



Article D-dCNN: A Novel Hybrid Deep Learning-Based Tool for Vibration-Based Diagnostics

Ugochukwu Ejike Akpudo 🕩 and Jang-Wook Hur *

Department of Mechanical Engineering, Department of Aeronautics, Mechanical and Electronic Convergence Engineering, Kumoh National Institute of Technology, 61 Daehak-ro (Yangho-Dong), Gumi 39177, Gyeongbuk, Korea; akpudougo@gmail.com

* Correspondence: hhjw88@kumoh.ac.kr

Abstract: This paper develops a novel hybrid feature learner and classifier for vibration-based fault detection and isolation (FDI) of industrial apartments. The trained model extracts high-level discriminative features from vibration signals and predicts equipment state. Against the limitations of traditional machine learning (ML)-based classifiers, the convolutional neural network (CNN) and deep neural network (DNN) are not only superior for real-time applications, but they also come with other benefits including ease-of-use, automated feature learning, and higher predictive accuracies. This study proposes a hybrid DNN and one-dimensional CNN diagnostics model (*D*-*dCNN*) which automatically extracts high-level discriminative features from vibration signals for FDI. Via *Softmax averaging* at the output layer, the model mitigates the limitations of the standalone classifiers. A diagnostic case study demonstrates the efficiency of the model with a significant accuracy of 92% (F1 score) and extensive comparative empirical validations.

Keywords: parallel learning; vibration monitoring; fault detection and isolation; convolutional neural network; deep neural network

1. Introduction

Unexpected equipment failure has a significant contribution to production/operation costs and sometimes results in life-threatening situations. This has motivated the development (and improvement) of condition-based maintenance technologies with DL algorithms at their core for accurate modeling and real-time applicability [1]. Vibration monitoring for FDI of industrial components—pumps, bearings, gears, actuators, valves, etc.—has proven effective over the past decades, and with the right diagnostic models, accurate real-time condition monitoring can be achieved for improved safety, remaining useful life extension, cost minimization, and downtime avoidance [2].

Vibration monitoring is one of the most popular (and reliable) FDI techniques for industrial purposes. Its success is highly attributed to the availability of time domain, frequency domain, and time-frequency domain (TFD) signal processing techniques for discriminative feature engineering—hand-crafted feature extraction, selection, and manipulation [3,4]. For instance, the invention of TFD signals processing techniques like the short-time Fourier transform (STFT), empirical mode decomposition (EMD), wavelet transform (WT), Mel frequency cepstral coefficients (MFCCs), etc. provided strong comparative efficiencies over the more traditional statistical time domain and frequency domain techniques for vibration monitoring [4,5]. Against their efficiencies, the reliability of all these hand-crafted features is limited by expensive statistical assumptions and trade-offs. Furthermore, the use of hand-crafted features demands a reasonable level of domain knowledge and the use of conventional machine learning algorithms which are not as efficient as state-of-the-art methods. Further, feature engineering remains one of the major problems faced with the optimum efficiency of these traditional techniques due to the uniqueness in strength (and weakness) of every feature/descriptor for modeling/reflecting condition



Citation: Akpudo, U.E.; Hur, J.-W. D-dCNN: A Novel Hybrid Deep Learning-Based Tool for Vibration-Based Diagnostics. *Energies* 2021, 14, 5286. https://doi.org/ 10.3390/en14175286

Academic Editor: Remigiusz Wisniewski

Received: 2 August 2021 Accepted: 24 August 2021 Published: 26 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). from non-stationary/noisy vibration (or other sensor) measurements [3]. Consequently, ongoing research suggests the use of bioinspired mathematical models with deep architecture (deep learning methods) for automatic feature extraction [6,7].

State-of-the-art reliability studies (and applications) on industrial cyber-physical systems (ICPSs) is indelibly skewing towards DL models—the more efficient artificial intelligence (AI)-based solutions due to the increasing process/system complexities and the inherent need to model them accurately. Thanks to super computers and the fastemerging industry 4.0 revolution, a significant improvement in dynamic modeling, design, maintenance, and high-level decision-making processes is being witnessed across diverse applications with emphasis on optimum utility of available real-time information [1]. Against the traditional methods which rely on hand-crafted feature engineering, these DL methods are fully automated in their architecture to perform both feature engineering and classification (and regression) tasks efficiently. This transition to DL-based models has greatly improved process efficiencies [7], equipment condition monitoring [4], spatiotemporal forecasting [8], and a host of many other solutions [5,9–11]; however, they are faced with challenges ranging from over-fitting, interpretability, optimal hyperparameter selection (and optimization), standardized weight initialization paradigm, and finding the optimal decision criteria between power consumption and performance [1,12]. Nonetheless, considering the need for accurate real-time solutions for ICPS components especially with the growing need for uncertainty modeling, sensor data discrepancies, dynamic environmental and operating conditions, etc., DL methods remain preferable even at the cost of computational power.

