Computers, Materials & Continua
DOI: 10.32604/cmc.2023.033448
Article

Improved Symbiotic Organism Search with Deep Learning for Industrial Fault Diagnosis

Mrim M. Alnfiai*

Department of Information Technology, College of Computers and Information Technology, Taif University, Taif P.O. Box 11099, Taif, 21944, Saudi Arabia
*Corresponding Author: Mrim M. Alnfiai. Email: m.alnofiee@tu.edu.sa
Received: 16 June 2022; Accepted: 04 August 2022

Abstract: Developments in data storage and sensor technologies have allowed the cumulation of a large volume of data from industrial systems. Both structural and non-structural data of industrial systems are collected, which covers data formats of time-series, text, images, sound, etc. Several researchers discussed above were mostly qualitative, and ceratin techniques need expert guidance to conclude on the condition of gearboxes. But, in this study, an improved symbiotic organism search with deep learning enabled fault diagnosis (ISOSDL-FD) model for gearbox fault detection in industrial systems. The proposed ISOSDL-FD technique majorly concentrates on the identification and classification of faults in the gearbox data. In addition, a Fast kurtogram based time-frequency analysis can be used for revealing the energy present in the machinery signals in the time-frequency representation. Moreover, the deep bidirectional recurrent neural network (DBiRNN) is applied for fault detection and classification. At last, the ISOS approach was derived for optimal hyperparameter tuning of the DL method so that the classification performance will be improvised. To illustrate the improvised performance of the ISOSDL-FD algorithm, a comprehensive experimental analysis can be performed. The experimental results stated the betterment of the ISOSDL-FD algorithm over current techniques.

Keywords: Industrial systems; data science; fault diagnosis; deep learning; time frequency analysis

1 Introduction

In industrial production, any faults of rotating machines might lead to massive economic losses. Hence, fault diagnosis (FD) and fault detection become vital to the operative safety of latest production system [1]. Owing to the lack of fault samples, complications of fault kinds, problems of FD, and higher cost of manual annotation samples, the under small samples are still a research domain which requires continuous update, continual progress, and optimization [2,3]. With the continual growth of modern manufacturing sector, rotating machinery plays a crucial role in complicated electromechanical equipment, can be advanced toward complexity, intelligence, and large scale [4,5].

Rotating machinery generally operates for longer duration under high speeds and heavy loads. Bearing, gearboxes, and other significant elements were probably to suffer from fracture, wear, deformation, and other faults at the time under time-varying operating conditions and inherent degradation process [6]. The failure of such elements would affect the normal function and equipment usage [7].

Serious cases could lead to damage or downtime and they might occur huge economic losses and casualties [8,9]. Thus, research relating to condition monitoring and FD of rotating machinery was one such key which assures the normal function of equipment, minimizes unplanned maintenance, ignored catastrophic failures, and assures the security of industrial production that was most important one [10]. Main tasks of FD research on rotating machinery involve determination of functioning state, determining the fault kind, and estimating the fault trend. Recently, authors have made many research studies in the domain of rotating machinery FD that reached productive research outcomes and were implied in actual working states [11,12]. With the speedy advancement of intellectual methods, the recent study on deep learning (DL) in the FD of rotating machinery was rising year by year, and it grabbed the interest of many authors. For the promotion of the research growth in this domain, the paper assesses the related research from the two prospects of DL concept and its applications in FD [13]. (Research could grant inspiration and convenience for related authors and offer references for promoting and understanding the advancement of FD research.

This article presents an improved symbiotic organism search with deep learning enabled fault diagnosis (ISOSDL-FD) model for gearbox fault detection in industrial systems. The proposed ISOSDL-FD technique applies a Fast kurtogram based time-frequency analysis for revealing the energy present in the machinery signals in the time and frequency representation. In addition, the deep bidirectional recurrent neural network (DBiRNN) is applied for fault detection and classification. At last, the ISOS method can be derived for optimal hyperparameter tuning of the DL approach. To illustrate the improvised performance of the ISOSDL-FD algorithm, a comprehensive experimental analysis was performed.

The rest of the paper is organized as follows. Section 2 offers the literature review, Section 3 elaborates the proposed ISOSDL-FD model, Section 4 provides performance validation, and Section 5 concludes the paper.

2 Literature Review

Wang et al. [14] suggest a novel fault prognosis technique has infrared thermal images and least-square-interactive-support-matrix-machine (LSISMM). In this technique, a new matrix-form classifier known as LSISMM can be created over the idea of non-parallel interactive hyperplane to completely use the structure data of infrared thermal imageries. To enhance the computational efficacy, innovative least square loss constraints were projected for LSISMM. Additionally, the researchers extract an effectual solution structure on the basis of the alternative direction technique of the multiplier (ADMM) structure. In [15], an adaptive convolutional neural network (CNN) was built on the basis of the ontology representation of pump. Firstly, time–frequency features of pressure signals were acquired through continual wavelet transformation. Secondly, establishment of a deep CNN by setting initial hyperparameter. Thirdly, Bayesian optimization can be leveraged for achieving automatic learning of main significant hyperparameters for constructing an adaptive CNN. In [16], a new FD technique for rotating machinery on the basis of deep residual neural network (DRNN) and data fusion was suggested. Initially, the frequency domain and time domain features of the original signal were derived via the Short-time Fourier transform (STFT) layer, after which the DRNN and

the fusion embedding layers were utilized for fusing the frequency domain, spatial domain, and time domain features for obtaining high quality low-dimension fusion features.

An FD technique for rotating machinery in time varying circumstances on the basis of DL and tacholess order tracking (TOT) was suggested in [14]. At first, estimated speed information and frequency domain periodic signals were attained by tracking the order. Next, the frequency domain periodic signals refer to a speed which can be standardized with the help of estimated speed data. At last, standardized features were derived by DL method for making feature vectors. The feature vector can be fed to softmax layers to finish FD of the gearboxes. Li et al. [17] advance an FD technique with the help of CNN for InfRared Thermal (IRT) images. Firstly, IRT method can be used for capturing the IRT images of rotating machinery. Secondly, the CNN can be implied for extraction of fault features from the IRT images. Finally, the acquired features were fed to the Softmax Regression (SR) classification for fault pattern identification. In [18], a novel DL related FD technique, that derives features from frequency as well as time domains were suggested. A set of 2 deep features from several domain was merged into global principle component analysis and intrinsic low-dimensional features by local. And a novel ensemble kernel extreme learning machine was suggested for fault pattern classifier on the basis of the fused feature.

3 The Proposed Model

In this study, a new ISOSDL-FD technique was introduced for the recognition of faults in industrial systems. The proposed ISOSDL-FD technique accomplished proficient identification and classification of faults in the gearbox data. In addition, a Fast kurtogram based time-frequency analysis is used for revealing the energy present in the machinery signals in the time and frequency representation. Moreover, the ISOS-DBiRNN is applied for fault detection and classification. Fig. 1 depicts the overall process of ISOSDL-FD approach.

