Received May 30, 2020, accepted June 20, 2020, date of publication June 23, 2020, date of current version July 2, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.3004489*

# A Novel 2-D Current Signal-Based Residual Learning With Optimized Softmax to Identify Faults in Ball Screw Actuators

**NAVEED RIAZ[1], SYED IRTIZA ALI SHAH[2], FAISAL REHMAN[3], SYED OMER GILANI[3], AND EMAD UDIN[1]**

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

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

[3]Department of Robotics and 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** Ball screw electro-mechanical actuators are commonly found in high precision motion control applications including aerospace systems as well as automated setups for industries. These actuators perform flight / application critical job and ball screw drives are responsible to provide precise linear motion while carrying thrust loading. A failure in ball screw drive may disturb positioning accuracy of overall system. At present, few techniques are available to monitor electro-mechanical actuators for aerospace and industrial systems. This paper provides a deep learning based intelligent technique to monitor condition of ball screw actuators. The proposed scheme utilizes modified residual learning scheme to extract features from two-dimensional transformed motor current signals. The current signal data was collected under different load domains in terms of magnitude and direction reversal. A 2D-Remanant-CNN (2D-Rem-CNN) model was developed for features extraction with proposed optimized softmax for classification of mechanical faults. The proposed technique was validated against different ball screw fault cases. The testing results prove the superiority of 2D-Rem-CNN model against different state of the art techniques. The proposed framework was also tested for system's stability under different load domains.

**INDEX TERMS** Electro-mechanical actuators, intelligent technique, motor current signals, positioning accuracy, residual deep learning.

## I. INTRODUCTION

Linear electro-mechanical actuator (LEA) is one of the best choices for precise motion control applications and covers broad range of actuation requirements in manufacturing and aerospace industries. Ball screw drive is the most critical element in LEA. Ball screw driven actuators, in comparison with other linear drives, offer higher efficiency, better positioning accuracy and less backlash problems. Different electromechanical systems including aircraft steering and surface control, robotic arms, feed drives etc., demand high precision motion control. The functionality, availability and safe working of these systems may severely affect due to degradation in their actuators. In the case of aircraft control surfaces, mechanical friction (jamming) or dead-band (backlash) of actuators may become flight critical as the case of 261 Alaska Airline freezing stabilizer, which caused loss of vertical control of airplane due to jack-screw excessive wear and became the reason of plane crash [1]. Furthermore, excessive backlash reduces ball screw stiffness and ultimately lead to loss of positional accuracy. This indicates the importance of early identification of actuator faults to avoid unexpected failure of entire system.

Linear electro-mechanical actuators work in combination of electric motor with necessary speed reduction gear-train along with rack–pinion or screw–nut arrangement as shown in Figure 1. Linear actuators comprising of ball-screw drive have gained huge attention by researchers and industries for automation and precision control systems. The failure and degradation analysis in these systems has still not studied

The associate editor coordinating the review of this manuscript and approving it for publication was Victor Hugo Albuquerque.

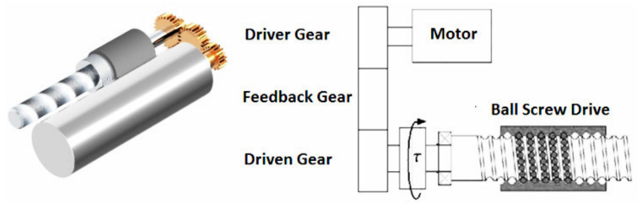


**FIGURE 1.** **Schematics of ball screw electric actuator.**

in detail, however critical failure modes and their effects for ball-screw systems have been studied and available [2]. Different mechanical faults that occur more frequently and become more severe in these systems include inadequate lubrication, friction, mechanical backlash, structural faults (wear, cracks, spalling) and jammed return channel.

Linear electro-mechanical actuators work in combination of electric motor with necessary speed reduction gear-train along with rack–pinion or screw–nut arrangement as shown in Figure 1. Linear actuators comprising of ball-screw drive have gained huge attention by researchers and industries for automation and precision control systems. The failure and degradation analysis in these systems has still not studied in detail, however critical failure modes and their effects for ball-screw systems have been studied and available [2]. Different mechanical faults that occur more frequently and become more severe in these systems include inadequate lubrication, friction, mechanical backlash, structural faults (wear, cracks, spalling) and jammed return channel.

The increased utilization of ball-screw actuators in critical aerospace applications needs more knowledge of possible fault modes since these actuators have to perform critical mission. The condition assessment of aerospace ball screw drive is quite challenging due to complex dynamics of these systems. Limited condition maintenance tools are available for these systems due to variation in operating conditions, mechanical inaccuracies and non-consistent operation. This demands customized monitoring setup with improved sensors; and therefore requires expertise in electro-mechanical systems and their control interface. This typically becomes a high cost solution.

Another simple approach is to fix a suitable sensor (like force sensor) to measure ball-nut force. This indicates preload loss which gives a measure of increased backlash. This method, however, is impracticable due to high cost and difficulty in sensors installation. Moreover, the data from sensors inside the linear actuators are limited for servo control purpose which becomes a constraint to implement this methodology for monitoring.

Apart from these limitations, a number of studies have been done recently for electric linear actuators in automation industries. The application of linear actuators in aerospace sectors also carries big potential. The concept of ''Fly Electric'' considers more to all electric flight in terms of electric powered flight along with control surfaces operation, steering nose control, and retractable landing gear; all controlled with electro-mechanical actuators. The data monitoring system for these actuators is being the primary concern for researchers [3], [4]. For these systems, actuator data should be monitored after installing in aircraft to ensure system reliability related to acquired data [5] to be as close as compared with conventional systems (hydraulics). To be specific, ball-screw drive actuators are the most common electric actuators explored in many experimental and simulation research studies; not only in comparison to hydraulic actuators [6], [7] as well as by installing it on actual aircrafts.

This research proposes deep learning based intelligent technique to diagnose mechanical faults for ball-screw driven linear electro-mechanical actuators using current signal data. The theme of this paper is to explore data monitoring based on significant changes in current signal features. These remarkable signal features indicate system's dynamic behavior due to any noticeable change in its performance. For improved representation of signal features and their extraction, a 2D-Remnant Convolutional Neural Network (2D-Rem-CNN) fault classifier is proposed that comprises modified residual learning blocks along with two dropout and two fully-connected layers. Finally, an optimized softmax is applied for improved faults classification.

The contribution of our work in this paper is summarized below:

a) Current signal data was collected for normal and fault conditions. Different faults were induced in ball screw drive to study their effects.
b) A data pre-processing scheme is suggested which transforms 1-D current signal into 2-D gray scale image to make use of two-dimensional features.
c) Improved domain adaptability and performance is achieved by 2D-Rem-CNN with two Dropout and two fully connected layers in addition to modified residual blocks.
d) Improved multi-class boundary discrimination is achieved using proposed optimized softmax.
e) The developed approach successfully deals with different fault modes in ball screws.
f) The formulated scheme gives well competitive results comparable with state of the arts.

The remaining paper is organized as follows: Section II includes brief description of previous works. Section III gives detail description of our proposed methodology. Section IV provides experimentation details in including testing setup details and description of faults cases. Section V presents the model results followed by detailed discussion in section VI.

## II. BRIEF DESCRIPTION OF RELATED WORK

Most of the techniques developed previously rely on engineering modeling of actuator [8]. Health assessment of electro-mechanical actuators using analytical and experimental methodology was considered [9]. System level analysis and prognostics was considered for failure prevention [10] and numerical methods were used to solve complex models [11]–[14]. These techniques were although successful, however the limited application of these approaches due to characteristic models and requirement of dedicated model developer of each system or mechanism make them ineffective. Other techniques utilize measurement of physical variables like temperature or vibration using suitable sensors to acquire and process signal data and apply some feature extraction tools to identify fault signature [15], [16]. For electric actuators, motor faults have been previously studied in detail [17] and are not considered in this work.

At present, different fault diagnostic techniques are in practice using machine learning and deep learning approaches to identify as well as classify faults in different mechanical systems [18], [19]. The intelligent fault identification approach works by collecting signal data from installed sensors and process the signal data automatically to identify system condition. Many research studies have been done for different intelligent fault detection techniques. A feature extraction technique for combined faults in rotating systems was developed on the basis of Empirical mode decomposition. Bayesian classifier was applied for intelligent fault detection of extracted signal features [20]. Other contributions include faults classification based on nearest neighbor classifier [21], [22] and vibration signals analysis using Support vector machine [23].