Today, there is a plethora of work showcasing diverse algorithms and learning rules for implementing DL-based diagnostics/classification tasks; however, a closer look into our proposed case study suggests the need for exploring these algorithms and subsequently, proposing a befitting FDI model while considering costs, ease-of-use, real-time applicability, and other necessary factors. Most of these algorithms are standalone models which obviously come with their shortcomings and may be component-specific and/or application-specific. On the other hand, hybrid models come with broader efficiencies by integrating the strengths of its constituent standalone models. In our quest for assessing (and validating) the proposed hybrid *D-dCNN* FDI tool, this study chiefly makes the following contributions:

- The proposed algorithm comes with advantages including ease-of-use, automation capabilities, and accuracies/trustworthiness for vibration monitoring of ICPS components. This makes it an almost universal FDI tool for equipment monitoring.
- The proposed algorithm mitigates the exhaustive feature engineering/signal processing steps which are associated with most traditional methods. This minimizes the need for domain knowledge and physics-of-failure (PoF) analysis.
- By integrating the strengths of its constituent standalone models, the proposed model invariably compensates for the constituent models' weaknesses for a more reliable predictive accuracy. Furthermore, with fewer parameterization process, the model complexity/computational remains within an acceptable range when compared to the predictive efficiencies it provides.

The remainder of the paper is organized as follows. Section 2 discusses related works to state-of-the-art vibration-based FDI of ICPS equipment while Section 3 describes the proposed hybrid *D*-*dCNN* FDI diagnostic. An experimental case study is presented in Section 4 while Section 5 discusses its implications alongside open issues inherent from the study. Last, Section 6 concludes the paper.

2. Related Works

Research studies on vibration-based FDI of ICPS equipment have been on the rise, including, but not limited to, pumps, bearings, gears, etc. These have motivated several manufacturers to integrating on-board accelerometers into their products but the challenge of utilizing the complex nonstationary signals from these sensors remains a

problem facing optimum real-time condition monitoring and predictive maintenance of ICPS equipment/components [13]. The past few decades recorded several innovations on advanced signal processing for time domain, frequency domain, and the more robust TFD diagnostic feature extraction with reliable successes; however, because each technique has its unique pros and cons for dynamic modeling alongside significant assumptions/trade-offs during dynamic modeling, their reliability for safety-critical systems are still being questioned [14,15]. Some studies have proposed the use of multi-domain feature extraction for a more comprehensive representation but such methods require high domain expertise, experience, and computational power for optimum performance [4,5,15,16]. On the contrary, state-of-the-art methods suggest the use of the more user-friendly, time-efficient, accurate, and real-time-capable date-driven methods [15]. These methods are primarily formed with artificial neural networks (RNNs) at their core including but not limited to CNNs, DNNs, recurrent neural networks (RNNs), echo state networks (ESNs), etc.

Apparently, the inevitable shift towards data-driven plant-wide monitoring and safety control is currently being challenged by issues such as online fault localization without prior knowledge of target component PoF, adaptive control of highly dynamic systems, sensor data discrepancy, interpretability, standardized weight initialization paradigm, overfitting, model faithfulness/trustworthiness, optimal hyperparameter optimization, and uncertainty modeling [17,18]. Notwithstanding, efforts are currently being put in place for solving these problems with remarkable successes. These include the use of explainable AI (XAI) for understanding the complex connections among multiple neural layers with the goal of revealing why a deep model made a decision with the given inputs(s) [19], transfer learning—using the weights of a successfully pretrained ANN on a set of new inputs to solve a similar task [20,21], adaptive spatiotemporal feature learning [22], etc.

Although highly reliable, standalone DL algorithms have limited efficiencies which may be compensated for by integrating other DL algorithms to form a hybrid model for improved learning and predictive efficiencies [23,24]. For instance, beyond the general retraining issues inherent in closed set classification systems, CNNs are strongly affected by inputs with dynamic transient behaviour while DNNs are easily fooled by inputs due to their high dependence on a priori knowledge [25]. In contrast, RNNs and ESNs have very limited diagnostic capabilities as they are primarily designed for time-series forecasting and learning transient information from inputs [25]. Consequently, the quest for complementing the inefficiencies of diagnostic models (CNNs and DNNs) can be achieved by designing a hybrid network. Hybrid ANN architectures range across series, parallel, and/or series/parallel learning architectures whereby the series learning approach usually provide solutions to prognostic/forecasting problems as seen in most CNN-RNN or DNN-RNN structures [20,26,27]. On the other hand, parallel learning primarily entails a simultaneous training of input(s) by two or more DL models to compute an output(s) in a unified network architecture and are more befitting for classification/diagnostic problems [28-30].

