




Article

Monitoring Pumping Units by Convolutional Neural Networks for Operating Point Estimations

Hanbing Ma ^{*}, Lukas Gaisser  and Stefan Riedelbauch 

Institute of Fluid Mechanics and Hydraulic Machinery, University of Stuttgart, 70569 Stuttgart, Germany

^{*} Correspondence: hanbing.ma@ihs.uni-stuttgart.de; Tel.: +49-711-685-60234

Abstract: To avoid the failure of pumping units, the monitoring of operating points with a subsequent assessment of the condition of the pump may support the decision for required maintenance. For that purpose, convolutional neural networks (CNNs) are implemented to predict the operating points of pumping units. Instead of using traditional flowmeter and manometer, vibration and acoustic signals are used to estimate the head and volume flow rate. An appropriate pre-processing of raw data is applied, enabling our method to predict well on different datasets. For the datasets measured in an anechoic chamber, the best model of each subset achieves relative errors smaller than 4.9% for the prediction of head and 7.6% for the volume flow rate. For cases where only small amounts of data exist, it is furthermore demonstrated that transfer learning from one dataset to another dataset provides an improvement in performance.

Keywords: standard water pump; operating point estimations; convolutional neural networks



Citation: Ma, H.; Gaisser, L.; Riedelbauch, S. Monitoring Pumping Units by Convolutional Neural Networks for Operating Point Estimations. *Energies* **2023**, *16*, 4392. <https://doi.org/10.3390/en16114392>

Academic Editors: Bernhard Semlitsch and Christian Bauer

Received: 17 March 2023

Revised: 26 May 2023

Accepted: 26 May 2023

Published: 29 May 2023



Copyright: © 2023 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/>).

1. Introduction

Nowadays, pumping units are installed in many plants. Due to the long operating time, wear and tear cannot be prevented, and the extent of deterioration in most technical applications depends on the operating points and the pumped fluid. To avoid the failure of the pump, the monitoring of operating points with a subsequent assessment of the condition of the pump may support the decision for required maintenance. Ultimately, the risk of sudden failure of the pump system could be minimized and at the same time, too frequent maintenance is avoided as much as possible while ensuring the safe operation of pumping units.

In pump systems, vibration signals and acoustic signals are commonly used to identify the operating condition or diagnose mechanical faults. Chao et al. proposed a multi-sensor fusion model for leakage levels classification of an axial piston pump [1]. By comparing the contribution of three sub-models based on three vibration signals and adjusting the class probability distribution, the accuracy is increased. J Černetič analyzed the vibration and noise spectra for cavitation detection [2]. It is observed that the spectral differences between cavitation and non-cavitation are much larger in the frequency range above 1000 Hz. Casoli et al. proposed a vibration-based method for fault identification. Through feature extraction and feature reduction, different classifiers are tested and satisfying accuracy is achieved [3]. Feng and Böhle conducted a feasibility study for implementing a neural network (NN) for operating the condition detection of a centrifugal pump [4]. Using a simple NN with only one hidden layer, the vibration signal combined with volume flow class is utilized to detect the operating condition, achieving the best congruence rate of 92.86%. Nguyen and Huang proposed a fault-detection method in water pumps using sound signals [5]. The method shows high accuracy, regardless of whether the test set is from the same pump as the train set or a different one. Furthermore, accelerometers and microphones are easier and more flexible to install or exchange than flowmeters and pressure sensors, eliminating the need to change the piping. Hence, we are inspired to monitor the operation condition of pumping units with vibration and acoustic signals.

In recent years, neural networks (NNs) [6] have rapidly gained acceptance and shown outstanding competence in the field of image processing and feature extraction. Instead of the great dependence of traditional methods on the experience and background knowledge, neural-network-based approaches extract the useful features from input signals automatically. Huang et al. proposed a hybrid neural network incorporating a theoretical loss model to predict the energy performance of centrifugal pumps [7]. It achieves low mean-square-error values for head, power, and efficiency, outperforming linear regression and the original neural network. Wu et al. developed a flow rate estimation model based on a back-propagation neural network, using valve opening as the input variable [8]. The drawback of the affinity law's inapplicability due to significant speed differences can be disregarded. He et al. proposed a four-layer neural network to identify the four typical flow patterns in a centrifugal pump's impeller. The application of the SMOTE (synthetic minority oversampling technique) algorithm for oversampling the original data resulted in a substantial improvement in the model's identification rate, increasing it from 89.91% to 94.88% [9]. The greater flexibility of NNs enables the proposed method to be implemented in a variety of situations, even on existing plants, where there was not enough metering equipment originally installed. CNN-based algorithms develop rapidly and are successfully applied in many industrial fields, such as fault diagnosis [10,11] and cavitation detection [12–14].