Compared with machine learning techniques mentioned above, deep learning techniques perform better results, still its application in this area is in developing phase. Among deep learning algorithms, Convolutional Neural Network (CNN) is one of the most dominant classifiers especially when working with complex features [24], [25].

There are two main concerns regarding intelligent fault detection of mechanical systems. First, the fault signal data collected from mechanical systems is difficult to acquire than from normal system operation. This means for normal operation; data can easily be collected while for fault modes inadequate data is available. This causes imbalance distribution of system health data which may bias the classifier towards major health data [26]. This ultimately causes inadequate system learning towards minor acquired data and thus gives incorrect fault diagnosis.

Second, there are different fault diagnostic methods available but small number of ways exist to process time-based signal data directly. Also, same classifier is used by different methods like Back propagation neural network and Support vector machine. Fault classifiers may also sometimes fail for correct classification of samples due to poor domain adaptability.

This paper deals with data monitoring of Ball screw (LEA's) using signal features from current data. These signal features represent change in system dynamics due to any impending fault. Improved features extraction and classification is achieved using proposed 2D-Rem-CNN and optimized softmax as explained in next section.

## III. PROPOSED METHODOLOGY

Actuator mechanical faults like those stated in the last sections may disturb system operation and degrade the performance of entire system which can be analyzed by measuring different variables like motor current signature, motion profile, vibration signals etc. For this specific precise application, actuator needs to accelerate and decelerate continuously in each motion cycle from one extreme to other. In different fault modes, the measured parameter shows variation in certain features of signal data which are controlled by system dynamics and controller behavior like signal overshoots and other observable peaks at position or direction change etc. Any variation in these signal features may indicate undesirable change in system mechanical behavior due to different impending faults.

In this technique, time-based 1-D current signal data is transformed in to 2-D textured image that explores certain two dimensional features from signal data and thus time-based signal classification can be analyzed as image feature recognition [27].

A testing setup was developed to collect signal data for normal and fault scenarios as discussed in Section IV. This generated dataset was re-shaped into 2-D image using proposed signal to image conversion approach to identify and classify different faults with Convolutional Neural Network (CNN) algorithm.

### A. PROPOSED 1-D TO 2-D SIGNAL CONVERSION METHOD

Various approaches have been used for signal transformations including continuous wavelet transform (CWT) [28], Recurrence plots which transform time signal data into 2D texture images [29], Hilbert-Huang transform [30] and combination of different transforms. This paper presents an effective approach for signal data pre-processing that converts 1-D time domain raw signal data into 2-D gray image. This approach gives an opportunity to find 2-D features from acquired raw signal. This signal to image conversion approach doesn't need any preset parameter.

This signal transformation approach works by successively filling the image pixels with time domain signal data. A random segmented sample of length $M^2$ is selected from time domain signal data as shown in Figure 2. 2-D image is obtained with $M \times M$ pixel points by processing and ordering samples. By normalizing signal segments from 0 to 255, the intercepted segments of signal give the pixel values of gray level image. $224 \times 224$ image size is selected determined by signal data volume. K (m); where m = 1, 2, 3..., $M^2$ represents signal segmented values. P (m, n) where, m = 1, 2, 3..., M and n = 1, 2, 3..., M gives pixel strength values of generated image. The above process is defined in mathematical notation as:

$$P(m,n) = \frac{Max\ K - K(nM+m)}{Max\ K - Min\ K} * 255 \tag{1}$$

### B. PROPOSED FEATURES EXTRACTION & CLASSIFICATION ARCHITECTURE

#### 1) PROPOSED 2D-REM-CNN ARCHITECTURE

Convolutional neural network (CNN) is one of the most common solutions for effective faults identification in

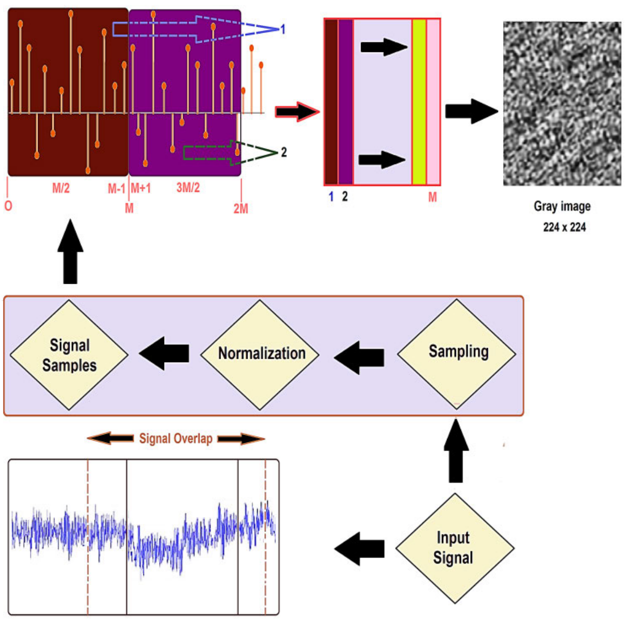


**FIGURE 2.** 1-D Signal to 2-D signal (image) conversion.

practice today. CNN includes multiple filtering and classification layers to extract desired features from input data and classify it accordingly. Filtering stage constitutes convolutional and pooling layers. Convolutional layer contains multiple kernels that give new feature maps when input is convolved with convolutional kernels sensitive to specific feature. The purpose of pooling layer is to reduce spatial size without decreasing number of feature maps. The classification stage includes fully connected layer which is a multi-stage perceptron in which high level features are extracted and taken as input to the output layer where CNN model output is generated. Mathematically, convolution process can be expressed as;

$$x^{l(i,j)} = K_i^l * m^{l(n^j)} = \sum_{j'=0}^{W} K_i^l\left(j'\right) m^{l(j+j')} \tag{2}$$

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

The general model [32] can be expressed as:

$$CNN_{out} = \begin{bmatrix} P\left(y=1\,|\,x;\,W_1, b_1\right) \\ P\left(y=2\,|\,x;\,W_2, b_2\right) \\ \cdots \\ P\left(y=m\,|\,x;\,W_m, b_m\right) \end{bmatrix} = \frac{1}{\sum_{j=1}^{k} \exp\left(W_{jx}+b_{j}\right)} \begin{bmatrix} \exp\left(W_1x+b_1\right) \\ \exp\left(W_2x+b_2\right) \\ \cdots \\ \exp\left(W_kx+b_k\right) \end{bmatrix} \tag{3}$$

where, $b$ is the bias neuron of each layer. More literature details can be found in [29]–[32].

The above mentioned CNN model has been used in different combinations to solve problems. However more deep networks increase computation time and create over-fitting along with decreased network accuracy. To solve these problems, residual blocks were added by Kaiming He [25] in deep networks as shown in Figure 3.

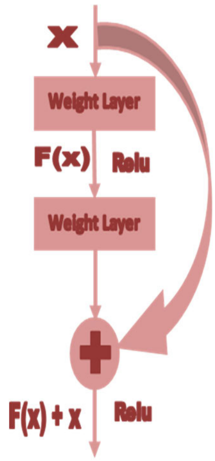


**FIGURE 3.** Original residual learning block.

The 2D-Rem-CNN architecture proposed in this paper is inspired from residual learning block as shown above. A much deeper network is established to identify ball screw actuator faults for critical operating conditions by improving network performance and learning ability. We have added modified residual connections. Two residual connections are added in our network. Each residual connection comprises $3 \times 3$ convolutional layer sandwiched between convolutional layers having size $1 \times 1$.

To enhance classification accuracy and decreasing the effects of high frequency noise, wide ($9 \times 9$) size first convolutional layer is considered. Remaining multi-layer convolutional kernels having size $3 \times 3$ give better representation of input data and makes the network deeper. Max Pooling layer is used at the end of residual connections to avoid over-smoothing of image data which may give error in extracting desired features. ReLU (Rectified Linear Unit) activation function was applied in each CNN stage to develop non-linearity. Moreover, we have used two Dropout and two Fully Connected layers in classification stage to improve structure adaptability.

Mathematically, assuming *Res* $(x)$ representing residual learning function, the output $O(res)$ of residual blocks can be given by,

$$\begin{aligned} O\left(res\right) &= f_r\left(W_1.x + O_{res1} + m\left(x\right) + W_2 + O_{res2}\right) \\ &= f_r\left(W_1.x + O_{res1} + m\left(x\right) + W_1.O'_{res1} + b_2 + m\left(x\right)\right) \end{aligned} \tag{4}$$

where, $f_r$ is the non-linear ReLU activation function, $O_{res1}$ is the output signal feature of first residual block obtained by sliding convolutional kernel size $3 \times 3$ with selected signal stride. The output of first residual block was added to the identity mapping $m(x)$ and passed through the ReLU activation function.

