Transfer Learning Strategies for Anomaly Detection in IoT Vibration Data

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

21-10-2021 / 27-10-2021

CITATION

Heistracher, Clemens; Jalali, Anahid; Strobl, Indu; Suendermann, Axel; Meixner, Sebastian; Holly, Stephanie; et al. (2021): Transfer Learning Strategies for Anomaly Detection in IoT Vibration Data. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.16844521.v1>

DOI

[10.36227/techrxiv.16844521.v1](https://doi.org/10.36227/techrxiv.16844521.v1)

Transfer Learning Strategies for Anomaly Detection in IoT Vibration Data

Clemens Heistracher
Digital Safety and Security
Austrian Institute of Technology
clemens.heistracher@ait.ac.at

Anahid Jalali
Digital Safety and Security
Austrian Institute of Technology
anahid.jalali@ait.ac.at

Indu Strobl
Distributed AI Systems
Siemens AG Austria
indu.strobl@siemens.com

Axel Suendermann
Distributed AI Systems
Siemens AG Austria
axel.suendermann@siemens.com

Sebastian Meixner
Distributed AI Systems
Siemens AG Austria
sebastian.a.meixner@siemens.com

Stephanie Holly
Distributed AI Systems
Siemens AG Austria
stephanie.holly@siemens.com

Daniel Schall
Distributed AI Systems
Siemens AG Austria
daniel.schall@siemens.com

Bernhard Haslhofer
Digital Safety and Security
Austrian Institute of Technology
bernhard.haslhofer@ait.ac.at

Jana Kemnitz
Distributed AI Systems
Siemens AG Austria
jana.kemnitz@siemens.com

Abstract—An increasing number of industrial assets are equipped with IoT sensor platforms and the industry now expects data-driven maintenance strategies with minimal deployment costs. However, gathering labeled training data for supervised tasks such as anomaly detection is costly and often difficult to implement in operational environments. Therefore, this work aims to design and implement a solution that reduces the required amount of data for training anomaly classification models on time series sensor data and thereby brings down the overall deployment effort of IoT anomaly detection sensors. We set up several in-lab experiments using three peristaltic pumps and investigated approaches for transferring trained anomaly detection models across assets of the same type. Our experiments achieved promising effectiveness and provide initial evidence that transfer learning could be a suitable strategy for using pre-trained anomaly classification models across industrial assets of the same type with minimal prior labeling and training effort. This work could serve as a starting point for more general, pre-trained sensor data embeddings, applicable to a wide range of assets.

Index Terms—datasets, anomaly classification, vibration signal, framework

I. INTRODUCTION

Due to the recent advances in the development and installation of IoT sensor platforms, there is also an increasing need for data-driven maintenance activities, such as anomaly detection or error classification. Standardized sensor tags can nowadays be deployed to various industrial assets to record data, which is then relayed to a server infrastructure for further analysis. The industry expects a massive reduction of resource consumption and costs by implementing a generic data-driven maintenance solution, which can be deployed across industrial assets with minimal configuration and prior training effort. A major challenge in industrial setups is the positive economic impact of a machine learning model. This

is the case when the costs and effort to create, deploy and update the model are below the savings gained, for instance, through downtime reduction or an increased asset life span. Therefore, models that can be deployed to multiple assets with no to little additional training are required.

However, building such a generic solution is challenging because anomalous behavior of industrial assets depends on numerous real-world factors, such as minimal divergence in production, wear and tear, or environmental factors like current fluctuation. All this is reflected in sensor data recordings of individual assets and must be considered when building an anomaly detection model and, most importantly, when training and applying models across assets.

Ongoing, directly related efforts are not transferable across assets because they rely on labels collected for a specific asset. The creation of such a dataset is costly, as it should cover all operating conditions, including asset faults, which usually occur infrequently and are hard to recreate in simulation environments [1], [2]. Therefore, this work aims to design and implement a solution that reduces the required amount of data for training anomaly detection models, which should reduce the overall deployment effort of IoT sensors across industrial assets of the same type.

Our approach addresses this real-world problem and is based on two observations: first, unlabeled sensor data for industrial assets of the same type are often available in large quantities because of low marginal costs. Second, labeled sensor data are rare because this requires rather expensive manual interventions and annotations. We believe that transfer learning could be a possible strategy to optimize the use of available labeled and unlabeled datasets across industrial assets and reduce the overall number of samples required for reliable predictions.

