers variable precision probability which is more robust against noisy data. A probabilistic precision threshold δ is added which depends on noise magnitude. Based on this threshold, three distinct feature regions (FR) are defined which are as follows.

FR1: δ – True Region: $T_\delta (V) = \{v \in O | P(V| [v]) \geq 1 - \delta\}$;

FR2: δ – False Region: $F_\delta (V) = \{v \in O | P(V| [v]) \leq \delta\}$;

FR3: δ – Border: $B_\delta (V) = \{v \in O | \delta < P(V| [v]), 1 - \delta\}$

Using feature regions, the correlation degree between target and equivalent class can be found by,

$$T_\delta ([v], V) = \frac{|T_\delta(V) \cup F_\delta(V)|}{|O|} \quad (18)$$

D. MULTI-CLASS SVM (MC-SVM) FAULT CLASSIFIER

For classification of faults, a strong multi-class Support vector machine (Mc-SVM) classifier was applied based on Gaussian radial kernel function that gives efficient fault classification [41]. The Gaussian radial kernel function has already proved its outstanding ability and performance in

handling nonlinear faults classification. Mathematically, this function can be given by Eq. 19 [41].

$$R_G (u_m, u_n) = exp \left(- \frac{\|u_m - u_n\|^2}{2\sigma^2} \right) \quad (19)$$

where, $R_G(u_m, u_n)$ gives radial kernel function, u_m and u_n shows input features and σ represents hyper parameter selected on the basis of effective kernel width.

E. PROPOSED OVERALL HYBRID STRUCTURE

A basic outline of our proposed hybrid framework is shown in Figure 10. The flowchart shows the sequential working of model for faults classification in BS-LMS as summarized below.

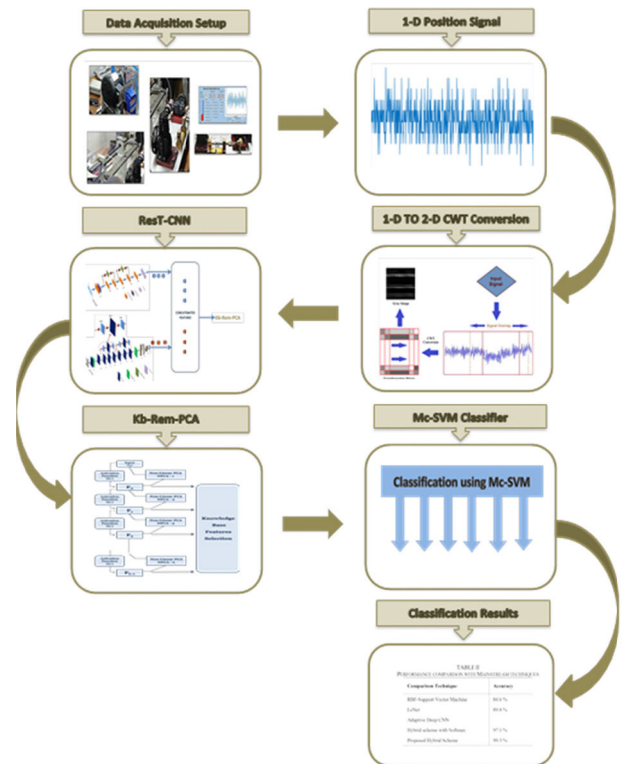


FIGURE 10. Proposed overall hybrid structure.

IV. EXPERIMENTATION DETAILS

To evaluate the performance of proposed algorithm, an instrumented setup is developed. The testing setup is based on BS linear motion mechanism where BS shaft provides necessary ball bearings channel retained inside suiting ball nut. A number of distinctive faults were induced in BS drive and load bearing under varying conditions of external load to consider the effect of faults on measured system parameter as well as to acquire the signal data for detection of these faults.

A. TESTING SETUP DETAILS

The testing setup was developed using available mechanical components including precision rolled miniature BS drive (that consists of BS threaded shaft fitted with ball nut). Ball nut is provided with threaded attachment collar for

assembly integration. Other components required for motion system include motor in combination with necessary gear train along bearings, linear miniature guides and position feedback device etc. The experimental setup was also equipped with different sensors for data acquisition purpose. The resolution, bandwidth and repeatability of sensors should be well analyzed for reliable data acquisition. Linear transducer having linearity error 0.05 %, Hall-effect current sensor with 32.7 mV rated sensitivity @ 12V DC and a tension-compression load cell having 0.02 % rated output error were also integrated in the setup for data monitoring.

The developed experimental setup provides 100mm linear movement with a load capacity of 1000N (in both directions). Figure 11 illustrates some details of experimental setup.

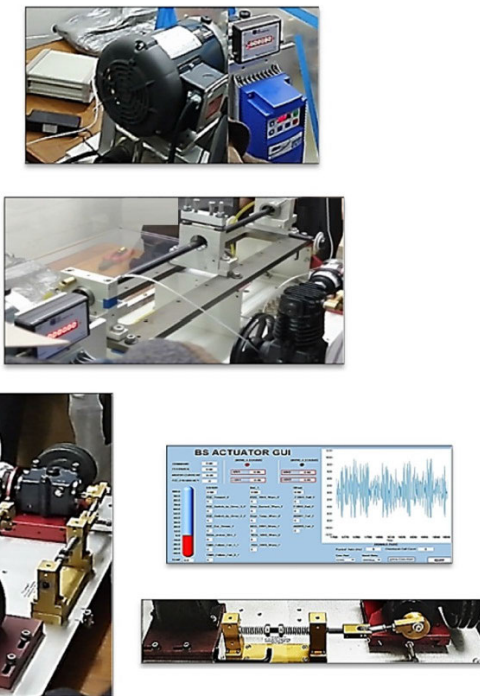


FIGURE 11. Testing setup to simulate fault modes.

B. BS LINEAR DRIVE FAULT CASES

The testing setup was initially operated with no induced fault to get significant signal information and observe LMS behavior under different load and motion profiles. This data was labeled as reference data with acceptable behavior of mechanical system. In the next phase, different faults were induced in LMS and the system was run to collect position measurement data. This data was used to evaluate the performance of developed monitoring scheme. The faults were selected based on repetitive nature and high criticality obtained from previous history. The faults were induced in ball screw (BS) and load bearing (LB) for testing and include the following.

- BS Friction (Insufficient lubrication) fault, developed by ceasing the lubricant and pressing the ball nut seals (as shown in Figure 12).