Hybrid models often offer a comparatively better solution than standalone models and with appropriate dynamic modeling processes, their robustness can be explored across diverse disciplines and applications. Their inherent global efficiencies can be well appreciated beyond theoretical validations to changing real-world environments. Particularly, CNNs and DNNs are known for their individual superior discriminative feature extraction and diagnostic capabilities for supervised cases, and have been explored in many works of literature including but not limited to automatic modulation recognition [26], intelligent fault diagnosis for rotating machinery [31], DDoS attack detection [27], and a host of other applications. In this work, we propose a hybrid DNN and one-dimensional CNN diagnostics model which automatically extracts diagnostic features from vibration signals through parallel training and predicts the target labels via *SoftMax Averaging* at the output layer. Experimental results provide comparative advantages over stand–alone diagnostic models.

3. The Proposed *D*-*dCNN* FDI Model

The DNN typically consists of stacked multilayer perceptrons (MLPs) whereby inputs are exported (note-to-node) successively between the layers via an activated forward propagation process. The automatic (supervised) learning process of DNNs by gradient descent enables for minimizing the squared error in the predicted outputs via a back-propagation of weights [27]. Contrary to this architecture, CNNs are basically configured with convolution, pooling, and fully connected layers, whereby the convolution layers act as filters for extracting discriminative features from inputs while the pooling layer reduces the feature dimension for computational efficiency. With the help of the fully connected layer(s)—classical artificial neural network (ANN), multi-label output predictions can be made via nonlinear activation functions.

Arguably, the efficiencies and reliability advantages associated with parallel hybrid networks cannot be overemphasized, especially for classification problems [20,28]; however, certain factors come into play and may increase computational costs, reduce transferability potentials of the model, and heighten model stochasticity and overfitting [30,32,33] and they include the following.

- The number of layers and nodes. In excess, multiple layers adversely affect computational efficiency and may lead to overfitting.
- Batch normalization—a technique used to standardize the inputs to a network accelerates training, and provides some regularization, thereby minimizing generalization error.
- The choice of activation functions. Against the use of a linear activation function, nonlinear activation functions provide improved learning efficiencies; however, for designing hybrid models (parallel), multiple trials reveal that the *SoftMax* activation function is more reliable; unlike the rectified linear activation units (ReLU), Tanh, and Leaky ReLU activation functions that increase model stochasticity, overfitting, and computational power, and flourish in time-series forecasting problems.

The overall pipeline of the proposed model is illustrated in Figure 1 while the full network architecture and parameter values employed in this study are summarized in Table 1.

Consequently, the proposed novel DL-based FDI tool—hybrid *D-dCNN*—is a tree structure with two branches: each branch extracts high-level discriminative features separately in different representations of data; thereby providing more reliable paradigm for making empirical judgments. Each model predicts the probability of the class labels using the *SoftMax* activation function and the final prediction is obtained by averaging the probability of same class label. In this way, high-level features are transmitted simultaneously from the constituent models to the *SoftMax-averaging* layer.

Given a set of multi-class inputs $X_n^m = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, where $x_n \in R^m$ and $y_n \in \{1, 2, \dots, n\}$, as inputs, the *SoftMax*-activated prediction by the CNN and DNN models are summarized in Equations (1) and (2), respectively.

$$R_n^{CNN} = CNN[SoftMax \otimes X_n^m] \tag{1}$$

$$R_n^{DNN} = DNN[SoftMax \otimes X_n^m]$$
⁽²⁾

Outputs obtained from both branches of the hybrid model are averaged using Equation (3).

$$R_n^{Hybrid} = \sum_{i=1}^2 \frac{R_n^{CNN} \oplus R_n^{DNN}}{2}$$
(3)

Like other ANNs, the stochastic learning process (due to random weight initialization) of the proposed model demands a reasonable number of learning iterations to ensure cost function minimization. As a norm, the categorical cross entropy is a befitting loss function for multi-class problems and is defined using Equation (4) [34].

$$L_{CE} = -\sum_{i=1}^{N} T_i \log\left(R_n^{Hybrid}\right)$$
(4)



where T_n is the truth label and R_n^{Hybrid} is the *Softmax* probability for the n^{th} class.

Figure 1. Proposed *D*-*dCNN* network architecture for FDI. C1, MP1, C2, and MP2 represent Convolution layer 1, Maxpooling layer 1, Convolution layer 2, and Max-pooling layer 2, respectively.