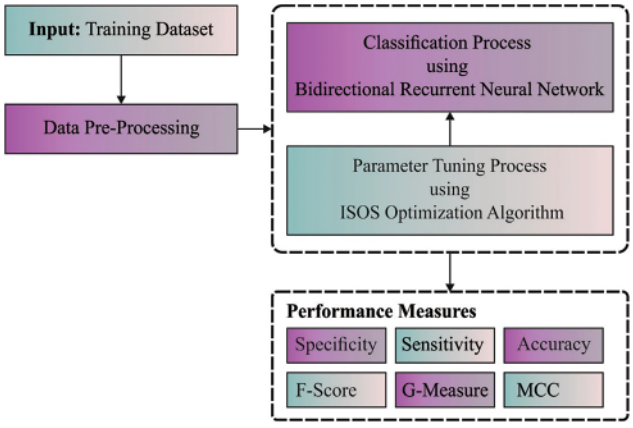

Figure 1: Overall process of ISOSDL-FD approach

3.1 Time Frequency Analysis

Time-frequency analysis was helpful for revealing the energy restricted from the machinery signal from the time and frequency representation [19]. Different kinds of time-frequency transformations like spectral kurtosis diagram (kurtogram), wavelet transform (WT), empirical mode decomposition (EMD), and STFT are utilized to machinery FD. But, every approach are flaws. It can be complex for

selecting the mother wavelet function to WT, but EMD models offer in mode mixing problems. Besides, in STFT approaches, the time and frequency resolutions were connected together. The fast kurtogram system offers any benefits on other methodologies as the present generation of fast kurtogram doesn't need some parameters that set, creating the model simple for utilizing. So, the fast kurtogram has been selected to the presented analysis model for reducing the amount of parameters which required that set. The kurtogram is 4th order of spectral diagnosis that is called spectral kurtosis. Antoni imposed 4 constraints on kurtogram diagnosis for increasing the versatility of procedure that creates the method further sensitive to non-stationary signals. The kurtogram is determined utilizing in Eq. (1):

$$k_x f\,(x) = \frac{\langle H^4\,(t,f)\rangle}{\langle H^2\,(t,f)\rangle^2} - 2, \tag{1}$$

whereas $\langle . \rangle$ implies the operator of time averaging and $H\,(t, f)$ denotes the enveloped procedure of signals $x\,(t)$. $H\,(t, f)$has been evaluated utilizing in Eq. (2):

$$H\,(t,\ f) = \Sigma_{n=t}^{t+N_n-1} w\,(n-t)\,x\,(n)\,e^{-j2\pi fn} \tag{2}$$

In which $w\,(t)$ represents the length of windows.

3.2 Fault Detection and Classification Model

In this study, the DBiRNN is applied for fault detection and classification. The prediction network is carried out by means of a Recurrent Neural Network (RNN) [20]. The neural network takes the temporal series of observation $X_t = (x_{t-T-1}, x_{t-T-2}, \ldots, x_t)$ as input, and generate a series of vectors $(h_{t-T-1}x_{t-T-2}, \ldots, h_t)$ through a nonlinear activation. The output vector is hidden variable that describes the distribution of observation. Unlike the conventional feedforward neural network, an RNN considers previous data or named "memory". The label is defined using a nonlinear mapping over this hidden variable and it is mathematically expressed in the subsequent formula.

$$h_t = H\,(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \tag{3}$$

$$\mathcal{Y}_t = F\left(W_{hy}h_t + b_y\right) \tag{4}$$

Now, H refers to a nonlinear function that is selected from sigmoid and *tanh*. $F$is generally selected from *softmax* or sigmoid when the output label is independent or mutually exclusive correspondingly. $W$and b are related weight (for example W_{xh} represent the input-hidden weight matrixes), and bias (for example, b_h is hidden bias vector) matrix correspondingly. The parameter is learned from the trained dataset. In recent times, researcher demonstrates a problem using simple RNN structure where the gradient disappears. In order to resolve the issue, the most common replacement is Gated Recurrent Unit (GRU) cells and Long Short-Term Memory (LSTM) that maintains their state over time, which leads to captured long term dependency within time sequence data. In this study, the version of GRU cell and LSTM utilized are correspondingly.

$$h_! = H_{LSTM}\,(x_t h_{t-1},\ c_{t-1}) \tag{5}$$

$$h_t = H_{GRU}\,(x_t h_{t-1}) \tag{6}$$

The main distinction among GRU and LSTM is that the content of GRU cell memory is exposed continuously to output and it is easily implemented because it needs fewer parameters in the network. Briefly, the RNN cell $H_X\,(.)$ can be LSTM (Eq. (5)) or GRU cell (Eq. (6)) RNN. One drawback of RNN is that they can utilize preceding context. But with a windowed approach where entire temporal

context is available, there is no reason not to utilize future context. The Bi-directional RNN (BiRNN) with elementary RNN unit achieves this aim by data processing in these two directions with two discrete hidden layers that are later fed to the output layer in similar fashion. A BiRNN calculates the output sequence y_t, $\vec{h}_t$ forward hidden sequence, and $\overleftarrow{h}_t$ the backward hidden sequence the by reiterating the forward layer from t to $t-T+1$, the backward layer from $t-T+1$ to t, and later update the output layer:

$$\vec{h} = H\left(W_{x\vec{h}}x_t + W_{\vec{h}\vec{h}}\vec{h}_{t-1} + b_{\vec{h}}\right) \tag{7}$$

$$\overleftarrow{h}_t = H\left(W_{x\overleftarrow{h}}x_t + W_{\overleftarrow{h}\overleftarrow{h}}\overleftarrow{h}_{t-1} + b_{\overleftarrow{h}}\right) \tag{8}$$

$$y_t = F\left(W_{\vec{h}y}\vec{h} + W_{\overleftarrow{h}y}\overleftarrow{h}_t + b_y\right) \tag{9}$$

Assisting LSTM with this bi-directional learning ability allows the system to access long-range context in input direction and enrich sequence data understanding. With the potential to disentangle complicated temporal dependency, deep RNN is constructed by stacking various RNN hidden layers, with the output series of one layer form the input series for the following. Assume the same hidden layer as for each N layer in the stack, the hidden vector sequence h^n is computed iteratively from $n = 1$ to N and $t - T + 1$ to t:

$$h_! = H\left(W_{h^{n-1}h^n}h^{n-1}_{+} + W_{h^nh^n}h^n_{t-1} + b_{h^n}\right) \tag{10}$$

$$y_t = F\left(W_{h^Ny}h^N_t + b_y\right) \tag{11}$$

where $h^0 = x$.

DBiRNN is implemented by replacing every h^m hidden sequence with $\vec{h}^m$ and $\overleftarrow{h}^m$ the backward and forward sequences, and ensures that each hidden layer receives input from the backward and forward layers. Deep architecture is attained by replacing every RNN unit in the DBiRNN by a deep uni-directional RNN. Fig. 2 showcases the infrastructure of Bi-directional RNN. Therefore, for M stacked BiRNN and every backward and forward cell have N stacked RNN structure, we attain a $2 \times M \times N$ RNN unit system. But we found that deep configuration doesn't produce improved performance on limited data as they overfit the training dataset. Hence, we use single Bi-LSTM stacked with single uni-directional GRU cell yields a three RNN unit system, as

$$\left(\vec{h}^1_t, \vec{c}^1_t\right) = H_{LSTM}\left(x_t, \vec{h}^1_{t-1}, \vec{c}^1_{t-1}\right) \tag{12}$$

$$\left(\overleftarrow{h}^1_t, \overleftarrow{c}^1_t\right) = H_{LSTM}\left(x_t, \overleftarrow{h}^1_{t-1}, \overleftarrow{c}^1_{t-1}\right) \tag{13}$$