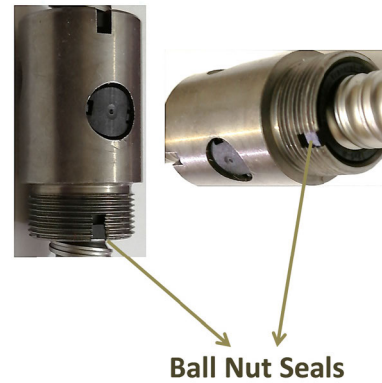


FIGURE 12. BS friction by pressing ball nut seals.



FIGURE 13. BS backlash by varying recirculating balls dimension.

- BS Backlash fault, induced by substituting original balls with undersize recirculating balls (as shown in Figure 13).
- LB Friction (Insufficient lubrication), developed by ceasing bearing lubricant.
- LB Roller Damage, developed by creating surface flaw at load bearing roller (as shown in Figure 14).

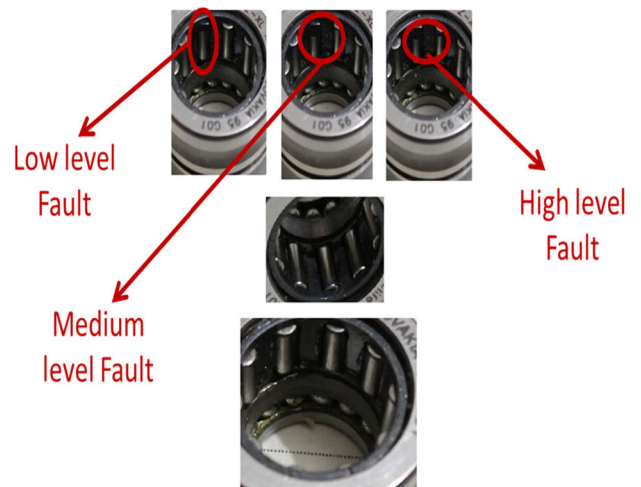


FIGURE 14. LB Roller damage stages.

Since it takes sufficient time for natural failure, therefore, the above mentioned faults were added to simulate system behavior and competency for signal features and faults classification.

In case of BS in-adequate lubrication fault, the lubricant carried by ball nut was initially removed. Signal data was observed and recorded as partial friction fault. Since BS drive offers low characteristic friction coefficient, major signal changes were not observed. Friction severity was increased in the next phase by tightening sealing locks provided at the end of ball drive nut as depicted in Figure 12. BS backlash was studied by substituting original dimensioned balls with undersize balls in ball nut channel. The original size of recirculating balls was 3mm in diameter which was replaced with 2.8mm and then 2.4mm successively. LB friction was analyzed by removing synthetic-based protective grease from bearing. LB roller damage fault was simulated in multiple phases, initiating with small sized flaw at contact surface of roller nearly 1mm wide (low level) as shown in Figure 14. In the next stage, the width of surface defect was increased to 2mm (medium level), simulating fault propagation. In the last stage (high level), the flaw size was increased to 3mm wide with additional surface marks on multiple rollers as shown in Figure 14.

The above faults were tested for BS linear drive against different external loadings to acquire signal measurements and observe performance of developed scheme. One motion cycle of BS linear drive under consideration is completed in 6sec including 1sec dwell time at direction reversals and gives 80mm linear movement. The loads applied include rated load (600 N), positive peak load (750 N) and negative peak load (-750 N). Initially, a number of motion cycles were given to attain system steady state (normalize thermal and mechanical effects). Later on, position measurements were observed for both normal and faulty conditions.

C. DATASET DESCRIPTION

1) BS LINEAR DRIVE DATA SET

The position measurements for mentioned fault cases were obtained using aforementioned experimentation. Signal data augmentation was performed, required for improved feature extraction accuracy and higher performance [35]. This gives increased number of signal samples, therefore improving CNN model generalization [35]. The observed signal data can generate large amount of training data by samples segmentation with signal data overlap (Figure 3). The original signal data with 40000 points can produce 100 samples, which can be extended to 800 training samples with a stride length of 360.

Dataset for four fault categories i.e., BS friction, BS backlash, LB friction and LB roller damage was generated with different severity levels under rated load. The experimental dataset details are given in Table 1. The faults are designated by BF (ball screw friction), BB (ball screw backlash), LF (load bearing friction), LR (load bearing roller damage) and ND (no defect). The intensity levels of these faults are given as H (high), M (medium) and L (low). For 1-D signal to 2-D conversion, as performed in section III-A, 224×224 gray scale image was generated using 50176 signal

TABLE 1. BS dataset description for proposed framework.

Fault Mode	Fault Category	Class	No of Samples		Label
			Training	Testing	
No Defect	All Faults in acceptable range	ND	240	80	0
BS Friction	Lubricant removed (Low Fault)	BF-L	240	80	1
BS Friction	Tightened Seals (High Fault)	BF-H	240	80	2
BS Backlash	2.8 mm Ball Size (Low Fault)	BB-L	240	80	3
BS Backlash	2.4 mm Ball Size (High Fault)	BB-H	240	80	4
LB Friction	Lubricant removed (Medium Fault)	LF-M	240	80	5
LB Roller damage	1mm wide flaw	LR-L	240	80	6
LB Roller damage	2mm wide flaw	LR-M	240	80	7
LB Roller damage	3mm wide flaw + other roller flaws	LR-H	240	80	8

sampling points. Gray images for different faults using CWT conversion are shown in Figure 15. It can be seen that each fault shows distinct gray image pattern which can be easily distinguished.

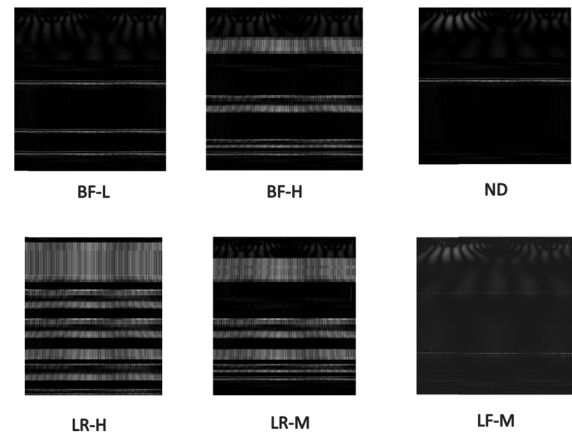


FIGURE 15. 2-D Gray images for different faults using CWT conversion.