Table 1. Architecture of the proposed <i>D</i> - <i>dCNN</i> m	odel.
--	-------

Layer	Output Volume	Description		
Input	<i>m</i> , 1, 80			
Conv1D-1	100, 1, 71	Number of filters: 100, Kernel size: 1×10 , Activation: <i>ReLu</i>		
Dropout	100, 1, 71	Gaussian Dropout: 0.3		
Pool	100, 1, 35	Max-pooling1D: 2×2		
Conv1D-2	50, 1, 26	Number of filters: 50, Kernel size: 1×10 , Activation: <i>ReLu</i>		
Dropout	50, 1, 26	Gaussian Dropout: 0.3		
Pool	50, 1, 13	Max-pooling1D: 2×2		
Dense_CNN	п	Fully connected <i>n</i> units, Activation: <i>SoftMax</i>		
MLP-1	100, 1, 80	Number of nodes: 100, Activation: <i>ReLu</i>		
Dropout	100, 1, 80	Gaussian Dropout: 0.3		
MLP-1	50, 1, 80	Number of nodes: 50, Activation: <i>ReLu</i>		
Dropout	50, 1, 80	Gaussian Dropout: 0.3		
Dense_DNN	п	Fully connected <i>n</i> units, Activation: SoftMax		
SoftMax Average	п	Average: [Dense_CNN, Dense_DNN] Fully connected n units Activation: SoftMax		

Designed to quantify the difference between true and *SoftMax*-predicted probability distributions of a multi-class problem, the objective is to iteratively minimize L_{CE} which invariably ensures accurate input–label modeling process (successful training). This can be monitored visually by observing the training convergence of the model over the iteration process. In addition, cross validation ensures a well–trained model is achieved while also evaluating the model's reliability over multiple trials. This helps eliminate possibilities of *accidental success*, overfitting/underfitting issues, and provide a range of accuracy values/horizon of the model on the test data.

4. Experimental Study

Vibration monitoring has become a highly reliable (and cost-efficient) condition monitoring technique hydraulic pumps; consequently, an experimental case study was conducted on VSC63A5 solenoid pumps at the Defense Reliability Laboratory, Kumoh National Institute of Technology (KIT), Republic of Korea. Figure 2 illustrates the experimental setup and data acquisition process.





Five pumps were operated under the conditions summarized in Table 2, while vibrations signals were collected from the respective casings from high-sensitivity accelerometers with the aid of a NI CompactDAQ 9178 and LabVIEW interface.

Table 2.	Pump	Running	Conditions.
----------	------	---------	-------------

Label	Input Power	Operating Condition	Failure Mode
V-1-CNT	220 V, 60 Hz	4 L Diesel, 5g Paper Ash, 1 L SAE40 Engine Oil	Contaminated Fluid
V-2-VSC	220 V, 60 Hz	3 L Diesel, 1 g Paper Ash, 3L SAE40 Engine Oil	High Viscosity Fluid
V-3-CLG	220 V, 60 Hz	4 L Diesel, 0.2 L Paraffin Solution, 100 g Pectin Powder	Suction Filter Clogging
V-4-NOM	220 V, 60 Hz	5 L Diesel	Healthy Condition
V-5-AMP	300 V, 40Hz	5 L Diesel, (300 V, 40 Hz)	Unspecified power supply

At room temperature and standard humidity level (as suggested by the KS A 0006 environmental standards for tests [35]), the experiments were conducted for approximately 10 days. A full description of the experimental procedure and results can be found in [4].

4.1. Signal Preprocessing and FDI

The vibration signals produced by the solenoid pumps are characterized by different wave-forms and amplitudes; however, using them in their raw unprocessed form will not contribute to the intended work purpose. Analyses from previous studies suggest the need for normalization. This simple but significant signal preprocessing step scales the signal values to a range of {0, 1}, thereby reducing redundancy while also presenting the signals in a format compatible with the hybrid model. Consequently, the raw vibration signals are

normalized, labeled appropriately, and decomposed into small frames. These frames are then fed simultaneously to the CNN and DNN models for parallel training.

As DNNs are less affected by inputs with dynamic transient behaviour than the CNN while CNNs are not easily fooled by inputs with high dependence on a priori knowledge (unlike the DNN) [25], the hybrid model mitigates both models' limitations with better FDI accuracy.

4.2. Test Results

Multiple empirical analyses were performed to confirm the effectiveness of the proposed model. These included a comparison between the proposed hybrid model and its constituent standalone models, past models [4,16] employed for the same purpose, and discussions on the overall investigation/study. Following a successful training and testing (over a 10-fold cross-validation), Figure 3 shows the validation accuracy and losses of the hybrid model and the standalone CNN and DNN models over 200 iterations.



Figure 3. Test/validation progress of proposed model and standalone models over 200 iterations. (a) accuracy and (b) loss.

As observed in Figure 3a, the accuracy of the hybrid model during over the iterations are comparatively higher than the standalone models. This comparative efficiency is also observed in Figure 3b where the hybrid model's losses over the iterations are relatively lower. Overall, the average test accuracy of the hybrid model, CNN, and DNN models over a 10-fold cross-validation was 92%, 90.1%, and 88.9%, respectively. These provide reliable insights on the superiority of the hybrid model; however, we conducted more in-depth comparative analysis on the models using other standard classification evaluation metrics.