The second filter layer weight $W_2$ was obtained by output of first layer $O'_{res1}$ to obtain final output of residual structure. The general architecture of proposed residual block is shown in Figure 4.

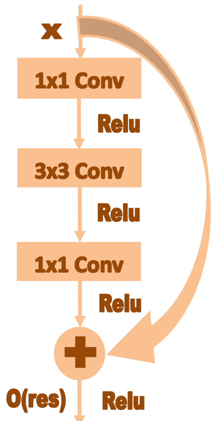


**FIGURE 4.** **Residual block architecture for proposed network.**

For output signal feature *x*, where $x \in \{x1, x2\ldots xk\}$; *k* being signal length, the data series *x* moves in the neural network without deterioration and learns residual function *Res(x)* more easily, providing better learning of CNN and improves system performance. This composition of CNN structure becomes deeper than CNN models previously existed. The remaining architecture is same as general CNN structure.

### C. PROPOSED SOFTMAX CLASSIFICATION

Softmax regression is widely employed for multi-class faults classification [35], [36] as it assures better performance and classification results with improved computational accuracy [33]. However, traditional Softmax loss is not capable for effective features classification due to biased sample distribution leading to misclassification. Furthermore, overfitting and noise problems also degrade performance classification. To overcome these problems, various improvements in Softmax classification have been proposed; for instance, Large margin (LM) softmax loss [37], Additive margin (AM) softmax loss [38] and Adaptive softmax [39].

For multi-class problems, softmax regression is utilized to evaluate the fault probabilities as mentioned in Equation 3 described earlier. Class 1 probability label will be assigned for $p_1 > p_2$; class 2 for $p_2 > p_1$. The decision boundary can be given by, $(W_1 - W_2)x + b_1 - b_2 = 0$; whereas, the classification results will be $W^T{}_1 x + b_1$; $W^T{}_1 x + b_1$. After weights normalization and zero biases with angular parameters, the classification results can be represented by, $p_1 = ||x||\cos\theta_1$; $p_2 = ||x||\cos\theta_2$. It can be seen that both prediction labels share same feature $x$ and result classification depends on angles $\theta_1$ and $\theta_2$ only. Moreover, for multi-class problems, above binary class case can be generalized. The improved softmax loss converge $k^{th}$ class features to smaller cosine angle in comparison to others which makes it a reliable parameter [37].

The traditional softmax loss is given by,

$$SL_{Orig} = \frac{1}{N}\sum\nolimits_i L_i \tag{5}$$

$$L_i = -\log\left(\frac{e^{f_{yi}}}{\sum_j e^{f_{yj}}}\right)\frac{1}{N} \tag{6}$$

where, $N$ shows training samples and $f_j$ shows $j^{th}$ element from [1, K] fault class of score vector $f$ which is generally output of FC layer. This gives $f_j = W_j^T x_i + b_j$. After the addition of angular boundaries with other improvements, we get modified softmax loss as given below,

$$SL_{Mod} = \frac{1}{N}\sum_i -\log\left(\frac{e^{||x_i||\cos\theta_{yi},i}}{\sum\limits_j e^{||x_i||\cos\theta_{yi},i}}\right) \tag{7}$$

The angular boundary parameters as given above with modified softmax loss successfully learn features, still these features show indiscrimination. The decision boundaries can be further manipulated by adding some marginal parameters that will improve discrimination ability. This can be achieved by introducing suitable scaling factor with angular parameter "$\theta$" which will boost classification ability by expanding angular margin. This scaling factor will improve classification in a balanced way.

The modified softmax loss for class 1 fault label requires $\cos\theta_1 > \cos\theta_2$ for accurate classification. By introducing scaling factor $\cos(2\sqrt{b}\,\theta_1) - \sqrt{b} > \cos\theta_2$ where b is an integral value greater than 3 gives more precise classification decision since it assures $\cos\theta_1 > \cos\theta_2$ in all cases. Formulating this concept in Equation 7 gives,

$$SL_{Opt} = \frac{1}{N}\sum_i - \log\left(\frac{e^{||x_i||\cos(2\sqrt{b}\theta_{yi}-\sqrt{b}),i}}{e^{||x_i||\cos(2\sqrt{b}\theta_{yi}-\sqrt{b}),i} + \sum\limits_{j\neq y_i} e^{||x_i||\cos(2\sqrt{b}\theta_{yi}-\sqrt{b}),i}}\right) \tag{8}$$

The above composition gives optimized angular margins and follows separate decision boundary specific for each class.

A schematic of 2D-Rem-CNN architecture is shown in Figure 5.

### D. PROPOSED FRAMEWORK

A basic outline for our proposed framework is shown in Figure 6. The flowchart gives the sequence of actions for classification of faults in ball-screw drive as summarized below.

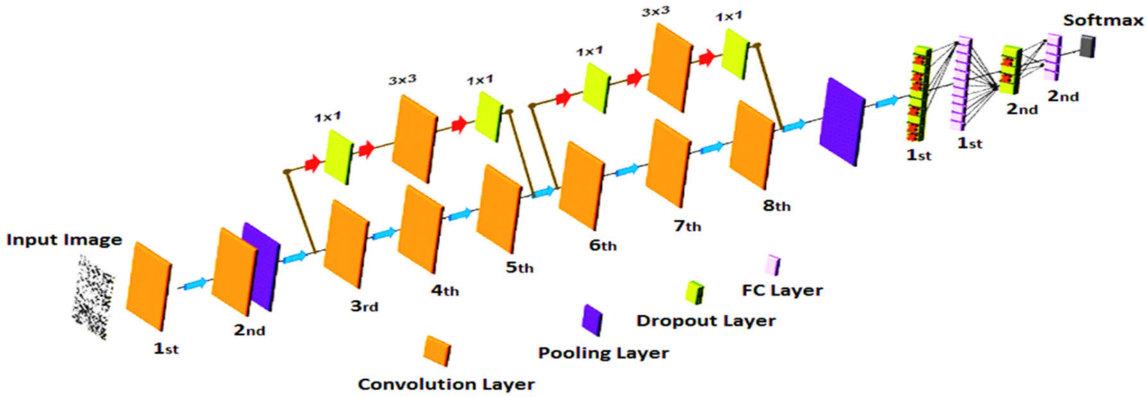


**FIGURE 5.** 2D-Rem-CNN architecture with coupled residual connection and Dropout & FC layers.

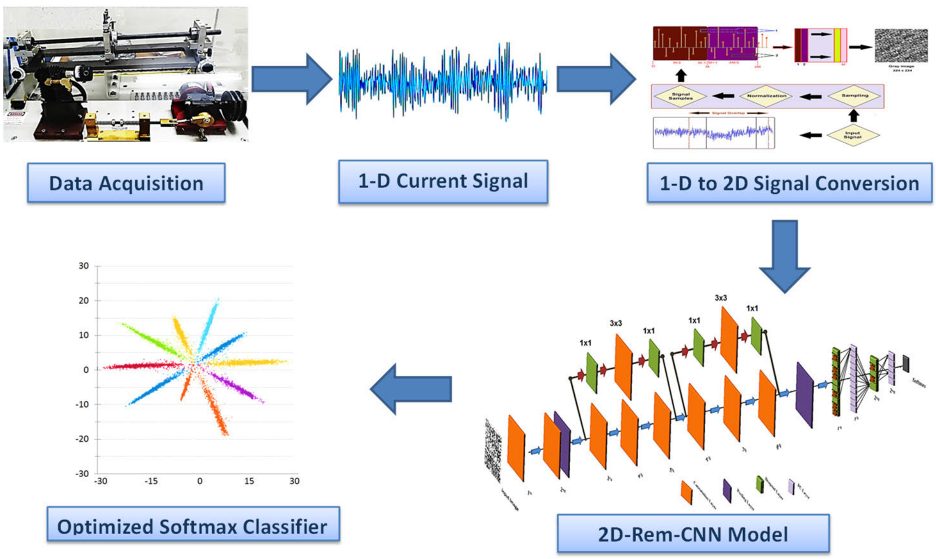


**FIGURE 6.** Proposed framework outline.

a) Measure ball screw current signal data for no-fault and different fault conditions to define signal patterns for each case.
b) Perform 1-D signal to 2-D gray image transformation using proposed technique. Split the data into training and testing data set.
c) Develop 2D-Rem-CNN and execute feature learning in successive layers.
d) Execute fault classification using optimized *Softmax* classifier to verify the proposed model.