$$h^2_t = H_{GRU}\left(\left[\vec{h}^1_t : \overleftarrow{h}^1_t\right], h^2_{t-1}\right) \tag{14}$$

$$\mathcal{Y}_t = F\left(W_{h^2y}h^2_t + b_y\right) \tag{15}$$

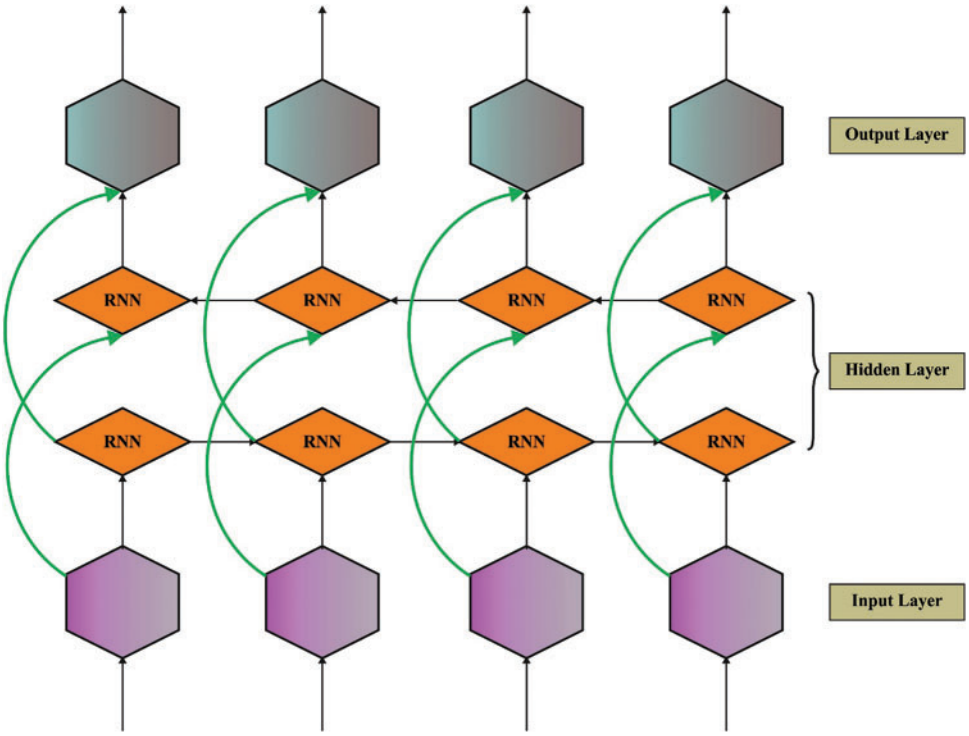

Figure 2: Structure of Bi-directional RNN

3.3 Hyperparameter Tuning Model

At the final stage, the ISOS algorithm is derived for optimal hyperparameter tuning of the DL model in such a way that the classification performance gets improved. The SOS algorithm was first introduced by Cheng et al. [21]. The SOS technique is the new methodology to solve optimization problems depends on organism interaction in nature. Organism rarely lives in isolation because of the dependence on other species for survival and nutrition. This relation depends on trust that is called a symbiotic relationship. The process initiates by a population initialization named the ecosystem. During the primary ecosystem, a group of organisms is randomly generated. Every organism represents a proposed solution for the problem. The time complexity in the SOS is characterized as n^2. The SOS renders a new solution according to the biological relations among the two living organisms. It consists of Mutualism, Commensalism, and Parasitism of biological relationships.

3.3.1 Mutualism

Currently, both organism benefits from the relationships they have with one another. Assume honeybee relationships with flowers, for instance, the bees fly over the flower for collecting the nectar needed to produce honey. Also, it is useful to the flower since bees, scattering the pollen, facilitate their pollination. In this study, Xi refers to an organism equivalent to $i-th$ individuals in nature. Then, Xj is selected at a random fashion and related to the Xi in nature. At last, Xi and Xj upgraded based on the following equation:

$$X_{inew} = X_i + rand\ (0,1) \times (X_{best} - Mutua1\_Vector\ \times BF_1) \tag{16}$$

$$X_{jnew} = X_j + rand\ (0,1) \times (X_{best} - Mutua1\_Vector\ \times BF_2) \tag{17}$$

$$Mutual_{Vector} = \frac{X_i + X_j}{2} \tag{18}$$

Now, rand (0, 1) signifies a randomly generated number. BF1 and BF2 denote the profit factors of *Xi* and *Xj* that showing every organism profit. In Eq. (16), Mutual_Vector signifies the relationships among the Xi and Xj. The $Mutual\_Vector * BF2$ in Eqs. (16) and (17) attempts to rise the survival of living organisms. Assumed the Darwinian Concept, that is survival of fittest, each organism has to improve the compatibility level with the surrounding. Now, the Xbest indicates the maximum level of compatibility [18].

3.3.2 Commensalism

Here, the organism benefit, and the other gains nothing from the relationships. For instance, assume the relationships among sharks and sticky fish. The shark gets little or no benefit and the sticky fish stick to the shark and feed on the residual food. As in the Mutualism stage, the organism *Xj* can be selected in a random fashion and is related to Xi. In such cases, Xi strives to gain profit, however, *Xj* get loss or no benefit.

$$X_{inew} = X_i + rand\,(-1,1) \times \left(X_{best} - X_j\right) \tag{19}$$

In Eq. (19), $X_{best} - X_j$ refers to the benefit given by X_j for increasing the survival of X_i.

3.3.3 Parasitism

In this phase, one organism gets benefits from the relationships, and others endure a loss. Once the parasite proliferates in human body causes that individual to die. In this study, the Xi, that is, malaria mosquito, creating an artificial parasite named "Parasite-Vector." The Parasite-Vector is generated in the searching space by imitating Xi. *Xj* is chosen at a random fashion by the ecosystem and serves the parasite as host. The Parasite-Vector attempts in taking *Xj* position in nature. The Xi and *Xj* are estimated for measuring competency. Once the Parasite-Vector destroyed *Xj* and take their position, it obtains its maximum competency. The maximum competency for *Xj* is accomplished once they resist the parasite, then the parasite no longer lives in the ecosystem.

During iterative optimization of complex problems, the SOS algorithm falls into local optima. In such cases, the solution value remains the same in the iteration. To address the shortcoming and increased the probability of algorithm jump out of the local optima location, the ISOS algorithm has been derived by adding a Gaussian mutation or a random perturbation (the function ran two hundred times and the outcome remains unchanged), and continue to run the process as follows:

$$x^{i,iter+1} = \begin{cases} x^{i,iter} + rand & if\ r \geq 0.2 \\ x^{i,iter} \times Gaussian\,(\mu,\ \sigma) & otherwise \end{cases} \tag{20}$$

In Eq. (20), rand refers to a random integer within [0, 1], r_j is the choosing probability of to implement random perturbation or Gaussian mutation. The Gaussian variation distribution function is given by:

$$Gaussian\ (\mu,\ \sigma) = \left(1/\sqrt{2\pi\sigma}\right)\ \exp\ \left(-\left(x - \frac{(\mu)^2}{2\sigma^2}\right)\right) \tag{21}$$

In Eq. (21), μ refers to the mean value, σ^2 denotes the variance

The ISOS approach extracts a fitness function for attaining enhanced classifier outcomes. It fixes a positive numeral for indicating superior outcome of the candidate solutions. In this paper, the reduction of a classifier error rate was regarded as a fitness function, as presented in Eq. (22).

$$\begin{aligned} fitness\,(x_i) &= ClassifierErrorRate\,(x_i) \\ &= \frac{number\ of\ misclassified\ samples}{Total\ number\ of\ samples} * 100 \end{aligned} \tag{22}$$

4 Results and Discussion

The experimental data employed in this paper is accumulated on a planetary gearbox. The planetary gearbox data samples were categorized into five health states: normal state (Normal), K1 planetary gear failure (fault 1), K1 planetary gear failure (fault 2), K2 planetary gear failure (fault 3), and K3 planetary sun gear failure (fault 4). A set of 3000 samples were utilized in this study as represented in Tab. 1.