For gaining a better understanding of technical feasibility, we set up an in-lab experimental environment with three peristaltic pumps. We then used these pumps to simulate various operating conditions ranging from normal operation to anomalous behavior potentially leading to faults. We collected vibration sensor data from all industrial assets (pumps) and then, to devise a transfer learning strategy that best uses available labeled and unlabeled data, ran a series of experiments that should help us answer the following questions.

Q1: If we have labeled data for one asset, to what extent can we use a trained model for another asset of the same type under varying sample sizes?

We approached this question by collecting vibration data from various assets and then trained a Convolutional Neural Network (CNN) classifier on one asset and applied the model to another one. As a reference benchmark, we trained and applied a model on the same asset. We found that models, which are applied on another asset without transfer learning, yield a lower F1-score (0.61) compared to the reference model (0.89) that was trained and evaluated using Short Term Fourier Transform Features (STFT).

Q2: To what extent can we employ transfer learning across assets, and which feature(s) are most effective?

We pre-trained a classifier with data from one asset and then transferred and evaluated it on another asset. The pre-trained model with STFT features showed average F1-score improvements of 0.14 over all sample sizes and had the highest effectiveness gains for small training samples. The improved F1-score was 0.86 for 5 training samples and 0.94 for the maximum number of samples.

Q3: If no labeled training data is available, to what extent can we employ unsupervised learning to improve overall effectiveness?

We trained a CNN-autoencoder on unlabeled data from two assets and used it as a feature embedding to extract a compressed, lower-dimensional representation of the signal. This improved the F1-score by 0.07 on average for STFT with a slightly stronger impact on cases with few training samples. At five training samples, the embedded features achieved a score of 0.74 instead of 0.63 by the benchmark.

Our experiments achieve promising effectiveness and provide initial evidence that transfer learning could be a strategy for applying pre-trained anomaly detection models across industrial assets with minimal prior labeling and training effort. They especially show strong effectiveness gains for small numbers of training samples, which is a realistic scenario in industrial applications. Our combined approach benefits from labeled as well as unlabeled data and is, in our opinion, the first step towards optimized use of all available datasets. Further, our autoencoders can serve as a starting point for more general, pre-trained sensor embeddings, similar to MobileNet [3] for images and Word2Vec [4] for text.

Several obstacles still must be overcome before a system can be put in operation: first, the experimental scope must be

widened to other types of assets and are more comprehensive dataset and label collection. Second, further experiments on the architecture of autoencoders are required and long short-term memory [5] or transformer-based networks [6] should be evaluated as potential feature extractors. Third, regarding the transfer learning itself, more advanced methods like gradually freezing layers need to be evaluated for our architecture.

As a starting point for further investigations, we make our peristaltic pump dataset and our model implementations available for further research (<https://git-service.ait.ac.at/dsai-public/TFAD>). We now commence by providing further background on transfer learning architectures and state of the art IoT anomaly detection approaches in Section II. Next, in Section III, we will introduce our dataset, propose our transfer learning strategy, and demonstrate our experimental results. Finally, we will discuss and conclude our work in Sections IV and V, respectively.

II. BACKGROUND

Transfer learning is a technique that uses learned knowledge on another related task. When applied for anomaly detection, one can use transfer learning to train a model on a general task with a large, high-quality dataset, which is then fine-tuned for another task with fewer available samples. That significantly reduces the number of required training samples and tackles the problem of insufficient ground-truth data, which is omnipresent in many application domains. So far, transfer learning strategies have often been used for image and text data and have become popular in application domains such as healthcare [7], chemistry [8], biology [9], or computer vision [10]. Recently, transfer learning is also gaining more attention in smart industrial applications (industry 4.0) for machine health prognostics tasks such as fault detection [11].

However, relatively little is known about the applicability of transfer learning techniques for sensor data, most importantly vibration sensor data, a well-known feature in machine health prognostics (c.f. [12]). Pittino et al. [13] discuss the importance of statistical learning approaches for sensor data anomaly classification. They experimented with several well-known classification methods (Support Vector Machine) and point out that the challenge does not rely on the algorithmic choice but instead on data availability and adequacy.