## IV. EXPERIMENTATION

The performance of proposed algorithm was evaluated on developed instrumented setup. The testing setup was based on ball screw drive mechanism where ball screw shaft provides necessary channel for ball bearings retained inside suiting nut. A number of distinctive faults were induced in ball screw drive under varying loading conditions to consider the effect of these faults on measured system parameters as well as to analyze the acquired signal for detection of faults.

### A. EXPERIMENTAL TESTING SETUP DETAILS

The experimental setup was designed and built using available mechanical components including precision rolled miniature recirculation ball-screw drive (that consists of ball-screw threaded shaft fitted with ball-nut). Ball-nut was provided with threaded collar for attachment in assembly. Other components include motor with necessary gear train along with position feedback device, bearings and linear miniature guides etc.

The mechanical components along with structural parts were assembled on a base structure made of aluminum. The testing setup was also equipped with multiple sensors for data acquisition and control purpose. The bandwidth, repeatability and resolution of sensors should be well analyzed for satisfactory data acquisition. Linear current sensor with rated

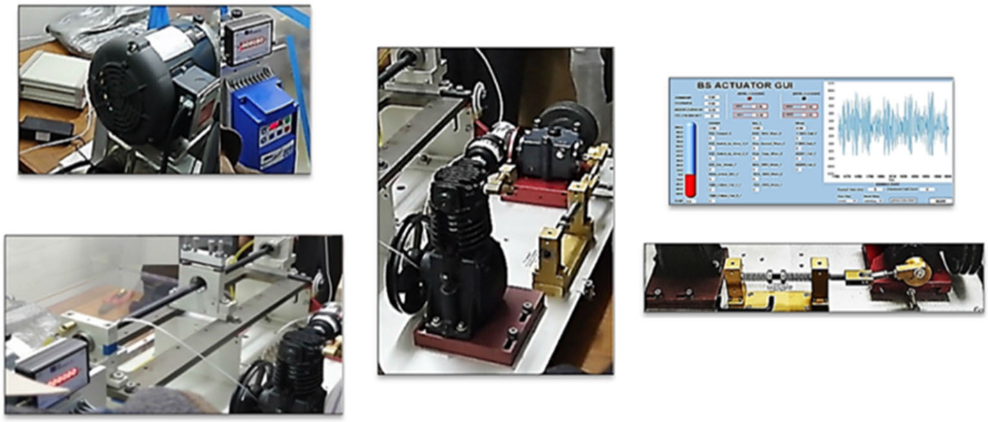


**FIGURE 7.** Experimental test setup to simulate faults.

sensitivity of 32.7 mV @ 12V DC, Linear transducer having linearity error 0.05 % and resistance of ± 20 % and a tension-compression load cell having 0.02 % error of rated output were also integrated for data monitoring purpose.

The developed testing setup was able provide 100mm linear travel (both directions) with load capacity of 1000N (positive and negative). Figure 7 illustrates some details of experimental setup with integrated sensors and ball screw.

### *B. FAULT CASES*

In the first phase, the testing setup was operated free of induced faults to get significant information and observe normal system behavior under varying motion and load profiles. This data was used as reference data with acceptable mechanical behavior. In the second phase, a number of different faults were induced in different areas of the system and was run to collect significant signal data. This data was used to examine health of actuator damaged due to these faults and to evaluate the performance of developed monitoring technique. The faults were selected based on high criticality and repetitive nature obtained from previous history and literature. The faults induced and tested include the following.

→ Mechanical friction (Inadequate lubrication) fault, developed by slowly ceasing the amount of lubricant and pressing firmly the ball nut seals (as shown in Figure 8).
→ Mechanical backlash fault, developed by substituting original recirculating balls with undersize balls of different sizes (as shown in Figure 9).
→ Fatigue (pitting or flaking) fault, developed by splitting of material surface (having varying sizes) from ball screw and balls (as shown in Figure 10).

Ball screw drive takes sufficient time to undergo natural failure, so the aforementioned faults were gradually induced in the system to understand the system behavior and capability to diagnose faults as well as to analyze the extracted signal data (features).

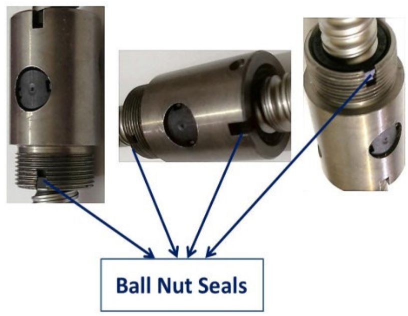


**FIGURE 8.** Mechanical friction by tightening ball nut seals.

**FIGURE 9.** Mechanical backlash fault by changing recirculating balls size.

In case of adequate lubrication fault, initially the lubricant inside ball nut was removed using suitable degreaser. The signal data was observed before and after removing grease. Due to the characteristic low friction of ball screw drive, no considerable signal changes were found. To increase the fault (friction) intensity, end sealing locks of ball nut were pressed tightly (Figure 8). The backlash fault was developed by replacing original (designed) size balls with smaller one. In this case original size was 3mm diameter, which was changed to 2.8mm, 2.6mm and 2.4mm (Figure 9). The fatigue fault was simulated in a number of phases initiating with a small surface defect of less than 1mm on ball screw channel

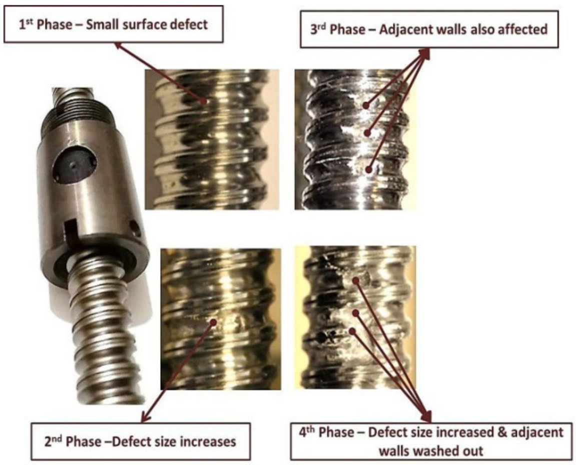


**FIGURE 10.** Fatigue fault phases.

**TABLE 1.** Dataset description for proposed framework.

| Fault Mode | Fault Category | Class | No of Samples | | Label |
|---|---|---|---|---|---|
| | | | Training | Testing | |
| No Fault | Faults in acceptable range | NF | 240 | 80 | 0 |
| Friction | Half lubricant cease | FR-1 | 240 | 80 | 1 |
| Friction | All lubricant cease | FR-2 | 240 | 80 | 2 |
| Friction | Tightened Seals | FR-3 | 240 | 80 | 3 |
| Backlash | 2.8 mm Ball Size | B-1 | 240 | 80 | 4 |
| Backlash | 2.6 mm Ball Size | B-2 | 240 | 80 | 5 |
| Backlash | 2.4 mm Ball Size | B-3 | 240 | 80 | 6 |
| Fatigue | 1st Phase | F-1 | 240 | 80 | 7 |
| Fatigue | 2nd phase | F-2 | 240 | 80 | 8 |
| Fatigue | 3rd Phase | F-3 | 240 | 80 | 9 |
| Fatigue | 4th Phase | F-4 | 240 | 80 | 10 |

in first phase (Figure 10). In the second phase, the fault was propagated by increasing size of spall up to 2mm, 3.5mm and 5mm successively on single channel. In the third phase, adjacent channels were also affected with 3.5mm and 5mm surface defects. In the last phase, the walls between adjacent channels were gradually washed away (Figure 10).

The testing was carried out for different load conditions to assess the capability of developed technique. The motion profile was tested with 80mm linear travel (1 complete cycle) which was completed in 6sec 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. The test setup was run initially for 20 minutes to reach steady behavior of the system (normalize temperature effects inside motor and ball nut) before data acquisition.

### C. DATASET DETAILS

#### 1) CURRENT SIGNAL DATA SET

The current signal dataset for different fault modes was collected using above experimentation. To achieve better performance [34] and higher feature extraction accuracy of deep learning networks, data augmentation is performed. Data augmentation increases number of data samples and improves CNN generalization [34]. A large number of training data is obtained by segmenting samples with signal overlap to extend original data as shown previously in Figure 2. The signal data with 50000 points can generate 120 samples which becomes 960 training samples with shift length of 400.