Table 1: Dataset details

Class	No. of samples
Normal	600
Fault1	600
Fault2	600
Fault3	600
Fault4	600
Total number of samples	3000

Fig. 3 highlights the confusion matrices produced by the ISOSDL-FD model on 70% of training (TR) data and 30% of testing (TS) data. With 70% of TR data, the ISOSDL-FD model has recognized 394 samples into normal, 393 samples into fault 1, 406 samples into fault 2, 400 samples into fault 3, and 405 samples into fault 4. Also, with 30% of TS data, the ISOSDL-FD method has identified 172 samples into normal, 177 samples into fault 1, 163 samples into fault 2, 171 samples into fault 3, and 172 samples into fault 4.

Tab. 2 illustrates a detailed classification result of the ISOSDL-FD model under 70% of TR and 30% of TS data. Fig. 4 shows the result analysis of ISOSDL-FD approach under 70% of TR data. The ISOSDL-FD model has recognized normal class with $accu_y$ of 98.19%, $sens_y$ of 94.48%, $spec_y$ of 99.11%, F_{score} of 95.40%, MCC of 94.28%, and $G_{measure}$ of 95.40%. Moreover, the ISOSDL-FD technique has detected Fault-1 class with $accu_y$ of 98.29%, $sens_y$ of 96.09%, $spec_y$ of 98.82%, F_{score} of 95.62%, MCC of 94.56%, and $G_{measure}$ of 95.62%. Furthermore, the ISOSDL-FD technique has recognized Fault-4 class with $accu_y$ of 97.67%, $sens_y$ of 95.52%, $spec_y$ of 98.21%, F_{score} of 94.30%, MCC of 92.84%, and $G_{measure}$ of 94.30%.

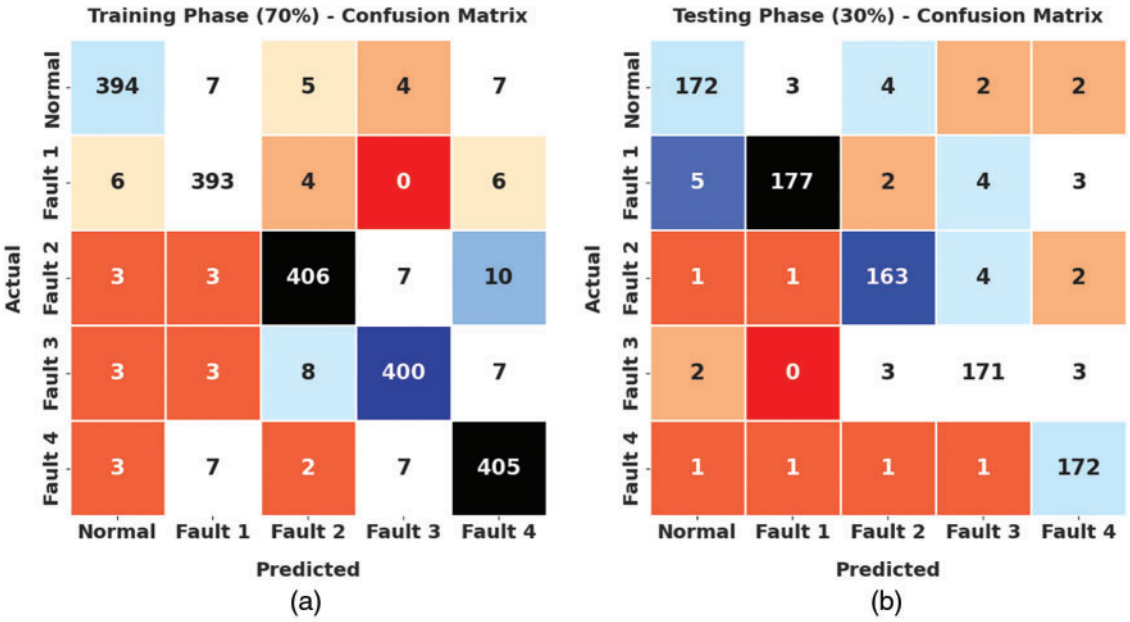

Figure 3: Confusion matrices of ISOSDL-FD approach (a) 70% of TR data and (b) 30% of TS data

Table 2: Result analysis of ISOSDL-FD approach with various measures under 70% of TR and 30% of TS data

Training/Testing (70:30)						
Labels	Accuracy	Sensitivity	Specificity	F-Score	MCC	G-Measure
Training phase						
Normal	98.19	94.48	99.11	95.40	94.28	95.40
Fault 1	98.29	96.09	98.82	95.62	94.56	95.62
Fault 2	98.00	94.64	98.86	95.08	93.83	95.08
Fault 3	98.14	95.01	98.93	95.35	94.19	95.35
Fault 4	97.67	95.52	98.21	94.30	92.84	94.30
Average	98.06	95.15	98.79	95.15	93.94	95.15
Testing phase						
Normal	97.78	93.99	98.74	94.51	93.11	94.51
Fault 1	97.89	92.67	99.29	94.91	93.62	94.93
Fault 2	98.00	95.32	98.63	94.77	93.53	94.77

(Continued)

Table 2: Continued

Training/Testing (70:30)						
Labels	Accuracy	Sensitivity	Specificity	F-Score	MCC	G-Measure
Fault 3	97.89	95.53	98.47	94.74	93.42	94.74
Fault 4	98.44	97.73	98.62	96.09	95.14	96.10
Average	98.00	95.05	98.75	95.00	93.77	95.01