In another research line, Wen and Keyes [14] propose a transfer learning framework for time series anomaly detection. They used a large-scale pre-trained model, fine-tuned it using a much smaller dataset, and applied a CNN architecture for the time series segmentation process. They tested their proposed approach on both real and synthetic data and achieved high precision. The few false positives and false negatives were due to the missing data of the occurred events.

We expand these works by investigating several transfer learning strategies on a dataset collected from industrial assets of the same type, which, in our case, are peristaltic pumps operated in a controlled in-lab environment. We set up several experiments to compare the effectiveness of models trained on features such as Short-Term Fourier Transform

Fig. 1: x (—), y (—) and z (—) for one sequence of 512 samples of the vibration signal from pump 3 during 250 Hz.

Spectrogram (STFT), which is a standard approach to encode a frequency spectrum that varies over time [15]. Inspired by the success of CNNs in image classification [16], we also use Gramian Angular Field (GAF) and Mel-Frequency Spectrogram (MEL) [17] features, as they provide a 2-dimensional representation of a time series, which a CNN can process as image data. Additionally, such transformations to images are commonly used in the related audio domain [18], [19].

III. EXPERIMENTS AND RESULTS

With our transfer learning experiments, we aimed at devising a strategy that efficiently (re)uses available labeled and unlabeled sensor data across industrial assets, which were peristaltic pumps in our case. In this section, we present more details on the collected dataset and the design as well as the results of our experiments, each of which aims to answer one of the questions outlined in Section I.

A. Dataset

Our experimental setup comprised three peristaltic pumps typically used for sterile or aggressive liquids. The flow of liquids is induced by a repeating sequential compression of a flexible tube that pushes the liquid in one direction. We mounted sensor tags on the pumps' rotational axis, documented their angles and horizontal axes (sensor 1: 0°, sensor 2: 45°, sensor 3: 0°), and recorded three-axis vibration data for each pump using the standard configuration of the used IoT sensor platform, which was 512 samples per minute with a sample rate of 6644 Hz. We rotated sensor 2 by 45° to evaluate the robustness of our transfer methods against rotation.

We operated the pumps under varying conditions: idle state and standard rotational speed of 100 Hz, which are considered as a healthy state, and five conditions, which are considered as non-healthy or an anomaly: a rotational speed of 50 Hz, 150 Hz, 200 Hz, and 250 Hz. In total, we collected at least 888 samples per healthy class and at least 134 samples per non-healthy class for each pump. Table I provides a more detailed summary of our dataset.

TABLE I: The number of samples per class and the total numbers for the healthy and the anomalous state.

Class	Freq. [1/min]	Pump 1	Pump 2	Pump 3
0	Idle	1118	888	1131
1	50	137	134	134
2	100	936	920	929
3	150	149	148	148
4	200	149	144	146
5	250	147	150	147
$\sum$ Healthy	-	2054	1808	2060
$\sum$ Anomalous	-	582	576	575

The imbalance between healthy and non-healthy, anomalous states in our dataset is on purpose and reflects the practical limitations of acquiring non-healthy data in real-world setups. We also acquired data as independent measurement series on different days and, for subsequent experiments, took data from the first pump as training data and data from the other two pumps as test data. Figure 1 shows the unprocessed vibration signal during 250 Hz from the train set of pump 3. Our dataset's dimension reflects the typical constraints of IoT sensor platforms due to limited bandwidth and battery power.

B. Experiment 1: No Transfer Learning

In a first step, to increase robustness and rotational invariance of the machine learning model, we normalized the 3-dimensional vibration signal from the sensor coordinate systems to the same world coordinate system using the Kabsch algorithm [20]. Next, we augmented data by applying a rolling window and splitting one sensor sample of dimension 512 into overlapping fragments of dimension 128 with an overlap of 16 steps. We ensured that all samples originating from a single sample in either the test or the training dataset remain in the corresponding set to avoid information leakage due to auto-correlation in the data.