The testing setup also simulates three types of fault cases under different loading. A total of 4 cases i.e. no-fault, mechanical friction, mechanical backlash and fatigue was considered for data collection. Dataset for these four cases was generated under different fault categories. The description of experimental data set is shown in Table 1 below. The fault modes are represented by FR for friction, B for backlash, F for fatigue and NF for no fault. The severity of faults is in the range of 1 – 3 where 1 is less severe 2 more severe and 3 being most severe. Fatigue fault is considered in four phases as per details mentioned in section IV-B.

The signal data was sampled at 50176 sampling points which was processed by 1-D to 2-D signal conversion method as proposed in section III-A to generate $224 \times 224$ gray image. 2-D converted gray images for different fault cases are shown in Figure 11. The variation in gray image textures for different fault cases can be easily seen from these images.

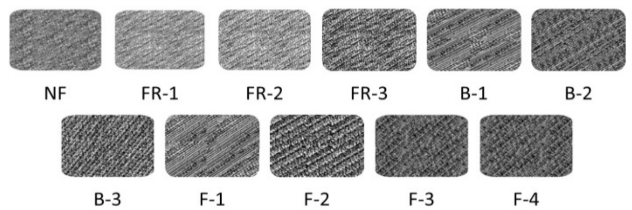


**FIGURE 11.** 2-D Gray Images for different fault cases.

#### 2) PADERBORN PUBLISHED BEARING DATASET

The developed model was also evaluated on available paderborn published datasets that comprise multi-stage bearings data. This include data from healthy as well as damaged bearings with real and artificial inner race and outer race damages with low (1), medium (2) and high (3) intensities. For comparison, 6203 type bearing is considered here. Artificial faults were generated using Electric discharge machining (EDM), electric Engraving (E) and Drilling (D) etc. Real damages include fatigue (f) and plastic (p) deformations. The dataset details for normal and artificial bearing faults are shown in Table 2 and Table 3. Here KA shows bearing number with outer ring damage, KI shows bearing number with inner ring damage, IR represents damaged inner race bearing and OR shows outer race fault. D and E shows drilling faults and engraved faults respectively. For comparison purpose, K001 and K002 data was selected as testing dataset and remaining as training data.

**TABLE 2.** Paderborn dataset description for artificial faults.

| Bearing Code | Fault Category | Fault Intensity | Fault Method |
|---|---|---|---|
| KA01 | OR | 1 | EDM |
| KA03 | OR | 2 | E |
| KA05 | OR | 1 | E |
| KA06 | OR | 2 | E |
| KA07 | OR | 1 | D |
| KA09 | OR | 2 | D |
| K101 | IR | 1 | EDM |
| K103 | IR | 1 | E |
| K105 | IR | 1 | E |
| K107 | IR | 2 | E |
| K108 | IR | 2 | E |

**TABLE 3.** Paderborn dataset description for real faults.

| Bearing Code | Fault Category | Fault Intensity | Fault Method |
|---|---|---|---|
| KA15 | OR | 1 | P |
| KA16 | OR | 2 | F |
| KA22 | OR | 1 | F |
| KA30 | OR | 1 | P |
| K114 | IR | 1 | F |
| K116 | IR | 3 | F |
| K117 | IR | 1 | F |
| K118 | IR | 2 | F |
| K121 | IR | 1 | F |

### D. IMPLEMENTATION DETAILS

The 2D-Rem-CNN was implemented in Python TensorFlow using windows operated desktop system Intel core i7 with 12 GB memory on GTX1070 GPU. The CNN training was performed with stochastic gradient descent (SGD) since good convergence was not achieved mini-batch SGD. The network was trained with 150 epochs. Other parameters include momentum of 0.8 along with 0.2 dropout rate and decaying learning rate of 0.002 reduce to half after every 100 iterations. The experimentation was carried out on 2-D transformed signal images.

The dataset was divided into training data and testing (including validation) data. The network was trained after data augmentation using large number of training samples to improve model performance and avoid data over-fitting. Initially, the weights of network model were generated randomly. Random testing samples with varying sample size were selected from testing dataset to observe its effect on model performance.

### E. PERFORMANCE FACTORS

To evaluate the performance of our system, we have considered two performance indicators: Accuracy and Precision. Accuracy is a comprehensive measure of classification performance of system. It can be defined mathematically as;

$$Acc = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{9}$$

where, $T_P$ and $F_P$ are true positive and false positive, $T_N$ and $F_N$ true negative and false negative samples. True positive includes positive samples which are classified correctly and true negative means negative samples which are classified correctly. Similar description for false positive and false negative applies.

Precision indicates the proportion of correct classified sample outcomes. Mathematically precision is defined as;

$$Prec = \frac{T_P}{T_P + F_P} \tag{10}$$

## V. RESULTS

The proposed network was trained with 150 epochs to learn no-fault and faulty features. 2D-Rem-CNN model was trained to learn automatically, features from training dataset samples for each condition. The trained model was then validated and tested. Figure 12 below shows the curves for training and testing accuracy along with precision results. The trained model converges after finite iterations. The model successfully learnt features from data with nearly 99 percent accuracy rate. The average precision of each fault was found as 96 percent. The results indicate that developed 2D-Rem-CNN model is very efficient to learn features from 2D signal images to classify ball-screw drive faults for 150 epochs accordingly.

### A. COMPARISON WITH MAINSTREAM TECHNIQUES

To validate the performance of 2D-Rem-CNN, the model was compared with some existing state of the art deep learning techniques and their performance attributes were compared. Other techniques used for comparison include Support Vector Machine (SVM) trained with raw signal data using time-domain feature parameters [21] and LeNet [40]. The achieved results are shown in Table 4. It can be noticed that the proposed residual connection network with other additional improvements outperformed different techniques. SVM and LeNet show low accuracy 81% and 88% respectively. However, the proposed technique shows higher accuracy of above 98%. The higher performance outcome of this framework shows superiority of the proposed methodology.

**TABLE 4.** Performance comparison with other techniques.

| Comparison Technique | Accuracy |
|---|---|
| Support Vector Machine | 81.9 % |
| LeNet | 88.7 % |
| 2D-Rem-CNN | 98.95 % |

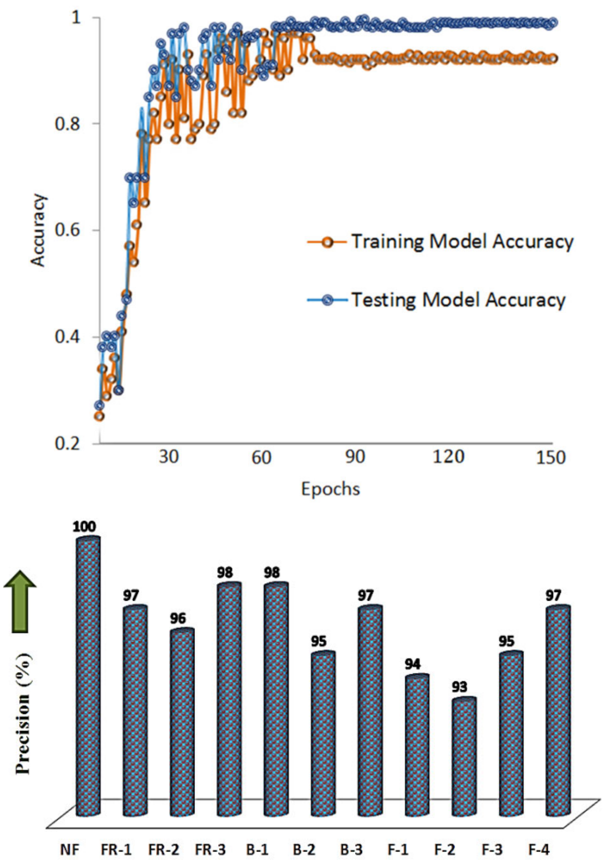


**FIGURE 12.** (a) Accuracy and (b) precision of 2D-Rem-CNN Model.

### *B. CONFUSION MATRIX*

For the testing dataset, confusion matrix was calculated using trained model. Looking at the confusion matrix in Figure 13, it can be visualized that the network perfectly identifies

| Desired Output \ Actual Output | NF | FR-1 | FR-2 | FR-3 | B-1 | B-2 | B-3 | F-1 | F-2 | F-3 | F-4 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| NF | 100.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % |
| FR-1 | 0.0 % | 97.0 % | 3.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % |
| FR-2 | 0.0 % | 3.0 % | 97.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % |
| FR-3 | 0.0 % | 0.0 % | 1.0 % | 99.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % |
| B-1 | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 96.0 % | 0.0 % | 4.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % |
| B-2 | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 4.0 % | 96.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % |
| B-3 | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 2.0 % | 98.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % |
| F-1 | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 94.0 % | 0.0 % | 6.0 % | 0.0 % |
| F-2 | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 97.0 % | 3.0 % | 0.0 % |
| F-3 | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 3.0 % | 0.0 % | 97.0 % | 0.0 % |
| F-4 | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 0.0 % | 2.0 % | 0.0 % | 98.0 % |

**FIGURE 13.** Confusion Matrix for ball-screw faults classification.

no-fault state with 100% accuracy. Furthermore, dead-band (friction) and mechanical backlash faults are classified with higher accuracy rate than with fatigue fault. In general, fatigue fault being complex in nature, is the most difficult one to classify and there is a chance that it can be confused with other fault modes. This may be due to complex fault nature or insufficient feature information for the said fault. There may also a chance that the system doesn't identify fatigue fault samples from other fault data.