2) IMS-UC PUBLISHED BEARING DATASET

IMS-UC (Intelligent Maintenance Systems - University of Cincinnati), USA published bearing dataset was also considered for validation of proposed scheme [55] since it is publically available and used by many researchers. IMS-UC dataset provides RTF (run to failure) maintenance testing results that gives vibration signal data recorded at 20 kHz with 20,480 input data points. 4 accelerometers were attached to acquire vibration signals for bearings faults that include outer race (O_R), inner race (I_R) and rolling element (R_E). The data packet includes three sets. Random samples were selected for each set that includes all types of faults data.

D. IMPLEMENTATION DETAILS

The ResT-CNN model was implemented in TensorFlow Python using Intel core-i7 CPU desktop system with

GTX1070 GPU, 12 GB memory. 1-D and 2-D CNN's were trained with 80 epochs and 150 epochs respectively. Adam optimization algorithm was adopted with a decaying learning rate of 0.001; reduce to one-half after 80 iterations. The dataset was split into training and testing data (along with validation) data. Augmented dataset was utilized for network training since it provides large training samples that avoid data over-fitting and improve model performance. Initial weights of network model were arbitrarily generated. Random samples were selected from defined testing dataset with different sample size to monitor its effect on performance of network.

The feature vectors obtained from 1-D and 2-D CNN's (C_{1f} and C_{2f}) give the characteristics of different faults. The information of both feature vectors can be combined, as shown in section III-B, by concatenating the two types to form a new feature vector C_f that carries strong and abundant information as compared with individual feature vectors. However, C_f contains high dimension features which not only require extra processing power, but a strong SVM classifier as well, for accurate classification. For this purpose, Kb-Rem-PCA was proposed and utilized to reduce redundant data from extracted features. Experiments have shown that the proposed combination gives low dimension feature mapping which gives better classification along with reduced memory and computation cost.

E. PERFORMANCE PARAMETERS

Two critical performance parameters considered commonly to evaluate the system performance include classification accuracy and precision. Accuracy gives a comprehensive classification performance of overall system. Mathematically, it can be expressed as;

$$Acc = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{20}$$

where, F_P and T_P show false positive and true positive samples whereas F_N and T_N represents false negative and true negative samples. False positive comprises positive samples which were classified incorrect and false negative means negative samples which were classified incorrect. True positive and true negative also gives similar interpretation.

Precision represents a measure of correct classification of samples outcome. Mathematically precision is given as;

$$Prec = \frac{T_P}{T_P + F_P} \tag{21}$$

V. RESULTS

The suggested hybrid scheme was successfully trained to learn features from BS linear drive dataset as well as IMS-UC dataset for each fault case. The model was then tested and validated. Figure 16 (a) shows the curves for training and validation results for BS dataset. It can be seen that the proposed hybrid combination gives close to 100% accuracy. Figure 16 (b) gives precision results for BS linear drive dataset. The average precision result for each fault is shown. The precision

ratio for BS friction, LB friction and LB roller damage fault is higher than BS backlash fault. The overall results indicate that the proposed scheme successfully deals with combination of critical fault in BS linear drives.

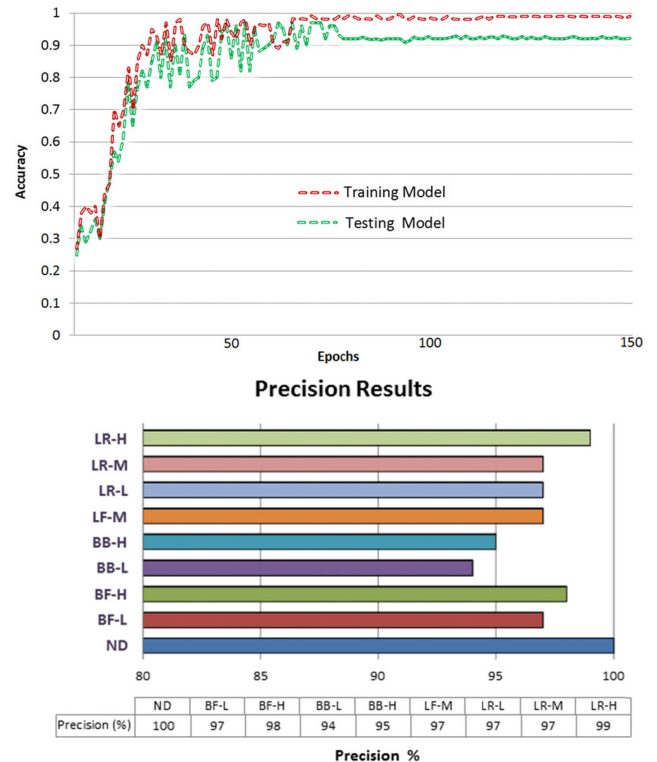


FIGURE 16. (a) Accuracy and (b) Precision results of proposed hybrid scheme.

A. PERFORMANCE COMPARISON WITH MAINSTREAM TECHNIQUES

The developed scheme was also compared with few selected state of the arts to evaluate classification performance. These include LeNet [47], Adaptive deep CNN [56] by converting 1-D signal to 2-D images and implementing adaptive deep CNN structure, RBF-SVM [21] by training time domain signal data feature parameters and our proposed hybrid scheme replacing Mc-SVM with Softmax [48]. Performance results are mentioned in Table 2. It can be observed that RBF-SVM and LeNet-5 gives low accuracy as compared with adaptive CNN. The proposed scheme replacing Mc-SVM with softmax (Hybrid scheme with Softmax) further promotes better results since it improves computational accuracy. The developed hybrid architecture gives much improved performance achieving higher accuracy of 99.3%. This proves the superior performance of developed hybrid methodology for high-level faults classification.

B. CONFUSION MATRIX FOR BS LINEAR DRIVE FAULTS

Confusion matrix for comparison of BS linear drive faults was computed using the trained network as given in Figure 17. The matrix shows fault accuracies for different predicted classes. It can be observed that the network model

TABLE 2. Performance comparison with mainstream techniques.