We performed our experiments using four feature variants, which are either 1-dimensional or 2-dimensional representations of the signal. Our feature selection is inspired by methods from the audio and image recognition domain:

- *Raw Signal*: the raw sensor signal, which is pre-processed and augmented as described above
- *Short-Time Fourier Transformation (STFT)*: a 2-dimensional representation of the signal, which is created by splitting the signal into samples of length 32 and applying a Fourier transform to each. The result is a matrix in which each column is the powers of the Fourier coefficients for each segment.
- *MEL spectrograms (MEL)*: Mel Frequency Coefficients are a method to extract features from audio signals used

in speech recognition and information retrieval in audio. Additionally to the STFT, logarithmization, grouping, and principal component analysis are performed [21].

- **Gramian Angular Field (GAF):** is an image obtained from a time series, representing some temporal correlation between each time point [22].

For anomaly detection, we trained a supervised CNN classification model with 32 filters, a kernel size of 5, and a relu activation function followed by a max-pooling layer of dimension 3, a dropout layer of 0.2, and a classification layer with a sigmoid activation function. Experiments on 2-dimensional signals are performed using a 2-dimensional CNN with the corresponding properties with 16 filters, 3×3 kernel, and a softmax activation layer. We implemented our models using Tensorflow Keras and ran our experiments on a Titan RTC GPU.

As a baseline for our transfer learning experiments, we trained models on datasets with few data samples from a single pump, representing the scenario without any knowledge transfer across assets. More specifically, we define two reference benchmarks:

- **Benchmark 1 (BM1):** we trained and evaluated the classifier on a single pump (no. 3) and, to evaluate the impact of the number of training samples on the result, performed each experiment multiple times with an increasing number of training samples x with $x \in [1, 3, 5, 10, 15, 20, 25, 30]$. The samples are randomly selected and reused for all further experiments. To ensure reliable results, we performed a variant of cross-validation and drew each sample four times and present the average results.
- **Benchmark 2 (BM2):** this benchmark is evaluated on the same dataset as BM1 and all further experiments but trained using a dataset from another pump recorded with the sensor rotated by 45° .

Finally, we measured the effectiveness of our models using the micro-averaged F1-score, which is the harmonic mean of the standard metric for classification precision and recall averaged over all samples [23].

Figure 2 shows the achieved F1 scores over the growing number of training samples per class. We can observe a gradual increase of accuracy for models using STFT and MEL-spectrogram features and lower overall effectiveness of GAF features. The highest F1-score of 0.61 was achieved with the maximum number of samples, while BM1, represented by the upper dotted line, achieved a maximum score of 0.89.

The result of our first experiment indicates that direct model transfers lead to effectiveness losses and points out the need for more advanced transfer techniques. For further reference, we denote the maximum score of this experiment as Benchmark 2 (BM 2) and plot it as the lower dotted line.

C. Experiment 2: Transfer Learning with Labeled Data

This experiment is inspired by the scenario in which annotated samples are missing for a specific asset but available for

Fig. 2: No Transfer Learning. Models were trained and tested on *different* assets with features GAF(—), MEL(—), STFT(—), and RAW(—). The best result (---) at 30 samples defines the lower baseline, the best result of BM1 trained and tested on the *same* asset with STFT features (---) defines the lower baseline (---) for our experiments.

another one. We evaluated to what extent a model can benefit from pre-training with labeled data from other assets.

We trained a classifier on labeled data from another asset and applied it to the actual asset after fine-tuning it with fewer samples. To investigate the impact of the number of training samples, we performed experiments with an increasing number of samples and, for the sake of computational efficiency, saved the classifier after the initial pre-training and re-loaded it for transfer learning and fine-tuning. To ensure reliable results, we performed each experiment four times with randomly selected samples. We pre-trained the classifier using data from pump 1 (0°) and pump 2 (45°) separately and evaluated it on pump 3 against BM1.

Our results indicate that pre-training on a labeled dataset from another pump can boost training performance with very few samples compared to the benchmark. The average improvement to the benchmark's F1-score is +0.14, a relative improvement of 18 %. The most substantial improvement could be shown for cases with few training samples, while the impact decreases with the sample size for experiments with 15 or more samples per class. In our experiments that use only five samples per class for training, we achieve an F1-score of 0.86 compared to 0.63 in the benchmark, which is an improvement of 26 %. The F1-score over samples per class is illustrated in Figure 3 for features based on MEL and STFT spectrograms.

Our results demonstrate that supervised pre-training can improve the F1-score for classification with few training samples using annotated data from other assets.