### *C. MODEL VALIDATION WITH PADERBORN PUBLIC BEARING DATASET*

2D-Rem-CNN model was trained and tested with Paderborn public bearing dataset. The same was used by Wang *et al* [41] to validate their developed fault diagnostic multi-scale network (MSN). The results are compared in Figure 14. It can be seen that the training and testing data accuracies for 2D-Rem-CNN with optimized softmax are higher than Multi-scale neural network for same dataset. Moreover, it was observed that OR faults were more accurately classified than with IR damages. Additionally, training and testing data classification for real faults has shown better performance than with artificially induced faults.

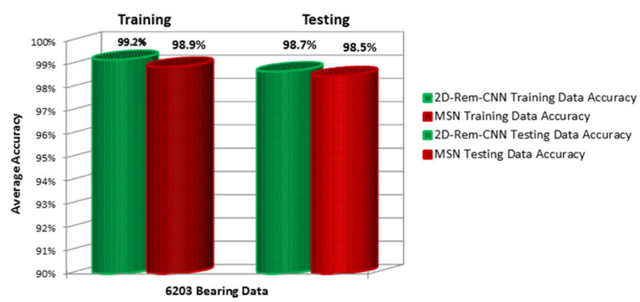


**FIGURE 14.** Comparison of 2D-Rem-CNN with MSN model.

## VI. DISCUSSIONS

The proposed model has shown higher accuracy and effective feature extraction for desired fault classification requirement. Some of the characteristic aspects of the proposed framework are discussed below.

### *A. SUPERIORITY OF PROPOSED RESIDUAL NETWORK*

Residual connections improve network accuracy and training speed by identity mapping of input data. Two residual connections, each comprising of $3 \times 3$ convolutional layer in between two $1 \times 1$ convolutional layers are added in the network. The stacked residual blocks minimize the difficulties of deep network training with improved performance. The model performance with and without residual connections was compared. The training and testing data accuracies were measured for the two models and shown in Figure 15. It can be observed from these results that the training data accuracy has increased by 6% average value with the addition of residual blocks. This indicates improved learning rate of the

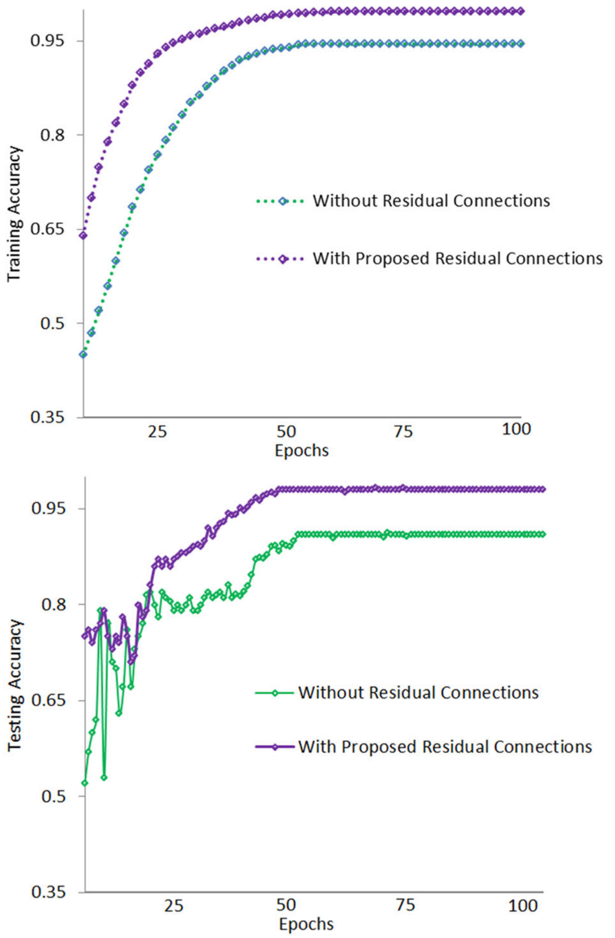


**FIGURE 15.** (a) Training and (b) Testing Accuracy Plots with and without proposed residual learning.

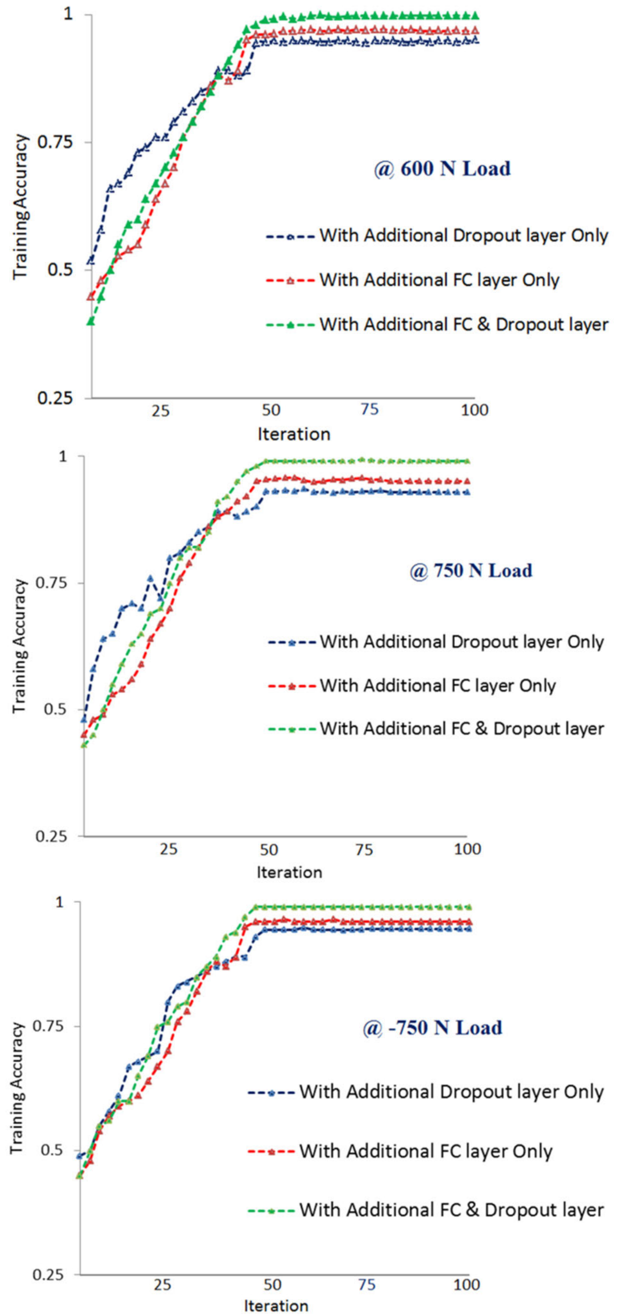


**FIGURE 16.** Accuracy under Different Loads (a) 600 N, (b) 750 N and (c) −750 N.

proposed network. More important outcome is the improvement of testing data accuracy which has increased by average of 8%. This means that network gives better performance with residual learning scheme by minimizing degradation issues. It is also noticeable that the testing data accuracy converges rapidly with residual network, which indicates that the system has improved generalization and gives effective fault classification especially with high level features. This demonstrates that residual learning has improved training difficulty problems of CNN model and gives better option than other feasible solutions.

### B. EFFECTIVENESS OF ADDITIONAL DROPOUT & FC LAYER

The effectiveness of proposed scheme was also compared using with and without additional Dropout and FC layer.