Although accuracy provides an insight on the model's comparative efficiencies, more metrics were employed to better understand the local and global performance of the model from different perspectives. Metrics like precision, recall, and F1-score provide reliable global classifier evaluation tools. While precision refers to the percentage of the predicted outputs that are relevant and recall refers to the percentage of total model predictions that are correctly classified, F1-score is another measure of the model's accuracy particularly for unbalanced data-sets. Invariably, F1-score is a more reliable metric for assessing global prediction accuracy of a model. Figure 4 provides a comparative summary of of the models over 200 iterations using the metrics recall, F1-score, and precision.



Figure 4. Test/validation progress of proposed model and standalone models over 200 iterations. (a) Recall (b) F1-score, and (c) Precision.

A visual observation of the green, blue, and red lines representing the training history for the CNN, DNN, and *D-dCNN* models, respectively, the superiority of the *D-dCNN* model is quite obvious as it returns the highest validation scores throughout the iteration process for recall, F1-score, and precision (see Figure 4a–c, respectively). In addition, it is also observed there are less stochasticity (low variance in the training history plots) from the proposed model unlike the other standalone models. This also hints at the model's stability—a yet another important criteria for trusting a model for real-life use.

Beyond the importance of global assessments for diagnostic models, it is also necessary to (as much as possible), assess their local diagnostic efficiencies. The confusion matrix provides a summary of the individual class predictions for class-specific judgments/evaluations. It provides a visual platform for assessing the true positives, true negatives, false positives, and false negatives for each class prediction. To better assess the models' class-specific predictive performance by the proposed model, Figure 5 present respective confusion matrices for the proposed model and standalone models.



Figure 5. Confusion matrices on test samples from (a) DNN model (b) CNN model, and (c) D-dCNN model.

As observed in Figure 4, although the hybrid model (in red lines) retains its superior performance above the other models throughout the the iterations, a closer look at Figure 5c shows there are more false alarms for the pump VSC-3-CLOG with only \approx 11% correctly classified samples; nonetheless, this is more acceptable than the 9.56% and 10% correctly predicted samples produced by the DNN and CNN models, respectively (see Figure 5a,b). A more detailed (averaged) performance evaluation of the models on each operating condition (class) is provided in Table 3.

As shown, the performance of the models across the various pump operating conditions reveal the better diagnostic efficiencies of the hybrid model over the standalone models and more interestingly, the model *SVM*–*RBF* employed in [4] for the same purpose.

Model	Pump Class	Precision	Recall	F1-Score	Accuracy
	VSC-1-CNT	73%	99%	84%	88%
	VSC-2-VISC	87%	99%	92%	88%
DNN	VSC-3-CLOG	90%	48%	62%	88%
	VSC-4-NORM	100%	97%	99%	88%
	VSC-5-AMP	99%	99%	99%	88%
	VSC-1-CNT	100%	100%	100%	90%
	VSC-2-VISC	67%	100%	80%	90%
CNN	VSC-3-CLOG	97%	51%	67%	90%
	VSC-4-NORM	100%	99%	99%	90%
	VSC-5-AMP	100%	100%	100%	90%
	VSC-1-CNT	73%	84%	84%	85%
	VSC-2-VISC	83%	88%	79%	85%
SVM–RBF [4]	VSC-3-CLOG	82%	42%	59%	85%
	VSC-4-NORM	91%	90%	94%	85%
	VSC-5-AMP	95%	93%	96%	85%
D-dCNN	VSC-1-CNT	100%	100%	100%	92%
	VSC-2-VISC	82%	100%	85%	92%
	VSC-3-CLOG	97%	64%	78%	92%
	VSC-4-NORM	100%	99%	99%	92%
	VSC-5-AMP	100%	100%	100%	92%

Table 3. Classification performance summary of diagnostic models on each operating conditions.

Contrary to the shallow opinion that standalone models offer cost efficiency, the authors believe that the marginal extra computational costs associated with the proposed model is quite negligible in comparison with the high predictive accuracies which the model offers. The analyses conducted in this study were made in the Python-based deep learning library Keras with Tensorflow back-end to provide compatibility with GPU environments on a PC whose specification is summarized in Table 4. Table 4 also contains the recorded computation times for the models, respectively.

Computation Time Computation Time Computation Time Manufacturer Processor Speed **RAM Size** (D-dCNN) (DNN) (CNN) Advanced Micro Ryzen 7, 3.20 GHz 16 GB 27.6 Secs 33.8 Secs 35.2 Secs Devices (AMD) 2700 Eight-core

Table 4. Specification of computational hardware used.

As shown in Table 4, under the same environment and computational capacity on the same data set, the DNN and CNN complete their respective tasks in 29.6 s and 33.8 s, respectively. In comparison, they are about 16% and 4% faster then the proposed *D-dCNN* model which completes the same task in 35.2 s. One may argue in favor of using either of the standalone models as they are both more time-efficient; however, considering the increasing need for optimum model trustworthiness especially in industrial safety-critical conditions, the authors believe a more accurate model would be more preferable even at the expense of a marginal increase in computational costs which can be readily compensated for by upgrading computational resources.