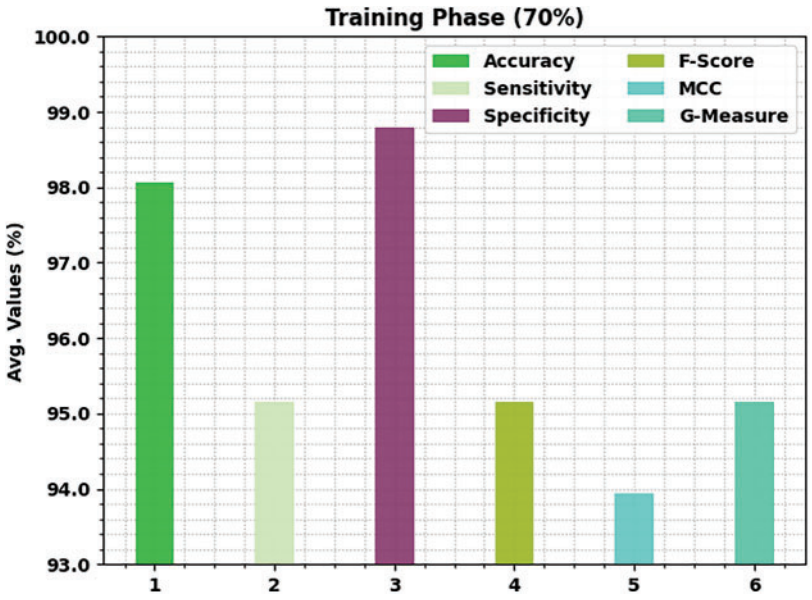

Figure 4: Average analysis of ISOSDL-FD algorithm under 70% of TR data

Fig. 5 displays the result analysis of ISOSDL-FD method under 30% of TS data. The ISOSDL-FD approach has identified normal class with $accu_y$ of 97.78%, $sens_y$ of 93.99%, $spec_y$ of 98.74%, F_{score} of 94.51%, MCC of 93.11%, and $G_{measure}$ of 94.51%. in addition to that, the ISOSDL-FD methodology has recognized Fault-1 class with $accu_y$ of 97.89%, $sens_y$ of 92.67%, $spec_y$ of 99.29%, F_{score} of 94.91%, MCC of 93.62%, and $G_{measure}$ of 94.93%. Also, the ISOSDL-FD technique has recognized Fault-4 class with $accu_y$ of 98.44%, $sens_y$ of 97.73%, $spec_y$ of 98.62%, F_{score} of 96.09%, MCC of 95.14%, and $G_{measure}$ of 96.10%.

Fig. 6 exhibits the confusion matrices generated by the ISOSDL-FD approach on 80% of TR data and 20% of TS data. With 80% of TR data, the ISOSDL-FD approach has recognized 469 samples into normal, 463 samples into fault 1, 468 samples into fault 2, 470 samples into fault 3, and 466 samples into fault 4. Moreover, with 30% of TS data, the ISOSDL-FD algorithm has identified 115 samples into normal, 118 samples into fault 1, 110 samples into fault 2, 119 samples into fault 3, and 123 samples into fault 4.

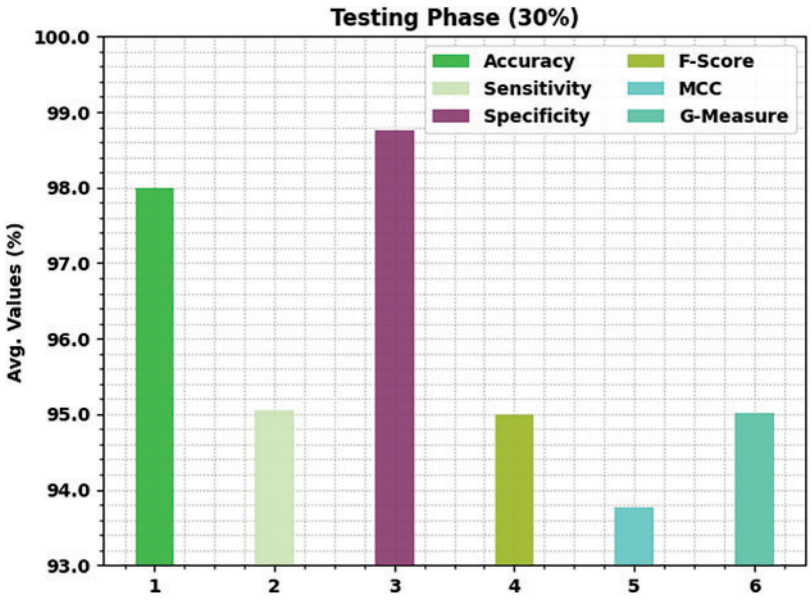

Figure 5: Average analysis of ISOSDL-FD algorithm under 30% of TS data

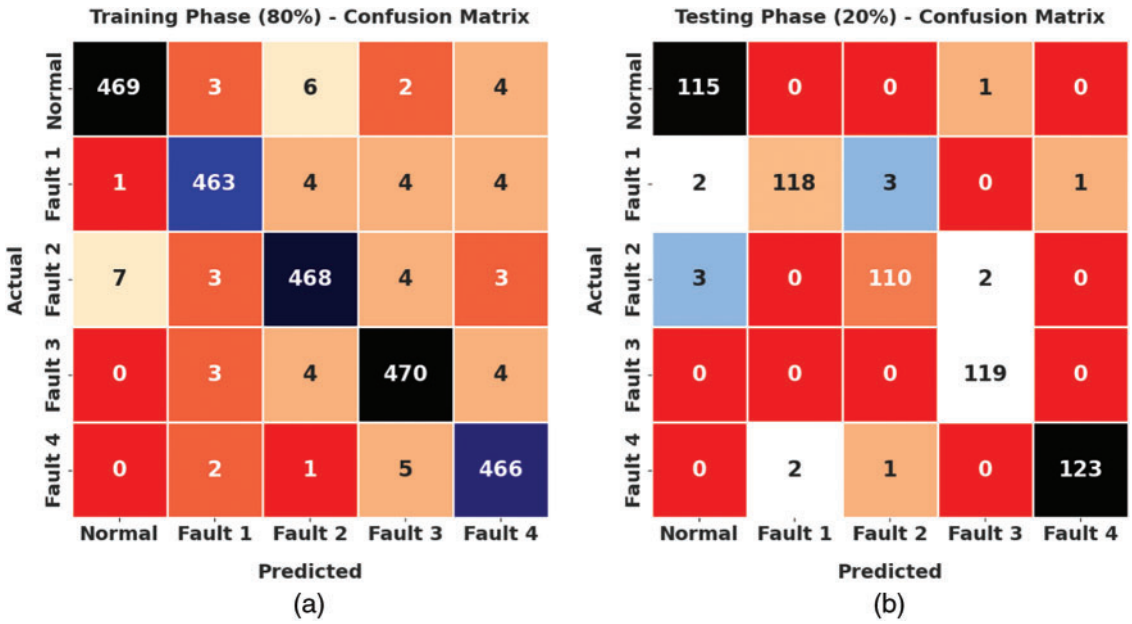

Figure 6: Confusion matrices of ISOSDL-FD approach (a) 80% of TR data and (b) 20% of TS data

Tab. 3 demonstrates a brief classification result of the ISOSDL-FD method under 80% of TR and 20% of TS data. Fig. 7 displays the result analysis of ISOSDL-FD methodology under 80% of TR data. The ISOSDL-FD algorithm has recognized normal class with $accu_y$ of 99.04%, $sens_y$ of 96.90%, $spec_y$ of 99.58%, F_{score} of 97.61%, MCC of 97.01%, and $G_{measure}$ of 97.61%. Furthermore, the ISOSDL-FD approach has recognized Fault-1 class with $accu_y$ of 99%, $sens_y$ of 97.27%, $spec_y$ of 99.43%, F_{score} of 97.47%, MCC of 96.85%, and $G_{measure}$ of 97.47%. Also, the ISOSDL-FD method has identified Fault-4

class with $accu_y$ of 99.04%, $sens_y$ of 98.31%, $spec_y$ of 99.22%, F_{score} of 97.59%, MCC of 97%, and $G_{measure}$ of 97.59%.