The experiments were performed in 3 phases; first with additional FC layer only, second with additional Dropout layer only and third with both additional FC and Dropout layers. Results are mentioned in Figure 16. The network training results are shown at 600N load, 750N and -750N. It can be seen that the accuracy plot converges rapidly with additional FC layer than with additional dropout layer. However, classification accuracy is low, which was improved using additional Dropout layer in combination with additional FC layer. The overall results confirm the superiority of proposed architecture in terms of classification accuracy, when compared with other models.

### C. NETWORK ENDURANCE TEST UNDER DIFFERENT LOADS

The effectiveness of system's stability to endure load variations was evaluated by comparing testing time for multiple

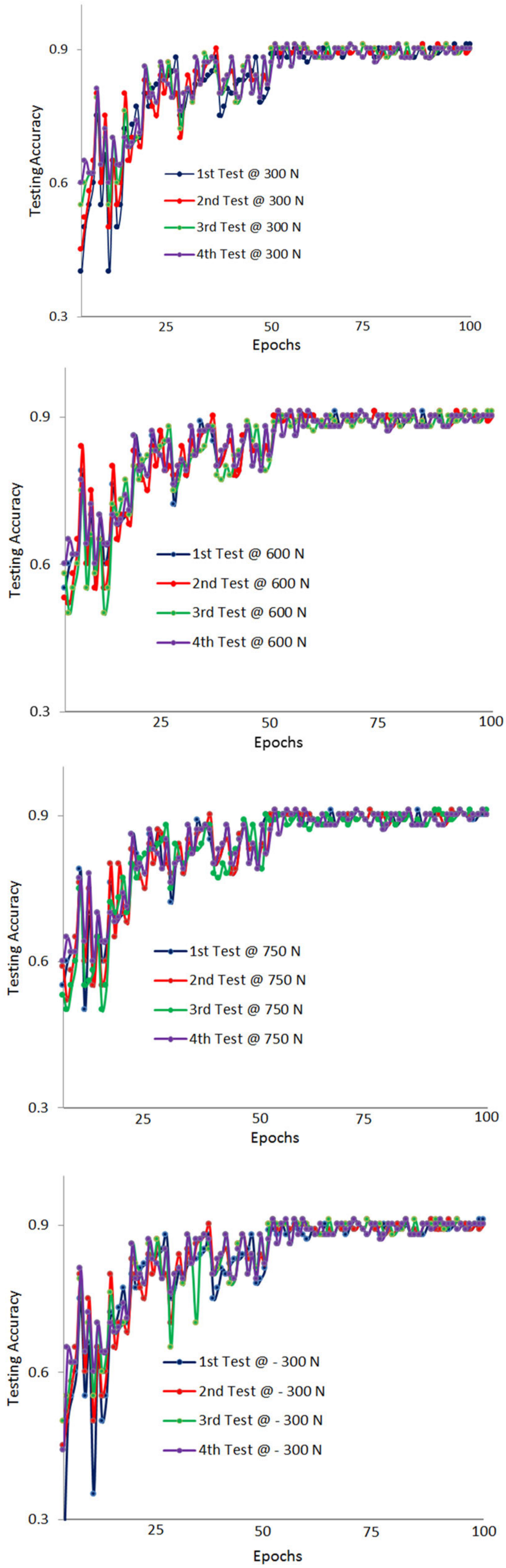


**FIGURE 17.** **Network Endurance Test Results under Load Variations (a) 300 N, (b) 600 N, (c) 750 N, (d) −300 N, (e) −600 N, (f) −750 N.**

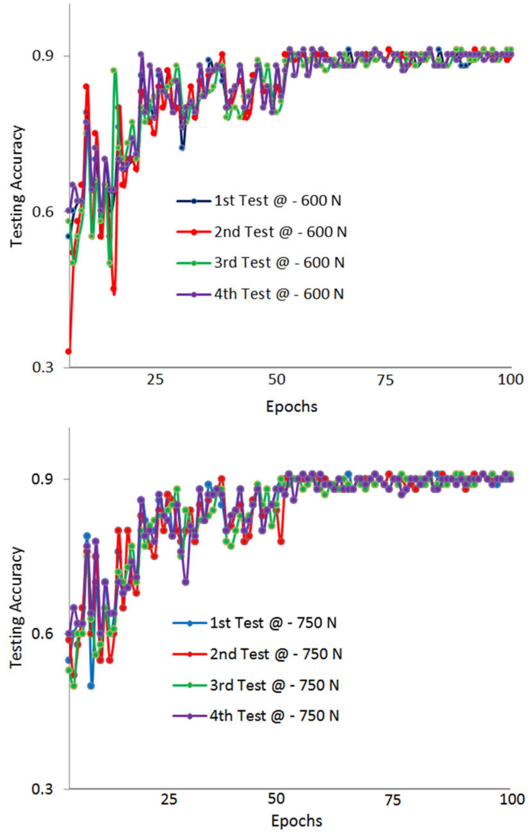


**FIGURE 17.** ***(Continued.)*** **Network Endurance Test Results under Load Variations (a) 300 N, (b) 600 N, (c) 750 N, (d) −300 N, (e) −600 N, (f) −750 N.**

experiments with changing loads on actuator's end attachment. This becomes very important for practical engineering approach. Since the actuator is linear, load was applied in terms of force (N). A series of 24 experiments were performed for six load cases (300N, 600N, 750N, −300N, −600N, −750N). The rated force capacity of actuator is 600N and peak force that it can sustain is 750N. Reversal loads were also applied for adaptive performance evaluation for different load domains. The loads were repeated consecutively, so each load was applied for four times sequentially. This gives network response to classify faults correctly and at constant time under variable load. The testing accuracy curves are shown in Figure 17. It can be seen that the model retains accuracy level in all repeated experiments with approximate variance of $\pm 0.0025$. This indicates that the proposed model endures good load stability with same testing time. The average classification time was found to be 0.9 milli-seconds for each test sample which is acceptable for practical engineering systems. This signifies the suitability of utilizing the proposed method for real-time data monitoring and faults classification.

### *D. PERFORMANCE OF OPTIMIZED SOFTMAX*

The addition of proposed marginal parameter greatly improves multi-class discrimination and promotes distinct separability between different classes. The angle between

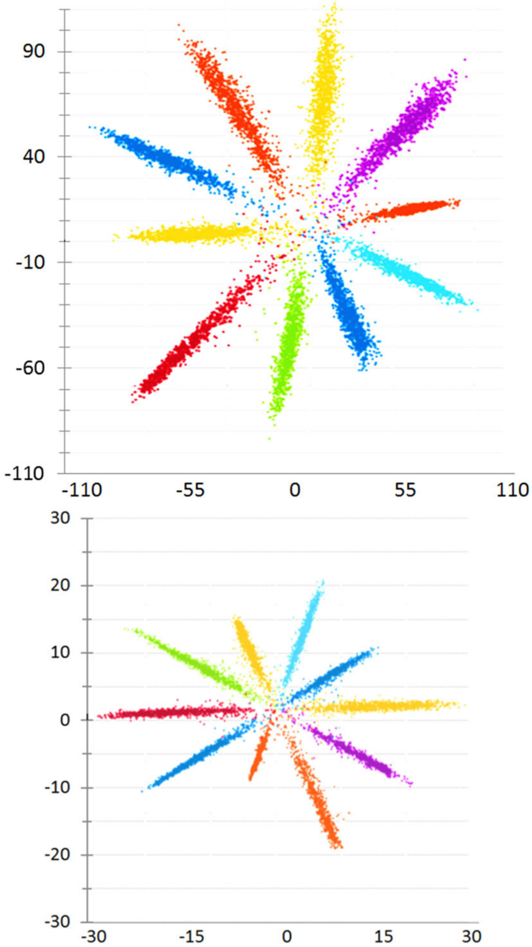


**FIGURE 18.** 2-D features visualization (a) Original Softmax loss (b) Optimized Softmax loss.

$W_{yi}$ and $x_i$ is desired to be at minimal than corresponding angle with remaining weights. Large gradients minimize variance between multiple classes which is attained by margins existence thus achieving optimization target. The performance of optimized softmax was analyzed by comparing intra-class variance for softmax loss using trained features. Figure 18 depicts low variance for optimized softmax loss observed by narrow lobes of spectrum. It can also be seen that feature magnitudes for softmax loss are high as compared to optimized softmax loss (110 magnitude as compared to 30). This is because high probability classification requires large feature norms. However, optimized softmax loss is not affected by feature norms since each feature is initially normalized before computing loss. This promotes same class features come to close locality and distinct features move away in the normalized space. The experiments have shown that our optimized softmax approach gives 2-3% higher performance with 8-10% reduction in error when compared with original softmax.