In this paper, CNNs are implemented to predict the operating points of pumping units. The vibration signal from accelerometers, which are installed on the surface of the bearing housing, or the noise during operation measured in the anechoic chamber, was selected as the input variables. The head and volume flow rate of the pump, which are used to characterize the pump performance, are selected as output variables. The motivation behind this is to see if it is possible to utilize machine learning methods to find a relationship between vibration/noise and operating conditions. If feature extraction works well, it has the potential to be extended to other tasks, such as fault diagnosis, cavitation detection and so on. Additionally, it is planned to measure data on different pumps to address the generalization problem in the future.

In actual situations, it is extremely time consuming and expensive to collect a huge amount of data to train a neural network. Furthermore, the data are usually measured under different experimental setups and hence have diverse distributions. To address this problem, the transfer learning method is investigated in this work, which focuses on reusing the learned knowledge from one task to another similar task.

The main contributions of this paper are summarized as follows:

- A test rig consisting of a standard centrifugal water pump, a motor with a frequency converter, water piping, and a sensor system is built, and an adequate data-acquisition system is designed. Vibration signals in three directions and acoustic signals in two measurement locations under different working conditions are collected.
- An appropriate pre-processing method of raw data is adopted, enabling the proposed method to predict well on different datasets. The proposed pipeline is verified on three datasets and shows satisfying performance on all datasets.
- To further improve the performance of the proposed method and obtain a robust and generalized model, the transfer learning method is investigated. The results are promising.

2. Methods

2.1. Network

Generally, a convolutional neural network consists of three types of layers [15]: convolutional layer, pooling layer, and fully connected layer (Figure 1). Three-dimensional filters in convolutional layers make it possible to extract specific features in the images. The pooling layer is used to reduce the spatial size of the feature maps. To some extent, using a stride bigger than 1 in the convolutional layer realizes a similar effect (in Figure 1: '/2' in convolutional block means the dimension of feature maps is reduced to half with

stride = 2). The last layer of the network is a fully connected layer, the flattened feature matrix passes through this layer, and the output is a vector.