Comparison Technique	Accuracy
RBF-Support Vector Machine	84.6 %
LeNet	89.8 %
Adaptive Deep CNN	95.6 %
Hybrid scheme with Softmax	97.1 %
Proposed Hybrid Scheme	99.3 %

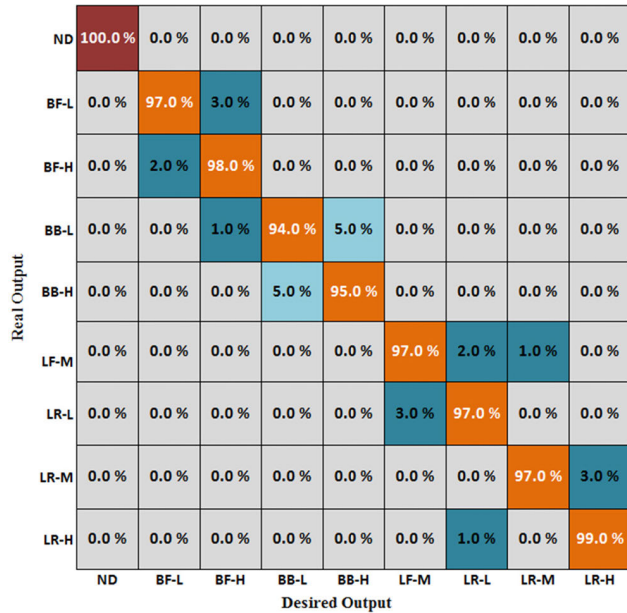


FIGURE 17. Confusion Matrix for BS linear drive faults.

is quite perfect for no defect (ND) class and gives 100% identification accuracy. BS friction (BF), LB friction (LF) and LB roller damage (LR) shows higher classification accuracy in comparison to BS backlash (BB). In general, BB faults have little bit higher confusion possibilities than with other fault modes. This implies BB faults are more complex to analyze and therefore need additional features information to avoid misclassification or false identification.

C. MODEL VALIDATION WITH IMS-UC BEARING DATASET

In addition to BS dataset, the developed network was also trained and validated by IMS-UC (Intelligent Maintenance Systems – University of Cincinnati) bearing dataset that include inner race (IR), outer race (OR) and rolling element (RE) faults data. The dataset were collected at 20 KHz sampling rate and 2000 RPM which gives 600 points per rotation. Continuous data samples were recorded for above 20K data points. 1360 data samples with 300 input points were used for model training and testing. The rolling element bearing considered for testing has 16 rollers with 8.4 mm roller diameter, 71.5 mm pitch diameter of bearing, 0.265 radians contact angle and 33.75 Hz rotational speed. Four

accelerometers were used to acquire vibration data. The same dataset was applied by Jagath et al. [57] and Eren et al. [58] for validation of their fault detection techniques. Results are compared in Figure 18. The work of Jagath et al. [57] provides early and real time faults identification for bearings using SVM. Time and frequency domain features were considered for SVM training. Eren et al. [58] apply their proposed adaptive 1-D CNN technique with 1224 training samples and 5440 testing samples (from 4 fault classes including no-fault data class). The proposed hybrid scheme was also trained and tested with similar data samples from each class. The results for each fault including inner race fault, outer race fault and rolling element fault are compared below. The classification accuracy results indicate more reliable and improved performance with developed model in comparison to other two techniques.

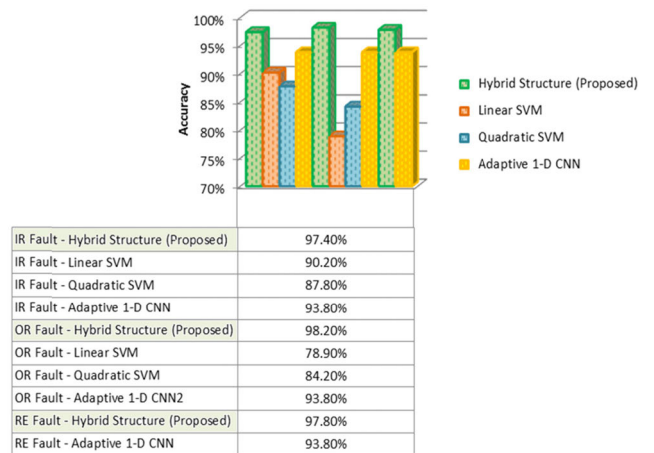


FIGURE 18. Performance comparison with IMS-UC bearing dataset.

VI. DISCUSSIONS

The combined hybrid model has achieved higher classification accuracy with improved feature extraction for desired diagnostic requirement. Typical characteristic aspects of our developed framework are discussed below.

A. SUPERIORITY OF RESIDUAL CONNECTIONS IN ResT CNN (1-D & 2-D CNN's)

The utilization of residual connections significantly improved the model accuracy and network training speed through identity mapping of input data. The stacked residual connections effectively minimized deep network training difficulties with better system performance. For ResT CNN, the comparison of model performance with and without residual connections is shown in Figure 19. It was found that the training data accuracy increases by an average of 5% using residual connections in 1-D and 2-D CNN's which indicate better learning rate of proposed architecture. Another important aspect is the improvement in performance accuracy of validation data which was increased by 7%. This shows that the overall performance of developed network is greatly

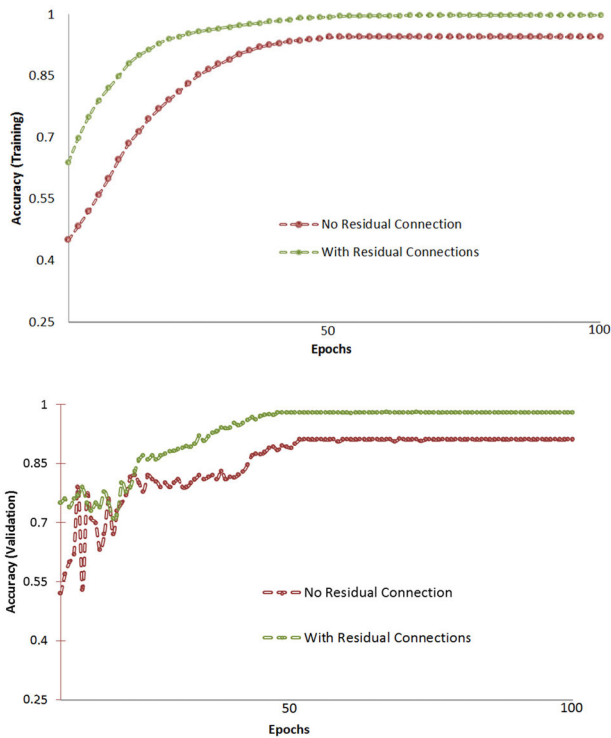


FIGURE 19. Accuracy Curves with and without Residual connections (a) Training (b) Validation.