## VII. CONCLUSION

This paper provides an efficient monitoring approach for linear electro-mechanical actuators based on improved deep residual learning. The main contribution of this work includes application of current measurements to detect and classify critical fault modes for ball screw actuators. The effectiveness of proposed scheme was verified by simulating different faults in ball screw system. By utilizing residual learning and additional dropout and Fc layers, the network capability to extract typical current signal features was improved. The developed 2D-Rem-CNN model not only extracts characteristic features from dataset more effectively; but also works more systematically with optimized softmax for classification of faults including mechanical friction, backlash and fatigue failure with different severity levels under different load domains. The classification results indicate superior performance of proposed algorithm in terms of accuracy and endurance for load variations. The model was also tested with Paderborn public bearing dataset which validated its performance.

This work can be extended, in future, to explore other failure scenarios in combination to these faults for training and testing of developed network. The failure of other mechanical components in the actuator can also affect the performance of ball screws; so, a current signal pattern for combined faults will be more helpful to predict actual condition of actuator.

## REFERENCES

[1] National Transportation Safety Board (NTSB). *Aviation Accident Investigation Report—Loss of Control and Impact with Pacific Ocean, Alaska Airlines Flight 261, MD-83, N963AS*. Accessed: Dec. 30, 2002. [Online]. Available: http://www.ntsb.gov/investigations/accidentreports/pages/AAR0201.aspx

[2] 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.

[3] G. Di Rito and F. Schettini, "Health monitoring of electromechanical flight actuators via position-tracking predictive models," *Adv. Mech. Eng.*, vol. 10, no. 4, Apr. 2018, Art. no. 168781401876814, doi: 10.1177/1687814018768146.

[4] 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.

[5] 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.

[6] S. Adams, P. A. Beling, K. Farinholt, N. Brown, S. Polter, and Q. Dong, "Condition based monitoring for a hydraulic actuator," in *Proc. Annu. Conf. Prognostics Health Manage. Soc.*, Oct. 2016, pp. 1–10.

[7] Z. Yao, Y. Yu, and J. Yao, "Artificial neural network–based internal leakage fault detection for hydraulic actuators: An experimental investigation," *Proc. Inst. Mech. Eng., I, J. Syst. Control Eng.*, vol. 232, no. 4, pp. 369–382, Apr. 2018, doi: 10.1177/0959651816678502.

[8] C. S. Byington, M. Watson, D. Edwards, and P. Stoelting, "A model-based approach to prognostics and health management for flight control actuators," in *Proc. IEEE Aerosp. Conf.*, vol. 6, Mar. 2004, pp. 3551–3562.

[9] 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.

[10] G. Swerdon, M. J. Watson, S. Bharadwaj, C. S. Byington, M. Smith, K. Goebel, and E. Balaban, "A systems engineering approach to electro-mechanical actuator diagnostic and prognostic development," in *Proc. Failure Prevention, Implement., Success Stories Lessons Learned-Conf. Soc. Machinery Failure Prevention Technol.*, 2009, pp. 1–15.

[11] A. J. Chirico and J. R. Kolodziej, "A data-driven methodology for fault detection in electromechanical actuators," *J. Dyn. Syst., Meas., Control*, vol. 136, no. 4, p. 136, Jul. 2014.

[12] A. J. Chirico, J. R. Kolodziej, and L. Hall, "A data driven frequency based feature extraction and classification method for EMA fault detection and isolation," in *Proc. Legged Locomotion, Mech. Syst., Mechatronics, Mechatronics Aquatic Environ., MEMS Control, Model Predictive Control, Modeling Model-Based Control Adv. IC Engines*, vol. 2, Oct. 2012, pp. 751–760.
[13] A. Chirico and J. R. Kolodziej, "Fault detection and isolation for electro-mechanical actuators using a data-driven Bayesian classification," *SAE Int. J. Aerosp.*, vol. 5, pp. 494–502, Oct. 2012.
[14] E. Balaban, A. Saxena. K. Goebel, C. S. Byington, M. Watson, S. Bharadwaj, M. Smith, and S. Amin, "Experimental data collection and modeling for nominal and fault conditions on electro-mechanical actuators," in *Proc. Annu. Conf. Prognostics Health Manage. Soc.*, 2009, pp. 1–15.
[15] H. Chen, Y. Sun, Z. Shi, and J. Lin, "Intelligent analysis method of gear faults based on FRWT and SVM," *Shock Vib.*, vol. 2016, pp. 1–9, Sep. 2016.
[16] 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.
[17] 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.
[18] 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.
[19] N. Hatami, Y. Gavet, and J. Debayle, "Classification of time-series images using deep convolutional neural networks," 2017, *arXiv:1710.00886*. [Online]. Available: https://arxiv.org/abs/1710.00886
[20] M. Y. Asr, M. M. Ettefagh, 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.
[21] X. Wang, Y. Zheng, Z. Zhao, and J. 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] Y. 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.
[24] 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.
[25] 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.
[26] 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/Jun. 2017.
[27] N. Hatami, Y. Gavet, and J. Debayle, "Classification of time-series images using deep convolutional neural networks," in *Proc. 10th Int. Conf. Mach. Vis.*, 2017, Art. no. 106960Y.
[28] 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.
[29] 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.
[30] 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.
[31] 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.
[32] 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.
[33] S. Ma and F. Chu, "Ensemble deep learning-based fault diagnosis of rotor bearing systems," *Comput. Ind.*, vol. 105, pp. 143–152, Feb. 2019, doi: 10.1016/j.compind.2018.12.012.
[34] 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.
[35] Y. Li, J. X. Gu, D. Zhen, M. Xu, and A. Ball, "An evaluation of gearbox condition monitoring using infrared thermal images applied with convolutional neural networks," *Sensors*, vol. 19, no. 9, p. 2205, May 2019, doi: 10.3390/s19092205.
[36] W.-L. Chu, C.-J. Lin, and K.-C. Kao, "Fault diagnosis of a rotor and ball-bearing system using DWT integrated with SVM, GRNN, and visual dot patterns," *Sensors*, vol. 19, no. 21, p. 4806, Nov. 2019, doi: 10.3390/s19214806.
[37] W. Liu, Y. Wen, Z. Yu, and M. Yang, "Large-margin softmax loss for convolutional neural networks," in *Proc. ICML*, 2016, pp. 507–516.
[38] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," 2018, *arXiv:1801.07698*. [Online]. Available: http://arxiv.org/abs/1801.07698
[39] E. Grave, A. Joulin, M. Cissé, D. Grangier, and H. Jégou, "Efficient softmax approximation for GPUs," 2016, *arXiv:1609.04309*. [Online]. Available: http://arxiv.org/abs/1609.04309
[40] W. Long, 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.
[41] D. Wang, Q. Guo, Y. Song, S. Gao, and Y. Li, "Application of multiscale learning neural network based on CNN in bearing fault diagnosis," *J. Signal Process. Syst.*, vol. 91, no. 10, pp. 1205–1217, Oct. 2019, doi: 10.1007/s11265-019-01461-w.

**NAVEED RIAZ** was born in Pakistan. He received the B.S. and master's degrees in engineering from NUST Islamabad. He has more ten years of vast experience in research industry and academics. His research interests include robotics, deep learning, medical image analysis and computer vision, artificial intelligence, and robot controls and aerial robotics. He has a lot of conference and journal research publications in these research areas.

**SYED IRTIZA ALI SHAH** was born in Pakistan. He received the B.E. degree from NED University, the M.S. degree in aerospace engineering (fluid dynamics) from NUST, 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 Georgia Tech.

**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 has also research experience in various research organizations as a junior and senior researcher. His research interests include deep learning, medical image analysis, and artificial intelligence. He has a lot of conference and journal research publications in these research fields.

**SYED OMER GILANI** received the Ph.D. degree in electrical and computer engineering from the National University of Singapore, in 2013. From 2006 to 2008, he worked on various vision and augmented-reality-based projects at Interactive Multimedia Lab, Singapore. His current research interests include human–computer interaction, biological vision, and machine learning.

**EMAD UDIN** received the Ph.D. degree from the Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology (KAIST), South Korea, in 2014. His research was focused on the computational modeling of the fluid flexible body interaction (FSI) for the real life flows, including anguilliform locomotion, swimming jellyfish, and the cell flowing through the blood.