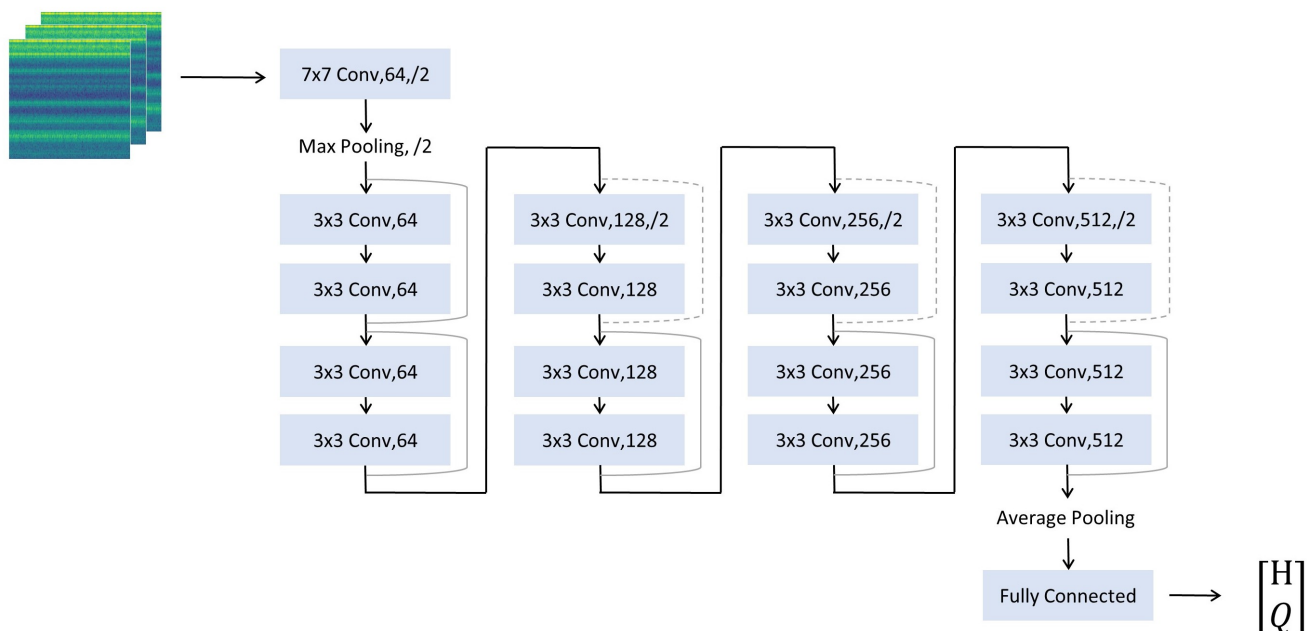


Figure 1. Network architecture. The network consists of 17 convolutional layers, 2 pooling layers and 1 fully connected layer. The input is a preprocessed spectrogram of vibration or sound signal. The output is the head and volume flow rate of the pump.

For deep neural networks, the problem of vanishing gradients sometimes occurs. To solve such a problem, He et al. [16] proposed a network implementing residual blocks instead of plain convolutional blocks, which specifically add an identity shortcut connection between intermediate layers. If x is the input and $F(x)$ is the output of the plain convolutional block, the basic definition of the output of the residual block is $F(x) + x$. For some blocks, including the downsampling operation, which means the output $F(x)$ and identity input x have different dimensions, an additional convolution operation is implemented to make the input to the required size (dashed lines in Figure 1). The definition is updated accordingly as $F(x) + Wx$, where W is a matrix of learnable parameters. The residual network is shown to outperform plain convolutional neural networks for various image-recognition tasks.

In our specific cases, there is a limited amount of data. To avoid the overfitting problem that occurs when training large networks with small datasets, the 18-layer ResNet with around 11 million trainable parameters is chosen for this task. The ResNet 18, Figure 1, consists of residual blocks with the following kernel numbers: 64, 128, 256, and 512. The output of the network is a 2-dimensional vector representing the head and volume flow rate. Hence, the output size of the fully connected layer is modified to 2. The input shape of the network is as default $(N, 3, 224, 224)$, where N is the batch size.

The training process of the network can be described as iterative parameter optimization over a train set. Starting with randomly initialized parameters, the network makes a prediction of the head and volume flow rate. The predefined loss function, as a metric for describing the error between the current prediction and true value (measured operating point), is calculated. The adopted loss function MSE loss represents the mean square error (squared L2 norm) between prediction x and target y . It is described as

$$l(x, y) = \text{mean}(\{l_1, \dots, l_N\}^T), l_n = (x_n - y_n)^2 \quad (1)$$

where N is the batch size. The derivatives of the loss function with respect to parameters are then computed. The parameters are optimized using gradient descent. New predictions can be generated and the next iteration is started. After repeating multiple times, the model converges to optimal parameters. Once the network is trained, the network is able to provide a prediction for each given vibration or acoustic signal input.

2.2. Data Preprocessing

The one-dimensional vibration signals and the acoustic data signal are split into one-second segments. To eliminate the influence of systematic biases in the measurement, the data are normalized by subtracting the mean and a division by the standard deviation.

The short time Fourier transform (STFT) is applied to produce the spectrogram. The value of each pixel in the spectrogram represents the amplitude of a specific frequency at a certain time point. As this transform contains both time and frequency information, it is a powerful preprocessing method. STFT is also widely employed for classification tasks and natural language processing. The spectrogram is computed using a Tukey window of length 256 with a shape parameter of 0.25 and 50% overlap. The spectrogram has a shape of $128 \times \lfloor \frac{n_s - 128}{128} \rfloor$ and a frequency range of $[0, \frac{f_s}{2}]$, where n_s is the number of data points in the segment and f_s represents the sampling rate.

Instead of directly using the spectrogram as an input feature of the neural network, the pixel-wise logarithm operation is performed on the spectrogram, showing considerably better performance in CNN experiments. The size of the spectrogram is kept constant.

Due to the structure of the network, the log spectrogram is resized to 224×224 using a resize function in the scikit-image package, which performs a spline interpolation. After that, min-max normalization is applied to the resized feature, which generally speeds up the training process. At last, the normalized log spectrogram is expanded and repeated three times in the first dimension, which is the expected input shape of the network. The preprocessing procedure is shown in Figure 2.