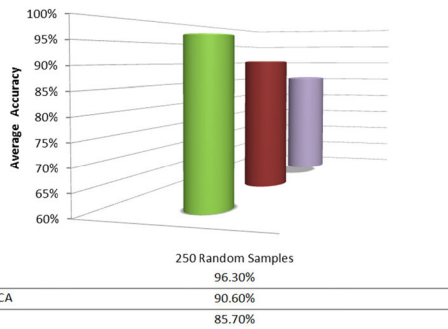


FIGURE 20. Accuracy comparison for PCA architectures.

improved with residual scheme by minimizing effects of network degradation. Additionally, it can be noticed that validation accuracy shows rapid convergence with residual architecture that explicitly indicates improved generalization and accurate fault classification of developed system with multi-class features. This illustrates that the proposed model gives better practical solution for BS linear drives faults classification.

B. EFFECTIVENESS OF KNOWLEDGE-BASED FEATURES SELECTION IN DEEP PCA (KB-REM-PCA)

The effectiveness of proposed knowledge-based PCA architecture was evaluated by computing average accuracies for proposed Kb-Rem-PCA with non-linear mapping, residual and deep PCA structures using typical parameters and G (·) activation function. BS linear drive dataset was considered for accuracy comparison. 250 random samples were selected (50 samples per class) from 5 classes that include high

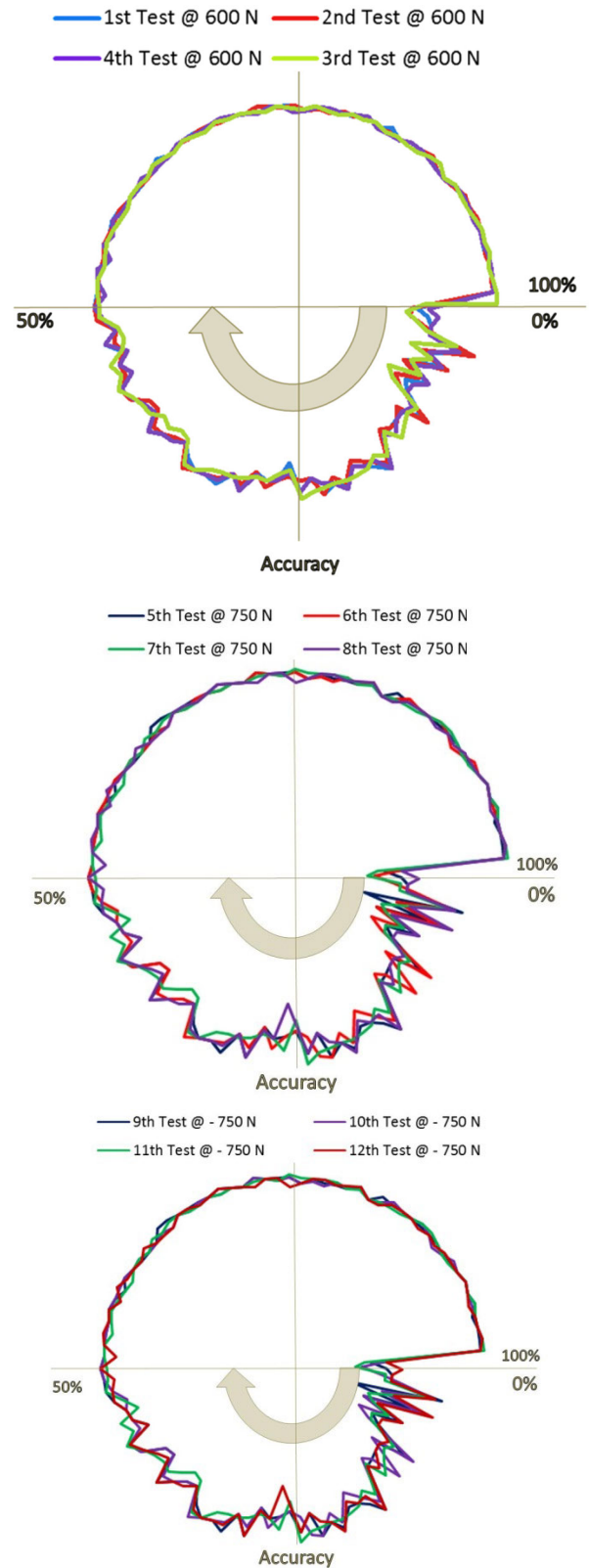


FIGURE 21. Performance against load variations (a) 600 N, (b) 750 N and (c) -750 N.

BS friction (BS-H), high BS backlash (BB-H), LB friction (LF-M) and high LB roller damage (LR-H). Results indicate

that Kb-Rem-PCA offers better features selection and quite stable behavior even with variation in input parameters settings. Although deep PCA shows good features discrimination, however there occurs loss of information in successive layers resulting in reduced accuracy especially when handling non-linear data. By adding residual connections in deep PCA, more effective features extraction is achieved than deep PCA. The addition of knowledge base rough set in combination with residual blocks significantly improves features classification accuracy by eliminating uncertain information from features. Kb-Rem-PCA, therefore, gives more promising classification results by removing data redundancy and uncertainty. Figure 20 gives outcome of this comparative analysis for 250 random sample points. The average accuracy for Kb-Rem-PCA has increased nearly 6% than with residual PCA structure. It can, therefore, be concluded that non-linear mapping with knowledge base structure minimizes redundancy and inaccuracies from PCA structure.

C. SYSTEM PERFORMANCE AGAINST LOAD PROFILE VARIATIONS

In addition to performance evaluation against rated load (600 N), the model was also studied for load variation in magnitude and direction. This include positive peak load (750 N) and negative peak load (-750 N) applied at attachment section of BS LMS under consideration. Multiple experiments were performed with above 3 load cycles in repetition. This evaluates system's performance for stability retention against different load domains. 12 experiments were performed for 3 load cases (4 tests per load cycle). Each load cycle was repeated 4 times in a sequence. These reversals of load cycles validate model's stability and adaptive performance to endure domain variations in profile. This scheme becomes very helpful for practical engineering problems and provides system response for correct classification at same computation time under variable load domains. From average accuracy results in Figure 21, it can be seen that in all experiments, the model retains high level of accuracy with average variance of ± 0.0025 (approx). The average computation time for each sample classification was found as 0.96 milli-sec. This ensures good network stability under varying load cycles with acceptable testing time which makes the proposed system suitable for real-time monitoring and fault detection of engineering systems.