5. Discussions and Open Issues

The comprehensive goal of our work is to model a high performing hybrid classifier that can effectively and readily predict fault conditions of ICPS equipment/components from vibration measurements with minimal cognitive experience. The case study presented in this study validates the model's robustness (with minimal losses) for solenoid pump diagnostics with extensive empirical assessments; however, the authors believe the following shortcomings should be put into consideration:

- Stochasticity: Just as it is experienced in virtually all ANNs, stochastic learning process (due to random weight initialization) is a significant factor worthy of consideration when using the proposed model [30,32,33]. It often causes ANN models to produce new (but closely related) prediction outputs for every run. Although cross-validation provides a reliable control paradigm for ensuring model reliability, the overall computational time—time for a complete trial + time for the multiple cross-validation trials—poses significant concerns for real-time applicability. For instance, the effect of stochasticity can be observed in Figures 3a and 4 for all the models' iteration results. Observing the iteration progress for the proposed model in all the figures (in red), it can be seen that at around the 70th and 150th iterations, respectively, there are some spikes which the authors believe were attempts made by the model to further minimize the cost function in Equation (4) but was fairly (on average) maintained uniform (approximate) validation scores of 92%, respectively. Overall, when compared with the standalone models, it is also observed there are less stochasticity (low variance in the training history plots) from the proposed model. This also hints at the model's stability—a yet another important criteria for trusting a model for real-life use [19].
- Parameterization: To date, no globally acceptable paradigm for optimal hyperparameter selection (and optimization) for ANNs exist and this has resulted in overfitting (*in attempts to achieve high model accuracy*) and underfitting (*in attempts to minimize computational times*). These conflicting challenges, although they may be controlled by *Dropout* and *batch normalization*, should be considered when using the proposed model [14]. Ideally, this study does not claim that the proposed model (in the recorded configuration herein) outperforms *all* DNN and CNN standalone models (and their various configurations) as the authors believe there are diverse model configurations out there for comparison. In reality, it would be futile to even attempt to assess and compare all standalone variants; however, under similar configurations of constituent DNN and CNN models, the proposed *D-dCNN* model provides more stable, reliable, and accurate FDI prediction results as verified in the cases study.
- Non-generality: From the results recorded herein, five (5) failure modes were considered for analysis with remarkable diagnostic results from the model. Considering that passive parameter/condition control improves a diagnostics/prognostics scheme (as the faults were pre-designed), the possibility for online performance of the proposed model is still limited since other failure modes are yet to be accounted for. This presents an opportunity for continued research towards assessing the proposed model's efficiencies.

With the findings in this study and strong theoretical support, we believe the proposed hybrid model provides a more reliable alternative to standalone DNN and CNN models for FDI. Although the results verify its effectiveness on vibration signals, it is uncertain at how much reliable the model would perform on different sensor measurements like temperature, pressure, etc.; however, the model's ability to mitigate each constituent ANN's weaknesses by pooling their strengths respectively provides a superior motivation for exploring its effectiveness for FDI cases whereby vibration sensors measurements are not available. With signal processing and data manipulation techniques available for diverse applications, the authors believe the model would flourish just as much (if not better) for other sensor measurements.

6. Conclusions and Future Works

Accurate vibration monitoring and fault detection/isolation demand reliable feature extraction and diagnostic technologies. Against the traditional approach of integrating handcrafted features with machine learning/Bayesian classifiers, deep learning-based algorithms provide better feature learning and diagnostic efficiencies.

This study presents a novel hybrid diagnostic tool—*D-dCNN*—for ICPS equipment diagnostics which consists of parallel DNN and CNN models with *SoftMax Averaging* at the concatenation layer for better feature learning and diagnostic efficiencies. With extensive empirical validations, the proposed hybrid model's strengths were validated on vibration signals from a practical case study with reliable results. Theoretical implications as well as possible extended practical applications of the model were discussed.

The proposed model receives one-dimensional vibration signals as inputs and computes complex neuron–neuron feature extraction simultaneously by the DNN and DNN modules, respectively. These features are received at the fully connected layers for *Softmax*-activated multi-class predictions which are then averaged for optimal target class prediction. The model's performance are compared with its constituent stand–alone models (and other reliable model(s)) using diverse evaluation metrics. Results on a case study reveal the model's superior predictive advantages; however, this advantage comes with increased marginal computational costs.

While future research shall be aimed at obtaining more experimental data to cover other failure modes not recorded herein, ongoing studies are aimed at exploring the model's efficiencies beyond vibration signals while also exploring improvement options for the model to account for its limitations. From a deeper perspective, considering that passive parameter/condition control improves a diagnostics/prognostics scheme (as the faults were pre-designed), the possibility for online performance is still limited. Therefore, future research directions would be geared towards investigating a more real-time functionality for the proposed model.