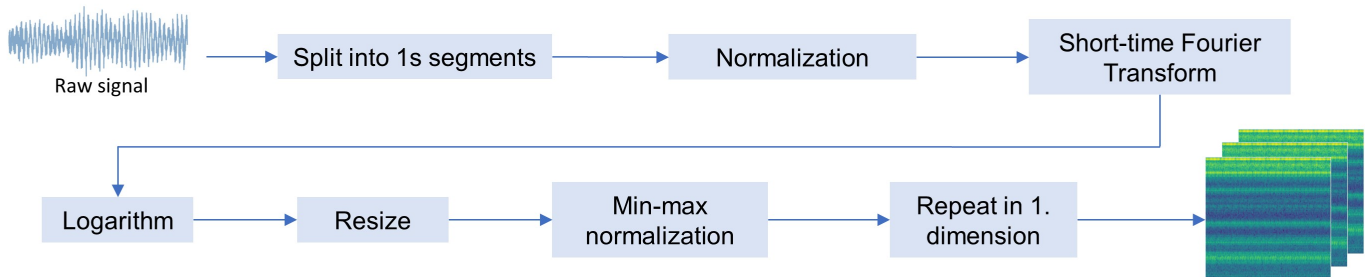


Figure 2. Preprocessing. The one-dimensional time signal is converted into an image representation.

In addition, the head is not directly measured. It is calculated based on the equation

$$H = \frac{\Delta p}{\rho g} + \frac{1}{2g} \left(\left(\frac{Q}{A_d} \right)^2 - \left(\frac{Q}{A_s} \right)^2 \right) + h, \quad (2)$$

where Δp is the static pressure difference between the discharge side and suction side, A_d and A_s are pipe cross-section areas, and h is the height difference between the pressure sensors on both sides.

2.3. Dataset

Three datasets are used to verify the effectiveness of the proposed method (Table 1). Dataset 1 consists of data collected from different sensors under increasing motor rotational speed (from 0 to 3500 r/min). A standard water pump P0 features a specific speed $n_q = 25.0$

with an impeller diameter of 219 mm. The specific speed characterizes the impeller shape and is calculated according to

$$n_q = n \frac{\sqrt{Q}}{H^{3/4}}, \quad (3)$$

where n is the pump rotational speed (rpm), Q is volume flow rate (m^3/s) at the point of best efficiency and H is the head (m) at the point of best efficiency.

Table 1. Available datasets.

	Dataset 1	Dataset 2	Dataset 3
Pump	P0	P1, P3	P1, P2, P3
Specific speed	25.0	15.6	15.6
Measurement condition	Increasing motor rotating speed, constant valve opening	Increasing motor rotating speed, constant valve opening	Constant rotating speed, changing valve opening
Sampling rate of acceleration and sound pressure	50 kHz	60 kHz	60 kHz

Each measurement records a signal with a length of 30 s. The slowly changing variables, such as volume flow rate, head, temperature, and fluid pressure, are sampled at 100 Hz, while fast-changing signals, e.g., acceleration in three directions and sound pressure, are collected with a sampling rate of 50 kHz.

Dataset 2 and dataset 3 are data obtained from another test bench. The three standard water pumps P1, P2, and P3 feature a specific speed $n_q = 15.6$ with an impeller diameter of 209 mm. The schematic diagram of the test bench and the location of accelerometers are presented in Figure 3. To capture key features of pumping units without the influence of environmental noise, the pumping unit is placed in an anechoic chamber during the first phase of this project.