VII. CONCLUSION

This paper presents a reliable and effective novel hybrid combination of collective techniques based on improvements in deep residual connections to identify and classify BS LMS faults. The major contributions include application of position measurement data for accurate detection of different BS linear drive faults including BS friction, BS backlash, LB friction and LB roller damage. ResT CNN structure was developed that comprises parallel learning of 1-D and 2-D improved deep residual CNN's followed by knowledge base Kb-Rem-PCA architecture for reduction of redundant data and multi-class SVM classifier. This composition greatly

improves system accuracy for typical features extraction and non-linear faults identification. The proposed methodology was also validated by IMS-UC published dataset. The results prove that the developed scheme not only gives effective extraction of characteristic features but also yields more systematic classification of faults with different severities and load variations. The accuracy results also show the superiority of proposed algorithm comparable with state of the art techniques.

This research can be extended to include some other failure modes of BS LMS elements to explore combined failure scenarios for more accurate system diagnosis. A comparison can also be done with fault signals from real failure case along with induced faults. This would be more helpful to predict BS drives performance with proposed scheme comparable with real complexities in actual framework.

REFERENCES

- [1] E. Balaban, P. Bansal, P. Stoelting, A. Saxena, K. F. Goebel, and S. Curran, "A diagnostic approach for electro-mechanical actuators in aerospace systems," in *Proc. IEEE Aerosp. Conf.*, Mar. 2009, pp. 1–13.
- [2] G. D. Rito and F. Schettini, "Health monitoring of electromechanical flight actuators via position-tracking predictive models," *Adv. Mech. Eng.*, vol. 10, no. 4, pp. 1–12, Apr. 2018, doi: [10.1177/1687814018768146](https://doi.org/10.1177/1687814018768146).
- [3] P. Phillips, D. Diston, and A. Starr, "Perspectives on the commercial development of landing gear health monitoring systems," *Transp. Res. C, Emerg. Technol.*, vol. 19, no. 6, pp. 1339–1352, Dec. 2011.
- [4] P. Phillips and D. Diston, "A knowledge driven approach to aerospace condition monitoring," *Knowl.-Based Syst.*, vol. 24, no. 6, pp. 915–927, Aug. 2011.
- [5] Y. H. Choi, S. M. Cha, J. H. Hong, and J. H. Choi, "A study on the vibration analysis of a ball screw feed drive system," *Mater. Sci. Forum*, vols. 471–472, pp. 149–154, Dec. 2004, doi: [10.4028/www.scientific.net/msf.471-472.149](https://doi.org/10.4028/www.scientific.net/msf.471-472.149).
- [6] C. S. Byington, P. Sloelting, M. Watson, and D. Edwards, "A model-based approach to prognostics and health management for flight control actuators," in *Proc. IEEE Aerosp. Conf.*, vol. 6, Dec. 2004, pp. 3551–3562.
- [7] M. J. Smith, C. S. Byington, M. J. Watson, S. Bharadwaj, G. Swerdon, K. Goebel, and E. Balaban, "Experimental and analytical development of health management for electro-mechanical actuators," in *Proc. IEEE Aerosp. Conf.*, Mar. 2009, pp. 1–14.
- [8] A. Chirico and J. R. Kolodziej, "Fault detection and isolation for electro-mechanical actuators using a data driven Bayesian classification," *SAE Int. J. Aerosp.*, pp. 494–502, Oct. 2012. [Online]. Available: <http://sae.aero.saejournals.org>, doi: [10.4271/2012-01-2215](https://doi.org/10.4271/2012-01-2215).
- [9] E. Balaban, A. Saxena, K. Goebel, C. S. Byington, S. Bharadwaj, M. Watson, and M. Smith, "Experimental data collection and modeling for nominal and fault conditions on electro-mechanical actuators," in *Proc. Annu. Conf. Prognostics Health Manage. Soc.*, 2009, pp. 1–5.
- [10] N. Saravanan and K. I. Ramachandran, "Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN)," *Expert Syst. Appl.*, vol. 37, no. 6, pp. 4168–4181, Jun. 2010.
- [11] H. Chen, Y. Sun, Z. Shi, and J. Lin, "Intelligent analysis method of gear faults based on FRWT and SVM," *Shock Vibrat.*, vol. 2016, pp. 1–9, Sep. 2016.
- [12] W. Jin, Y. Chen, and J. Lee, "Methodology for ball screw component health assessment and failure analysis," in *Proc. ASME Int. Manuf. Sci. Eng. Conf. (MSEC)*, Madison, WI, USA, vol. 2, Jun. 2013, doi: [10.1115/msec2013-1252](https://doi.org/10.1115/msec2013-1252).
- [13] A. Garinei and R. Marsili, "A new diagnostic technique for ball screw actuators," *Measurement*, vol. 45, no. 5, pp. 819–828, Jun. 2012, doi: [10.1016/j.measurement.2012.02.023](https://doi.org/10.1016/j.measurement.2012.02.023).
- [14] N. Riaz, S. I. A. Shah, F. Rehman, S. O. Gilani, and E. Udin, "A novel 2-D current signal-based residual learning with optimized softmax to identify faults in ball screw actuators," *IEEE Access*, vol. 8, pp. 115299–115313, 2020, doi: [10.1109/ACCESS.2020.3004489](https://doi.org/10.1109/ACCESS.2020.3004489).