Table 3: Result analysis of ISOSDL-FD approach with various measures under 80% of TR and 20% of TS data

Training/Testing (80:20)						
Labels	Accuracy	Sensitivity	Specificity	F-Score	MCC	G-Measure
Training phase						
Normal	99.04	96.90	99.58	97.61	97.01	97.61
Fault 1	99.00	97.27	99.43	97.47	96.85	97.47
Fault 2	98.67	96.49	99.22	96.69	95.86	96.69
Fault 3	98.92	97.71	99.22	97.31	96.63	97.31
Fault 4	99.04	98.31	99.22	97.59	97.00	97.59
Average	98.93	97.34	99.33	97.33	96.67	97.34
Testing phase						
Normal	99.00	99.14	98.97	97.46	96.86	97.47
Fault 1	98.67	95.16	99.58	96.72	95.91	96.73
Fault 2	98.50	95.65	99.18	96.07	95.14	96.07
Fault 3	99.50	100.00	99.38	98.76	98.45	98.76
Fault 4	99.33	97.62	99.79	98.40	97.98	98.40
Average	99.00	97.51	99.38	97.48	96.87	97.49

Fig. 8 illustrates the result analysis of ISOSDL-FD method under 20% of TS data. The ISOSDL-FD algorithm has recognized normal class with $accu_y$ of 99%, $sens_y$ of 99.14%, $spec_y$ of 98.97%, F_{score} of 97.46%, MCC of 96.86%, and $G_{measure}$ of 97.47%. Additionally, the ISOSDL-FD methodology has recognized Fault-1 class with $accu_y$ of 98.67%, $sens_y$ of 95.16%, $spec_y$ of 99.58%, F_{score} of 96.72%, MCC of 95.91%, and $G_{measure}$ of 96.73%. Along with that, the ISOSDL-FD technique has recognized Fault-4 class with $accu_y$ of 99.33%, $sens_y$ of 97.62%, $spec_y$ of 99.79%, F_{score} of 98.40%, MCC of 97.98%, and $G_{measure}$ of 98.40%.

The training accuracy (TA) and validation accuracy (VA) acquired by the ISOSDL-FD technique on test dataset is demonstrated in Fig. 9. The experimental outcome represents the ISOSDL-FD methodology has reached higher values of TA and VA. To be specific, the VA is higher than TA.

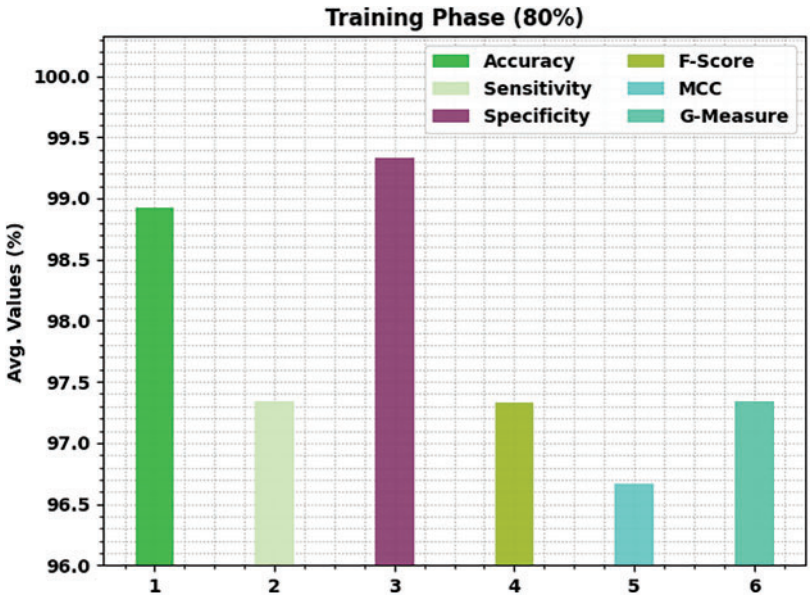

Figure 7: Average analysis of ISOSDL-FD algorithm under 80% of TR data

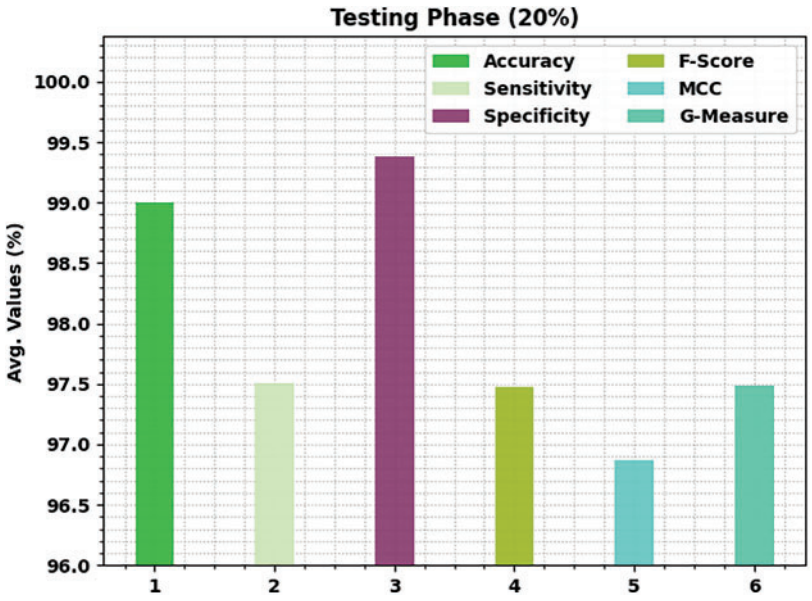

Figure 8: Average analysis of ISOSDL-FD algorithm under 20% of TS data

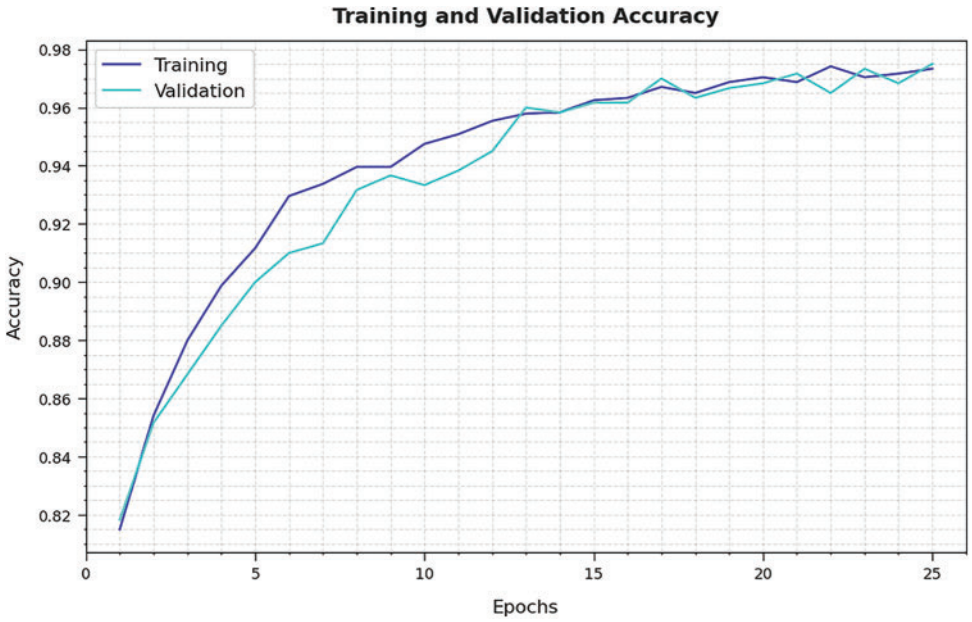

Figure 9: TA and VA analysis of ISOSDL-FD algorithm

The training loss (TL) and validation loss (VL) gained by the ISOSDL-FD method on test dataset are established in Fig. 10. The experimental outcome indicated the ISOSDL-FD technique has accomplished minimal values of TL and VL. Particularly, the VL is lower than TL.