Author Contributions: Conceptualization, U.E.A.; methodology, U.E.A.; software, U.E.A.; formal analysis, U.E.A.; investigation, U.E.A.; resources, U.E.A. and J.-W.H.; data curation, U.E.A.; writing—original draft, U.E.A.; writing—review and editing, U.E.A. and J.-W.H.; visualization, U.E.A.; supervision, J.-W.H.; project administration, J.-W.H.; funding acquisition, J.-W.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2020-2020-0-01612) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to laboratory regulations.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- 1. Yin, S.; Rodríguez-Andina, J.; Jiang, Y. Real-Time Monitoring and Control of Industrial Cyberphysical Systems: With Integrated Plant-Wide Monitoring and Control Framework. *IEEE Ind. Electron. Mag.* **2019**, *13*, 38–47. [CrossRef]
- 2. Guo, J.; Li, Z.; Li, M. A Review on Prognostics Methods for Engineering Systems. IEEE Trans. Reliab. 2020, 69, 1110–1129. [CrossRef]
- 3. Khaire U.M.; Dhanalakshmi, R. Stability of feature selection algorithm: A review. J. King Saud Univ.-Comput. Inf. Sci. 2019. [CrossRef]
- 4. Akpudo, U.E.; Hur, J. A Cost-Efficient MFCC-Based Fault Detection and Isolation Technology for Electromagnetic Pumps. *Electronics* **2021**, *10*, 439. [CrossRef]
- 5. Akpudo, U.E.; Hur, J. An Automated Sensor Fusion Approach for the RUL Prediction of Electromagnetic Pumps. *IEEE Access* **2021**, *9*, 38920–38933. [CrossRef]
- Mahony, N.O.; Campbell, S.; Carvalho, A.; Harapanahalli, S.; Velasco-Hernández, G.; Krpalkova, L.; Riordan, D.; Walsh, J. Deep Learning vs. Traditional Computer Vision. *arxiv* 2019, arXiv:1910.13796. [CrossRef]
- Weigold, M.; Ranzau, H.; Schaumann, S.; Kohne, T.; Panten, N.; Abele, E. Method for the application of deep reinforcement learning for optimised control of industrial energy supply systems by the example of a central cooling system. *CIRP Ann.* 2021, 70, 17–20. [CrossRef]
- 8. Medrano, R.; Aznarte, J.L. A spatio-temporal attention-based spot-forecasting framework for urban traffic prediction. *Appl. Soft Comput.* 2020, *96*, 106615. [CrossRef]