- [15] G. Zhao, G. Zhang, Q. Ge, and X. Liu, "Research advances in fault diagnosis and prognostic based on deep learning," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Chengdu)*, Oct. 2016, pp. 1–6.
- [16] W. Zhang, X. Li, and Q. Ding, "Deep residual learning-based fault diagnosis method for rotating machinery," *ISA Trans.*, vol. 95, pp. 295–305, Dec. 2019, doi: [10.1016/j.isatra.2018.12.025](https://doi.org/10.1016/j.isatra.2018.12.025).
- [17] M. He and D. He, "Deep learning based approach for bearing fault diagnosis," *IEEE Trans. Ind. Appl.*, vol. 53, no. 3, pp. 3057–3065, May/June 2017, doi: [10.1109/tia.2017.2661250](https://doi.org/10.1109/tia.2017.2661250).
- [18] L. Jing, M. Zhao, P. Li, and X. Xu, "A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox," *Measurement*, vol. 111, pp. 1–10, Dec. 2017, doi: [10.1016/j.measurement.2017.07.017](https://doi.org/10.1016/j.measurement.2017.07.017).
- [19] M. Y. Asr, M. M. Etefagh, R. Hassannejad, and S. N. Razavi, "Diagnosis of combined faults in rotary machinery by non-naive Bayesian approach," *Mech. Syst. Signal Process.*, vol. 85, pp. 56–70, Feb. 2017.
- [20] Y. C. Tang, Y. Q. Zhang, N. V. Chawla, and S. Krasser, "SVMs modeling for highly imbalanced classification," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 39, no. 1, pp. 281–288, Feb. 2009.
- [21] X. Wang, Y. Zheng, Z. Z. Zhao, and J. P. Wang, "Bearing fault diagnosis based on statistical locally linear embedding," *Sensors*, vol. 15, no. 7, pp. 16225–16247, Jul. 2015.
- [22] G. Georgoulas, P. Karvelis, T. Loutas, and C. D. Stylios, "Rolling element bearings diagnostics using the symbolic aggregate approximation," *Mech. Syst. Signal Process.*, vols. 60–61, pp. 229–242, Aug. 2015.
- [23] L. Zhang, H. Gao, J. Wen, S. Li, and Q. Liu, "A deep learning-based recognition method for degradation monitoring of ball screw with multi-sensor data fusion," *Microelectron. Rel.*, vol. 75, pp. 215–222, Aug. 2017, doi: [10.1016/j.microrel.2017.03.038](https://doi.org/10.1016/j.microrel.2017.03.038).
- [24] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Dec. 2005.
- [25] F. Rehman, S. I. A. Shah, N. Riaz, and S. O. Gilani, "A robust scheme of vertebrae segmentation for medical diagnosis," *IEEE Access*, vol. 7, pp. 120387–120398, 2019, doi: [10.1109/ACCESS.2019.2936492](https://doi.org/10.1109/ACCESS.2019.2936492).
- [26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [27] Z. Liu, W. Guo, J. Hu, and W. Ma, "A hybrid intelligent multi-fault detection method for rotating machinery based on RSGWPT, KPCA and twin SVM," *ISA Trans.*, vol. 66, pp. 249–261, Jan. 2017, doi: [10.1016/j.isatra.2016.11.001](https://doi.org/10.1016/j.isatra.2016.11.001).
- [28] M. Ye, C. Ji, H. Chen, L. Lei, H. Lu, and Y. Qian, "Residual deep PCA-based feature extraction for hyperspectral image classification," *Neural Comput. Appl.*, vol. 32, no. 18, pp. 14287–14300, Sep. 2020, doi: [10.1007/s00521-019-04503-3](https://doi.org/10.1007/s00521-019-04503-3).
- [29] D. Peng, Z. Liu, H. Wang, Y. Qin, and L. Jia, "A novel deeper one-dimensional CNN with residual learning for fault diagnosis of wheelset bearings in high-speed trains," *IEEE Access*, vol. 7, pp. 10278–10293, 2019, doi: [10.1109/access.2018.2888842](https://doi.org/10.1109/access.2018.2888842).
- [30] P. Santos, L. Villa, A. Reñones, A. Bustillo, and J. Maudes, "An SVM-based solution for fault detection in wind turbines," *Sensors*, vol. 15, no. 3, pp. 5627–5648, Mar. 2015, doi: [10.3390/s150305627](https://doi.org/10.3390/s150305627).
- [31] N. Hatami, Y. Gavet, and J. Debayle, "Classification of time-series images using deep convolutional neural networks," *Sensors*, vol. 15, pp. 5627–5648, 2015. [Online]. Available: <http://mdpi.com/journal/sensors>, doi: [10.3390/s150305627](https://doi.org/10.3390/s150305627).
- [32] J. Wen and H. Gao, "Degradation assessment for the ball screw with variational autoencoder and kernel density estimation," in *Proc. 10th Int. Conf. Mach. Vis. (ICMV)*, vol. 10696, Bellingham, WA, USA: SPIE, Apr. 2018, Art. no. 106960Y, doi: [10.1117/12.2309486](https://doi.org/10.1117/12.2309486).
- [33] Y.-C. Huang, C.-H. Kao, and S.-J. Chen, "Diagnosis of the hollow ball screw preload classification using machine learning," *Appl. Sci.*, vol. 8, no. 7, p. 1072, Jun. 2018, doi: [10.3390/app8071072](https://doi.org/10.3390/app8071072).
- [34] L. M. R. Baccarini, B. R. D. Menezes, V. V. R. Silva, and W. M. Caminhas, "SVM practical industrial application for mechanical faults diagnostic," *Expert Syst. Appl.*, vol. 38, no. 6, pp. 6980–6984, Jun. 2011, doi: [10.1016/j.eswa.2010.12.017](https://doi.org/10.1016/j.eswa.2010.12.017).
- [35] X. Cui, V. Goel, and B. Kingsbury, "Data augmentation for deep convolutional neural network acoustic modeling," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 4545–4549.
- [36] Z. Pawlak, "Rough sets," *Int. J. Inf. Comput. Sci.*, vol. 11, no. 5, pp. 341–356, 1982.
- [37] X. Deng, X. Tian, S. Chen, and C. J. Harris, "Deep learning based nonlinear principal component analysis for industrial process fault detection," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 1237–1243, doi: [10.1109/ijcnn.2017.7965994](https://doi.org/10.1109/ijcnn.2017.7965994).
- [38] V. E. Liang, J. Lu, and G. Wang, "Face recognition using deep PCA," in *Proc. 9th Int. Conf. Inf., Commun. Signal Process.*, Dec. 2013, pp. 1–5, doi: [10.1109/iccisp.2013.6782777](https://doi.org/10.1109/iccisp.2013.6782777).
- [39] Z. Pawlak, *Rough Sets—Theoretical Aspects of Reasoning About Data*. Dordrecht, The Netherlands: Kluwer Academic, 1991.
- [40] X. Chen, X.-M. Wang, Y. Huang, and X.-Y. Wang, "Fault diagnosis for tilt-rotor aircraft flight control system based on variable precision rough set-OMELM," *Kongzhi yu Juece/Control Decis.*, vol. 30, no. 3, pp. 433–440, 2015, doi: [10.13195/j.kzyjc.2013.1558](https://doi.org/10.13195/j.kzyjc.2013.1558).
- [41] S. U. Jan, Y.-D. Lee, J. Shin, and I. Koo, "Sensor fault classification based on support vector machine and statistical time-domain features," *IEEE Access*, vol. 5, pp. 8682–8690, 2017, doi: [10.1109/ACCESS.2017.2705644](https://doi.org/10.1109/ACCESS.2017.2705644).
- [42] X. Lou and K. A. Loparo, "Bearing fault diagnosis based on wavelet transform and fuzzy inference," *Mech. Syst. Signal Process.*, vol. 18, no. 5, pp. 1077–1095, 2004.
- [43] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mech. Syst. Signal Process.*, vol. 100, pp. 439–453, Feb. 2018, doi: [10.1016/j.ymssp.2017.06.022](https://doi.org/10.1016/j.ymssp.2017.06.022).
- [44] Z. Yuan, L. Zhang, L. Duan, and T. Li, "Intelligent fault diagnosis of rolling element bearings based on HHT and CNN," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Chongqing)*, Oct. 2018, pp. 292–296.
- [45] Y. Hsueh, V. R. Ittangihala, W. B. Wu, H. C. Chang, and C. C. Kuo, "Condition monitor system for rotation machine by CNN with recurrence plot," *Energies*, vol. 12, no. 17, p. 3221, Aug. 2019, doi: [10.3390/en12173221](https://doi.org/10.3390/en12173221).
- [46] Y. Li, L. Zou, L. Jiang, and X. Zhou, "Fault diagnosis of rotating machinery based on combination of deep belief network and one-dimensional convolutional neural network," *IEEE Access*, vol. 7, pp. 165710–165723, 2019, doi: [10.1109/ACCESS.2019.2953490](https://doi.org/10.1109/ACCESS.2019.2953490).
- [47] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018, doi: [10.1109/TIE.2017.2774777](https://doi.org/10.1109/TIE.2017.2774777).
- [48] H. Li, J. Huang, and S. Ji, "Bearing fault diagnosis with a feature fusion method based on an ensemble convolutional neural network and deep neural network," *Sensors*, vol. 19, no. 9, p. 2034, Apr. 2019.
- [49] S. Guo, T. Yang, W. Gao, and C. Zhang, "A novel fault diagnosis method for rotating machinery based on a convolutional neural network," *Sensors*, vol. 18, no. 5, p. 1429, May 2018, doi: [10.3390/s18051429](https://doi.org/10.3390/s18051429).
- [50] T. M. Lai, L. A. Snider, E. Lo, and D. Sutanto, "High-impedance fault detection using discrete wavelet transform and frequency range and RMS conversion," *IEEE Trans. Power Del.*, vol. 20, no. 1, pp. 397–407, Jan. 2005.
- [51] M. Azad, F. Khaled, M. I. Pavel, "A novel approach to classify and convert 1d signal to 2d grayscale image implementing support vector machine and empirical mode decomposition algorithm," *Int. J. Adv. Res.*, vol. 7, no. 1, pp. 328–335, Jan. 2019.
- [52] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. New York, NY, USA: Academic, 1999.
- [53] Z. K. Peng and F. L. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: A review with bibliography," *Mech. Syst. Signal Process.*, vol. 18, no. 2, pp. 199–221, Mar. 2004, doi: [10.1016/s0888-3270\(03\)00075-x](https://doi.org/10.1016/s0888-3270(03)00075-x).
- [54] J. Lin and L. Qu, "Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis," *J. Sound Vibrat.*, vol. 234, no. 1, pp. 135–148, Jun. 2000, doi: [10.1006/jsvi.2000.2864](https://doi.org/10.1006/jsvi.2000.2864).
- [55] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *J. Sound Vib.*, vol. 289, pp. 1066–1090, Feb. 2006.
- [56] X. Guo, L. Chen, and C. Shen, "Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis," *Measurement*, vol. 93, pp. 490–502, Nov. 2016, doi: [10.1016/j.measurement.2016.07.054](https://doi.org/10.1016/j.measurement.2016.07.054).
- [57] J. S. L. Senanayaka, S. T. Kandukuri, H. Van Khang, and K. G. Robbersmyr, "Early detection and classification of bearing faults using support vector machine algorithm," in *Proc. IEEE Workshop Electr. Mach. Design, Control Diagnosis (WEMDCD)*, Apr. 2017, pp. 250–255, doi: [10.1109/wemdc.2017.7947755](https://doi.org/10.1109/wemdc.2017.7947755).
- [58] L. Eren, T. Ince, and S. Kiranyaz, "A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier," *J. Signal Process. Syst.*, vol. 91, no. 2, pp. 179–189, Feb. 2019, doi: [10.1007/s11265-018-1378-3](https://doi.org/10.1007/s11265-018-1378-3).