D. Experiment 3: Transfer Learning with Unlabeled Data

In our next experiment, we examined to what extent we can employ unsupervised transfer learning techniques if no

Fig. 3: Transfer Learning with Labeled Data. Supervised pre-training (—), BM1 with Train/Test on same asset (---), BM2 with Train/Test on different asset (—, —), Baseline defined by best result at 30 samples for Train/test on same (· · ·) and different (· · ·) asset.

labeled training data is available for a specific asset or from other machines.

For that purpose, we implemented an autoencoder that learns features from STFT, GAF, or MEL-images. An autoencoder consists of an encoder that creates a lower-dimensional representation of the input, called the bottleneck, and the decoder part that reconstructs the original signal. The autoencoder is trained on unlabeled data to minimize the root mean squared error between the input signal and output signal. After training, only the encoder part is used to extract lower-dimensional features.

Our autoencoder implementation consists of three stacked CNNs for the encoder and the decoder with alternating max-pooling or up-sampling, respectively. For classification, we clip the autoencoder at the bottleneck and attach a classification layer to it, which is trained while the layers of the autoencoder are frozen. The fine-tuning and evaluation of the classifier remains as described in Experiment 1 and Experiment 2 in this section.

We trained the autoencoder with three variants of datasets to provide insights on the impact of the classes contained in the train set. As anomaly classes are harder to obtain in real life, we perform experiments with healthy data only.

Fig. 4: Transfer Learning with Unlabeled Data. Unsupervised pre-training with all classes (—), Unsupervised pre-training with all healthy classes (—), BM1 with Train/Test on same asset (---), BM2 with Train/Test on different asset (—, —) Baseline defined by best result at 30 samples for Train/Test on same (· · ·) and different (· · ·) asset.

Our embeddings improved the F1-score for STFT features and even in cases with few samples, as illustrated in Figure 4. The average improvement for the F1-score compared to the BM1 is 0.07. However, the embedding's positive effect diminishes for experiments with 15 or more samples per class.

Our results indicate that a vibration signal embedding of STFT can improve classification results for cases with few features.

Our experiments conclude that the addition of an embedding to our supervised pre-training method does not improve the F1-score for anomaly classification. Therefore, we state that the training of an embedding is not required in case annotated samples from other assets are available.

IV. DISCUSSION

In this paper, we investigated to what extent it is possible to transfer anomaly detection models, which were trained using vibration data of one industrial asset, to another asset of the same type. We evaluated i) the direct transfer of a model between assets, ii) supervised pre-training with labeled data from other machines, iii) unsupervised pre-training with an autoencoder as a feature extractor.

Our experiments demonstrate that pre-training with labeled data from other assets improves the F1-score compared to non-transfer learning. In case only unlabeled data is available, we can achieve improvements with only a few training samples from the target asset. It appears that sufficiently annotated samples can compensate the effect of an embedding. Overall, our work implies that both supervised and unsupervised methods should be considered for transferring models, which are trained on vibration sensor data, across industrial assets of the same type.

One limitation of our approach is that the dataset was recorded in a controlled environment and that it contains only data from a specific asset type. Hence we can not yet make any statement to what extent general, pre-trained anomaly detection models are transferable across asset types.

Subsequently, we can identify three challenges that need to be tackled for a profound transfer learning strategy on IoT sensor platforms. First, the dataset must be extended to include more types of assets. Second, the effects of additional feature engineering and novel transfer learning strategies must be explored. Finally, we believe that model explainability approaches could help in understanding and improving the transferability of models and ultimately help in achieving the goal of a generic, pre-trained anomaly detection model, which can easily be transferred across industrial assets.

V. CONCLUSION

Lack of labeled training data is currently a major obstacle that prevents the detection of anomalies in IoT sensor data recorded from industrial assets. In this paper, we followed the idea that trained models should be transferable across assets of the same type and substantiated this intuition with a series of experiments demonstrating overall feasibility and promising effectiveness gains when applying a transfer learning strategy that considers both supervised and unsupervised approaches. By that, we address a significant real-world problem that currently prevents the adoption of model-based anomaly detection approaches in industry and gained some fundamental insight on the current possibilities and most promising future directions. With this empirical evidence and the already available datasets, we can now commence and investigate more powerful, possibly generally deployable anomaly detection approaches based on publicly shared vibration sensor data.