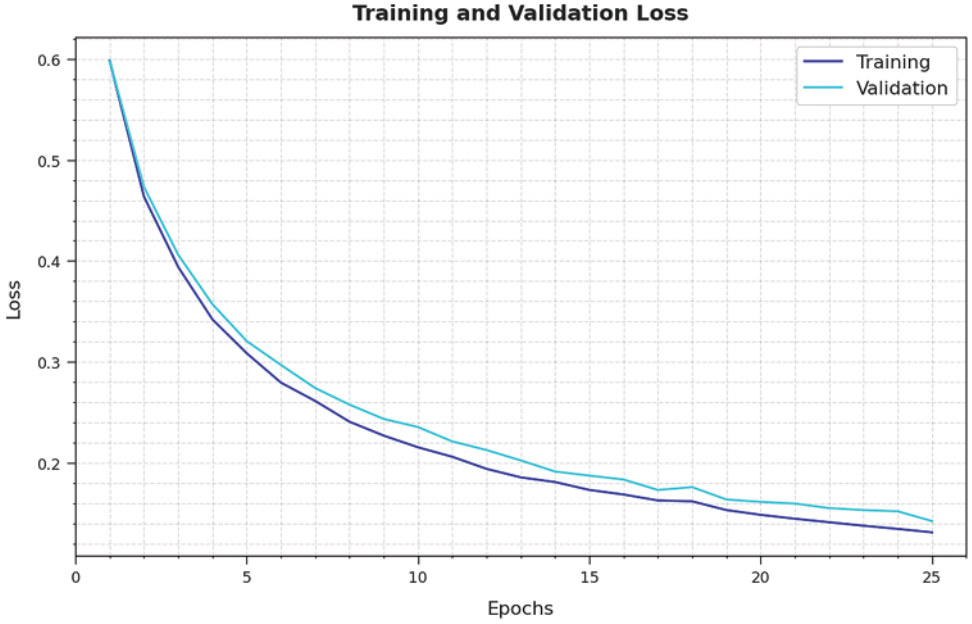

Figure 10: TL and VL analysis of ISOSDL-FD algorithm

At last, a comparative study of the ISOSDL-FD model with recent models interms of different measures is shown in Tab. 4 [22,23].

Table 4: Comparative analysis of ISOSDL-FD approach with existing methodologies

Methods	Accuracy	Sensitivity	Specificity	F-Score
ISOSDL-FD	99.00	97.51	99.38	97.48
CNN	95.60	94.02	95.37	97.13
CNN-FG	98.68	97.04	96.35	96.80
KNN	95.92	94.84	95.33	96.56
SVM	97.66	96.05	98.53	94.83
Autoencoder	94.14	95.61	95.64	96.69
DNN	96.78	95.84	94.55	95.76

Fig. 11 inspects a detailed result analysis of the ISOSDL-FD model with existing models interms of $sens_y$ and $spec_y$. The figure implied that the ISOSDL-FD model has gained enhanced results with maximum values of $sens_y$ and $spec_y$. For instance, with respect to $sens_y$, the ISOSDL-FD model has provided improved $sens_y$ of 97.51% whereas the CNN, CNN-FG, KNN, SVM, AE, and DNN models have resulted to reduced $sens_y$ of 94.02%, 97.04%, 94.84%, 96.05%, 95.61%, and 95.84% respectively. In addition, in connection to $spec_y$, the ISOSDL-FD method has rendered improvised $spec_y$ of 99.38% whereas the CNN, CNN-FG, KNN, SVM, AE, and DNN techniques have resulted to minimal $spec_y$ of 95.37%, 96.35%, 95.33%, 98.53%, 95.64%, and 94.55% correspondingly.

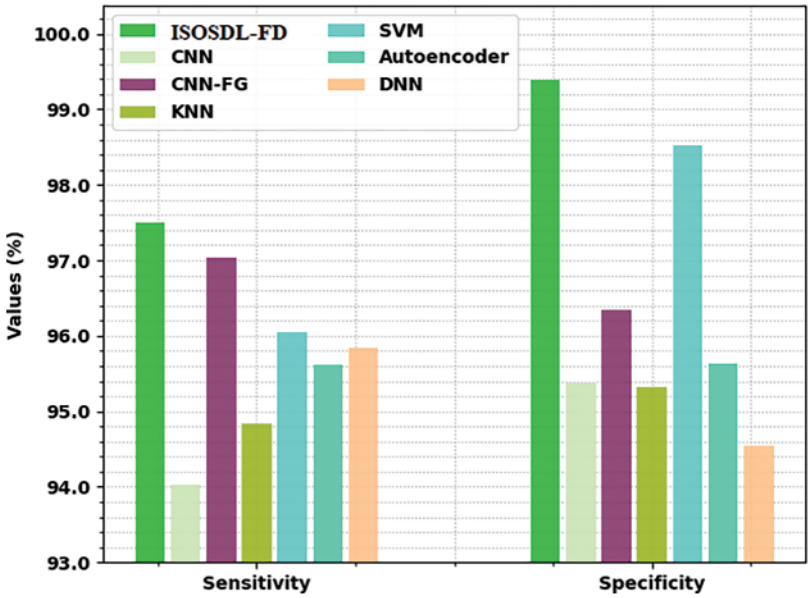

Figure 11: $Sens_y$and $Spec_y$ analysis of ISOSDL-FD approach with existing methodologies

Fig. 12 reviews a comparative result analysis of the ISOSDL-FD methodology with existing models interms of $accu_y$ and F_{score}. The figure inferred the ISOSDL-FD technique has obtained improved results with higher values of $accu_y$ and F_{score}. For example, in relation with $accu_y$, the ISOSDL-FD approach has rendered enhanced $accu_y$ of 99% whereas the CNN, CNN-FG, KNN, SVM, AE, and DNN methods have resulted in reduced $accu_y$ of 95.60%, 98.68%, 95.92%, 97.66%, 94.14%, and

96.78% respectively. Along with that with connection with F_{score}, the ISOSDL-FD model has offered improvised F_{score} of 97.48% whereas the CNN, CNN-FG, KNN, SVM, AE, and DNN approaches have resulted to minimal F_{score} of 97.13%, 96.80%, 96.56%, 94.83%, 96.69%, and 95.76% correspondingly.