- 9. Cabaneros, S.M.; Calautit, J.K.; Hughes, R. A review of artificial neural network models for ambient air pollution prediction. *Environ. Model. Softw.* **2019**, *119*, 285–304. [CrossRef]
- 10. Ahishakiye, E.; Bastiaan, M.B.; Tumwiine, J.; Wario, R.; Obungoloch, J. A survey on deep learning in medical image reconstruction. *Intell. Med.* **2021**. [CrossRef]
- Kim, H.; Lee, W.; Kim, M.; Moon, Y.; Lee, T.; Cho, M.; Mun, D. Deep-learning-based recognition of symbols and texts at an industrially applicable level from images of high-density piping and instrumentation diagrams. *Expert Syst. Appl.* 2021, 183, 115337. [CrossRef]
- 12. Sharma, O. Deep Challenges Associated with Deep Learning. In Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 14–16 February 2019; pp. 72–75. [CrossRef]
- 13. Vikas, R.D.; Kundan, P.; Pratik, K.; Atharva, R.; Amod, R. Failure analysis of fuel pumps used for diesel engines in transport utility vehicles. *Eng. Fail. Anal.* 2019, *105*, 1262–1272. [CrossRef]
- 14. Akpudo, U.E.; Hur, J.W. Towards bearing failure prognostics: A practical comparison between data-driven methods for industrial applications. *J. Mech. Sci. Technol.* **2020**, *34*, 4161–4172. [CrossRef]
- 15. Chen, H.; Jiang, B.; Ding, S.X.; Huang, B. Data-Driven Fault Diagnosis for Traction Systems in High-Speed Trains: A Survey, Challenges, and Perspectives. *IEEE Trans. Intell. Transp. Syst.* 2020. [CrossRef]
- 16. Akpudo, U.; Jang-Wook, H. A Multi-Domain Diagnostics Approach for Solenoid Pumps Based on Discriminative Features. *IEEE Access* 2020, *8*, 175020–175034. [CrossRef]
- 17. Jiang, Y.; Yin, S.; Kaynak, O. Data-Driven Monitoring and Safety Control of Industrial Cyber-Physical Systems: Basics and Beyond. *IEEE Access* **2018**, *6*, 47374–47384. [CrossRef]
- Wang, D.; Tsui, K.L. Two novel mixed effects models for prognostics of rolling element bearings. *Mech. Syst. Signal Process.* 2018, 99, 1–13. [CrossRef]
- 19. Ribeiro, M.T.; Singh, S.; Guestrin, C. Why Should I Trust You: Explaining the Predictions of Any Classifier. arXiv 2016, arXiv1602.04938.
- 20. Han, T.; Liu, C.; Yang, W.; Jiang, D. Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application. *ISA Trans.* **2020**, *97*, 269–281. [CrossRef] [PubMed]
- 21. Wei, Y.; Zhang, Y.; Yang, Q. Learning to Transfer. *arXiv* 2017, arXiv:1708.05629. Available online: https://arxiv.org/abs/1708.05629 (accessed on 13 October 2020).
- 22. Han, T.; Liu, C.; Yang, W.; Jiang, D. An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems. *Mech. Syst. Signal Process.* **2019**, *117*, 170–187. [CrossRef]
- 23. Rai, R.; Sahu, C.K. Driven by Data or Derived Through Physics? A Review of Hybrid Physics Guided Machine Learning Techniques With Cyber-Physical System (CPS) Focus. *IEEE Access* 2020, *8*, 71050–71073. [CrossRef]
- Suganthi, M.; Sathiaseelan, J. An Exploratory of Hybrid Techniques on Deep Learning for Image Classification. In Proceedings of the 2020 4th International Conference On Computer, Communication And Signal Processing (ICCCSP), Tamilnadu, India, 28–29 September 2020; pp. 1–4. [CrossRef]
- 25. Jirak, D.; Wermter, S. Potentials and Limitations of Deep Neural Networks for Cognitive Robots. *arXiv* **2018**, arXiv:1805.00777. Available online: https://arxiv.org/abs/1805.00777 (accessed on 12 April 2020).
- 26. Njoku, J.; Morocho-Cayamcela, M.; Lim, W. CGDNet: Efficient Hybrid Deep Learning Model for Robust Automatic Modulation Recognition. *IEEE Netw. Lett.* 2021, *3*, 47–51. [CrossRef]
- Bhardwaj, A.; Mangat, V.; Vig, R. Hyperband Tuned Deep Neural Network With Well Posed Stacked Sparse AutoEncoder for Detection of DDoS Attacks in Cloud. *IEEE Access* 2020, *8*, 181916–181929. [CrossRef]
- Zhang, Y.; Zhou, Y.; Lu, H.; Fujita, H. Traffic Network Flow Prediction Using Parallel Training for Deep Convolutional Neural Networks on Spark Cloud. *IEEE Trans. Ind. Inform.* 2020, 16, 7369–7380. [CrossRef]
- 29. Kang, B.; Jeong, J.; Jeong, C. Distributed Parallel Deep Learning for Fast Extraction of Similar Weather Map. In Proceedings of the 2018 IEEE Region 10 Conference, Jeju, Korea, 28–31 October 2018; pp. 1426–1429. [CrossRef]
- 30. Gavrishchaka, V.; Yang, Z.; Miao, R.; Senyukova, O. Advantages of hybrid deep learning frameworks in applications with limited data. *Int. J. Mach. Learn. Comput.* **2018**, *8*, 549–558. [CrossRef]
- Tang, S.; Yuan, S.; Zhu, Y. Convolutional Neural Network in Intelligent Fault Diagnosis Toward Rotatory Machinery. *IEEE Access* 2020, 8, 86510–86519. [CrossRef]
- 32. Panda, P.; Aketi, S.; Roy, K. Toward Scalable, Efficient, and Accurate Deep Spiking Neural Networks With Backward Residual Connections, Stochastic Softmax, and Hybridization. *Front. Neurosci.* **2020**, *14*, 653.[CrossRef] [PubMed]
- 33. Verma, A.; Liu, Y. Hybrid deep learning ensemble model for improved large-scale car recognition. In Proceedings of the 2017 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computed, Scalable Computing Communications, Cloud Big Data Computing, Internet Of People And Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA, USA, 4–8 August 2017; pp. 1–7. [CrossRef]
- Chen, C.H.; Lin, P.H.; Hsieh, J.G.; Cheng, S.L.; Jeng, J.H. Robust Multi-Class Classification Using Linearly Scored Categorical Cross-Entropy. In Proceedings of the 2020 3rd IEEE International Conference on Knowledge Innovation and Invention (ICKII), Kaohsiung, Taiwan, 21–23 August 2020; pp. 200–203. [CrossRef]
- KS A 0006:2001. K. Standard Atmospheric Conditions for Testing. Korean Agency For Technology And Standard, Eumseong-gun, Republic Of Korea. 2014, pp. 1–2. Available online: https://infostore.saiglobal.com/en-us/standards/ks-a-0006-2001-639645_saig_ksa_ ksa_1524462/ (accessed on 13 January 2021).