REFERENCES

- [1] Y. Zhang, P. Hutchinson, N. A. Lieven, and J. Nunez-Yanez, "Adaptive event-triggered anomaly detection in compressed vibration data," *Mechanical Systems and Signal Processing*, vol. 122, pp. 480–501, 2019.
- [2] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to rul prediction," *Mechanical systems and signal processing*, vol. 104, pp. 799–834, 2018.
- [3] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
- [4] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.
- [5] Y.-A. Chung, C.-C. Wu, C.-H. Shen, H.-Y. Lee, and L.-S. Lee, "Audio word2vec: Unsupervised learning of audio segment representations using sequence-to-sequence autoencoder," *arXiv preprint arXiv:1603.00982*, 2016.
- [6] K. Choi, C. Hawthorne, I. Simon, M. Dinculescu, and J. Engel, "Encoding musical style with transformer autoencoders," in *Proceedings of the 37th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, H. D. III and A. Singh, Eds., vol. 119. PMLR, 13–18 Jul 2020, pp. 1899–1908. [Online]. Available: <http://proceedings.mlr.press/v119/choi20b.html>
- [7] R. K. Samala, H.-P. Chan, L. M. Hadjiiski, M. A. Helvie, K. H. Cha, and C. D. Richter, "Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms," *Physics in Medicine & Biology*, vol. 62, no. 23, p. 8894, 2017.
- [8] J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev, and A. E. Roitberg, "Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning," *Nature communications*, vol. 10, no. 1, pp. 1–8, 2019.
- [9] A. Kaya, A. S. Keceli, C. Catal, H. Y. Yalic, H. Temucin, and B. Tekinerdogan, "Analysis of transfer learning for deep neural network based plant classification models," *Computers and electronics in agriculture*, vol. 158, pp. 20–29, 2019.
- [10] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection," *Construction and Building Materials*, vol. 157, pp. 322–330, 2017.
- [11] P. Cao, S. Zhang, and J. Tang, "Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning," *Ieee Access*, vol. 6, pp. 26 241–26 253, 2018.
- [12] R. K. Mobley, *An Introduction to Predictive Maintenance (Second Edition)*. Burlington: Butterworth-Heinemann, 2002.
- [13] F. Pittino, M. Puggl, T. Moldaschl, and C. Hirschl, "Automatic anomaly detection on in-production manufacturing machines using statistical learning methods," *MDPI-Sensors*, vol. 20, no. 8, p. 2344, 2020.
- [14] T. Wen and R. Keyes, "Time series anomaly detection using convolutional neural networks and transfer learning," *arXiv preprint arXiv:1905.13628*, 2019.
- [15] E. Sejdić, I. Djurović, and J. Jiang, "Time–frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S105120040800002X>
- [16] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural computation*, vol. 29, no. 9, pp. 2352–2449, 2017.
- [17] T. Hasegawa, J. Ogata, M. Murakawa, and T. Ogawa, "Tandem connectionist anomaly detection: Use of faulty vibration signals in feature representation learning," in *2018 IEEE International Conference on Prognostics and Health Management (ICPHM)*. IEEE, 2018, pp. 1–7.
- [18] G. Yu and J. Slotine, "Audio classification from time-frequency texture," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 1677–1680.
- [19] L. Cremer and M. Heckl, *Structure-borne sound: structural vibrations and sound radiation at audio frequencies*. Springer Science & Business Media, 2013.
- [20] F. L. Markley, "Optimal matrix algorithm," *The Journal of the astronomical Sciences*, vol. 41, no. 2, pp. 261–280, 1993.
- [21] M. Dörfler, R. Bammer, and T. Grill, "Inside the spectrogram: Convolutional neural networks in audio processing," in *2017 International Conference on Sampling Theory and Applications (SampTA)*. IEEE, 2017, pp. 152–155.
- [22] R. Damaševičius, R. Maskeliūnas, M. Woźniak, and D. Połap, "Visualization of physiologic signals based on hjorth parameters and gramian angular fields," in *2018 IEEE 16th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*. IEEE, 2018, pp. 000 091–000 096.
- [23] D. M. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.