NAVEED RIAZ was born in Pakistan. He received the bachelor's and master's degrees in engineering from the National University of Sciences And Technology (NUST), Islamabad. He has more than 11 years of vast experience in the research industry and academics. His research area includes robotics, deep learning, medical image analysis, computer vision, artificial intelligence, robot controls, and aerial robotics. He has a lot of conference and journal research publications in this research area.



SYED IRTIZA ALI SHAH was born in Pakistan. He graduated from the Georgia Institute of Technology, Atlanta, GA, USA, and received the B.E. degree from NED University and the M.S. degree in aerospace engineering from NUST in fluid dynamics, and the M.S. degree in flight mechanics and controls, the M.S. degree in machine vision and robot controls, and the Ph.D. degree in aerial robotics from GeorgiaTech, Atlanta, GA, USA.



FAISAL REHMAN was born in Islamabad, Pakistan. He received the B.S., M.S., and Ph.D. degrees in engineering from NUST Islamabad. He worked in both industry and academics. He also has research experience in various research organizations as a junior and senior researcher. His research area is deep learning, medical image analysis, and artificial intelligence. He has a lot of conference and journal research publications in his research field.



MUHAMMAD JAWAD KHAN received the B.E. and M.S. degrees in mechatronics engineering from Air University, Pakistan, and the Ph.D. degree in mechanical engineering from Pusan National University, South Korea, in 2018. He worked on brain-robot interfaces and control of haptic device during his Ph.D., where he published over 40 papers. His research interests include hybrid brain-computer interfaces, brain signal processing and control, AI, machine learning, computer vision, and rehabilitation.

...