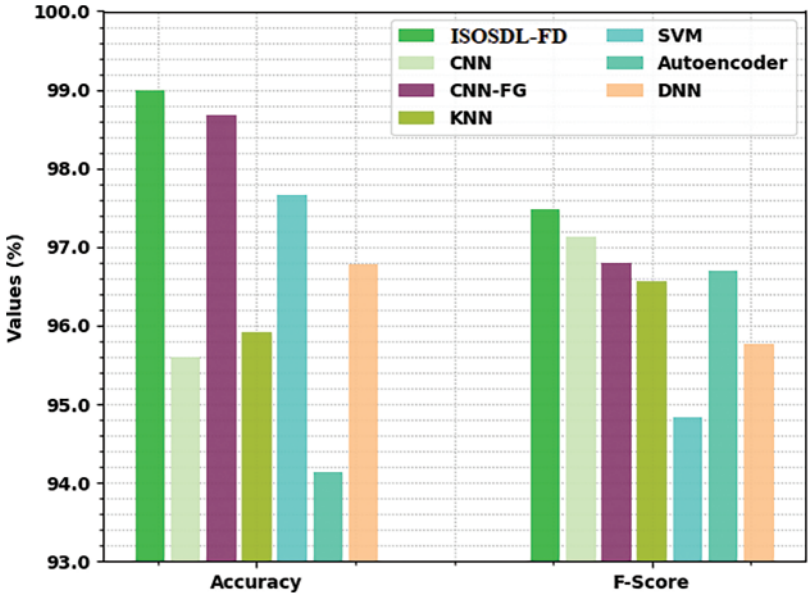

Figure 12: $Accu_y$and F_{score} analysis of ISOSDL-FD approach with existing methodologies

From the detailed results and discussion, it is apparent that the ISOSDL-FD model has shown enhanced classification performance over other models.

5 Conclusion

In this study, a new ISOSDL-FD technique was introduced for the recognition of faults in industrial systems**.** The proposed ISOSDL-FD technique accomplished proficient identification and classification of faults in the gearbox data. In addition, a Fast kurtogram based time-frequency analysis can be used for revealing the energy present in the machinery signals in the time and frequency representation. Moreover, the DBiRNN is applied for fault detection and classification. At last, the ISOS approach was derived for optimal hyperparameter tuning of the DL approach so that the classification performance will be improvised. To exhibit the enhanced performance of the ISOSDL-FD method, a comprehensive experimental analysis was performed. The experimental results stated the betterment of the ISOSDL-FD methodology over recent techniques with maximum accuracy of 99%. In future, the performance of the proposed model can be enhanced by the use of hybrid DL models.

Funding Statement: The author received no specific funding for this study.

Conflicts of Interest: The author declare that they have no conflicts of interest to report regarding the present study.

References

[1] R. Liu, B. Yang, E. Zio and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mechanical Systems and Signal Processing*, vol. 108, pp. 33–47, 2018.

[2] T. Zhou, T. Han and E. L. Droguett, "Towards trustworthy machine fault diagnosis: A probabilistic Bayesian deep learning framework," *Reliability Engineering & System Safety*, vol. 224, pp. 108525, 2022.

[3] Y. Qi, C. Shen, D. Wang, J. Shi, X. Jiang *et al.,* "Stacked sparse autoencoder-based deep network for fault diagnosis of rotating machinery," *IEEE Access*, vol. 5, pp. 15066–15079, 2017.

[4] A. M. Hilal, J. S. Alzahrani, I. Abunadi, N. Nemri, F. N. Al-Wesabi *et al.,* "Intelligent deep learning model for privacy preserving IIoT on 6G environment," *Computers, Materials & Continua*, vol. 72, no. 1, pp. 333–348, 2022.

[5] M. Kim, J. U. Ko, J. Lee, B. D. Youn, J. H. Jung *et al.,* "A domain adaptation with semantic clustering (DASC) method for fault diagnosis of rotating machinery," *ISA Transactions*, vol. 120, pp. 372–382, 2022.

[6] M. A. Alohali, F. N. Al-Wesabi, A. M. Hilal, S. Goel, D. Gupta *et al.,* "Artificial intelligence enabled intrusion detection systems for cognitive cyber-physical systems in industry 4.0 environment," *Cognitive Neurodynamics*, 2022. https://doi.org/10.1007/s11571-022-09780-8.

[7] Y. Ding, L. Ma, J. Ma, M. Suo, L. Tao *et al.,* "Intelligent fault diagnosis for rotating machinery using deep Q-network based health state classification: A deep reinforcement learning approach," *Advanced Engineering Informatics*, vol. 42, pp. 100977, 2019.

[8] I. Abunadi, M. M. Althobaiti, F. N. Al-Wesabi, A. M. Hilal, M. Medani *et al.,* "Federated learning with blockchain assisted image classification for clustered UAV networks," *Computers, Materials & Continua*, vol. 72, no. 1, pp. 1195–1212, 2022.

[9] Y. Zhang, T. Zhou, X. Huang, L. Cao and Q. Zhou, "Fault diagnosis of rotating machinery based on recurrent neural networks," *Measurement*, vol. 171, pp. 108774, 2021.

[10] T. Jin, C. Yan, C. Chen, Z. Yang, H. Tian *et al.,* "Light neural network with fewer parameters based on CNN for fault diagnosis of rotating machinery," *Measurement*, vol. 181, pp. 109639, 2021.

[11] W. Zhang, X. Li and Q. Ding, "Deep residual learning-based fault diagnosis method for rotating machinery," *ISA Transactions*, vol. 95, pp. 295–305, 2019.

[12] L. C. Brito, G. A. Susto, J. N. Brito and M. A. V. Duarte, "An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery," *Mechanical Systems and Signal Processing*, vol. 163, pp. 108105, 2022.

[13] W. Jiang, C. Wang, J. Zou and S. Zhang, "Application of deep learning in fault diagnosis of rotating machinery," *Processes*, vol. 9, no. 6, pp. 919, 2021.

[14] T. Wang, L. Zhang, H. Qiao and P. Wang, "Fault diagnosis of rotating machinery under time-varying speed based on order tracking and deep learning," *Journal of Vibroengineering*, vol. 22, no. 2, pp. 366–382, 2020.

[15] S. Tang, Y. Zhu and S. Yuan, "An adaptive deep learning model towards fault diagnosis of hydraulic piston pump using pressure signal," *Engineering Failure Analysis*, vol. 138, pp. 106300, 2022.

[16] B. Peng, H. Xia, X. Lv, M. A. Nyarko, S. Zhu *et al.,* "An intelligent fault diagnosis method for rotating machinery based on data fusion and deep residual neural network," *Applied Intelligence*, vol. 52, no. 3, pp. 3051–3065, 2022.

[17] Y. Li, X. Du, F. Wan, X. Wang and H. Yu, "Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging," *Chinese Journal of Aeronautics*, vol. 33, no. 2, pp. 427–438, 2020.

[18] S. Pang, X. Yang, X. Zhang and X. Lin, "Fault diagnosis of rotating machinery with ensemble kernel extreme learning machine based on fused multi-domain features," *ISA Transactions*, vol. 98, pp. 320–337, 2020.

[19] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong and M. H. Lim, "Gearbox fault diagnosis using a deep learning model with limited data sample," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6263–6271, 2020.

[20] O. Olabiyi, E. Martinson, V. Chintalapudi and R. Guo, "Driver action prediction using deep (bidirectional) recurrent neural network," arXiv:1706.02257, pp. 1–7, 2017.

[21] M. Y. Cheng and D. Prayogo, "Symbiotic organisms search: A new metaheuristic optimization algorithm," *Computers & Structures*, vol. 139, pp. 98–112, 2014.
[22] O. Essid, H. Laga and C. Samir, "Automatic detection and classification of manufacturing defects in metal boxes using deep neural networks," *PLoS One*, vol. 13, no. 11, pp. e0203192, 2018.
[23] P. Jiang, H. Cong, J. Wang and D. Zhang, "Fault diagnosis of gearbox in multiple conditions based on fine-grained classification cnn algorithm," *Shock and Vibration*, vol. 2020, pp. 1–15, 2020.