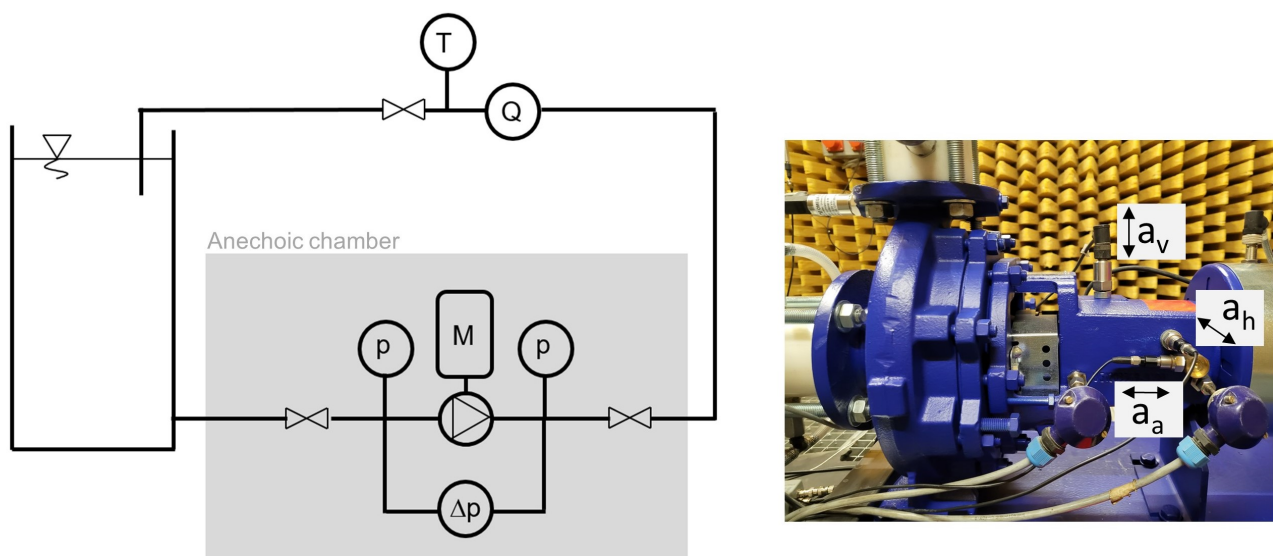


Figure 3. (Left) Schematic diagram of test bench (M : Motor; T : thermometer; Q : flowmeter; p , Δp : manometer). (Right) Locations of accelerometers in three directions. Accelerometers and microphones are located in the anechoic chamber.

The pump sucks water from the reservoir and pumps it back into the same reservoir. The built-in throttle valve allows a change in the volume flow rate and head at the pump and thus serves for operating point adjustment. To exchange the pumps, two additional ball valves are installed in the piping system upstream and downstream of the pumping unit. All measured operating conditions do not show cavitation at the pump, except about 10% of pump P1 in dataset 2 at very high volume flow rates.

Three accelerometers are magnetically fixed on the surface of the bearing housing between the pump and the motor. Type KS74C-100 (range: ± 60 g, sensitivity: $100 \pm 5\%$ mV/g) is used for the measurements in axial and horizontal directions, type KS80D (range: ± 55 g, sensitivity: $100 \pm 5\%$ mV/g) is adopted for the measurements in vertical direction. Additionally, two MM210 microphones (frequency range is 3.5 Hz to 20 kHz, and sensitivity is 50 mV/Pa) from Microtech Gefell are hung in the air. Each measurement takes a signal with a length of two seconds. The sampling rates of above mentioned slowly changing variables and fast-changing variables are 1 kHz and 60 kHz, respectively.

Dataset 2 is measured under similar test conditions as dataset 1, with constant valve opening (about 70%) and increasing rotational speed (from 100 to 3000 r/min). The dataset consists of data from two pumps: P1 and P3. For the measurement of different pumps, the position of the pumps and sensors is kept consistent.

Dataset 3 is measured under six constant rotational speeds: 500, 950, 1160, 1500, 2100, and 2400 r/min. This dataset consists of data from three pumps: P1, P2, and P3. For every speed, the valve opening is adjusted from totally open to almost closed so that the volume flow rate covers the admissible operation range.

3. Validation of Proposed Method

Experimental Setup

The experiments are conducted on GeForce GTX 1080 Ti with CUDA version 11.4. The open source PyTorch is used for the network construction and training. In the standard training process, the Adam optimizer with a learning rate of 5×10^{-4} and weight decay of 8×10^{-4} is used [17].

The dataset is randomly divided into a train set (60% of data), a validation set (20% of data), and a test set (20% of data). The training process is supervised learning, which means that not only the spectrograms but also the target outputs are input during the training process. The measured operating points are considered the target outputs.

For standard training, network architecture ResNet18 is trained from scratch with a random initialization. The network is trained for 500 epochs. During the training process, the parameters are updated iteratively. At the end of each epoch, the MSE loss of the model on the validation set is calculated and compared with the smallest validation loss so far. If the validation loss of current epoch is smaller, this model is to be saved. When all epochs are completed, the saved model is then used to predict the operating points and evaluate the performance on the test set.

For transfer learning, Figure 4, the idea is to transfer the weights learned from dataset A to another dataset B [18]. There are different ways to conduct transfer learning, for example, fix the parameters of convolutional part and only fine-tune the fully connected layer. Based on the results of pre-experiments, the following way is chosen: instead of random initialization, the network is initialized with a pretrained network (dataset A). Then the pretrained network is fine-tuned with another target dataset (dataset B). The rest of the training process remains the same as the standard training, and the best model is selected by evaluating the performance on the validation set during the fine-tuning process. The selected model is then used to predict the operating points on the test set.

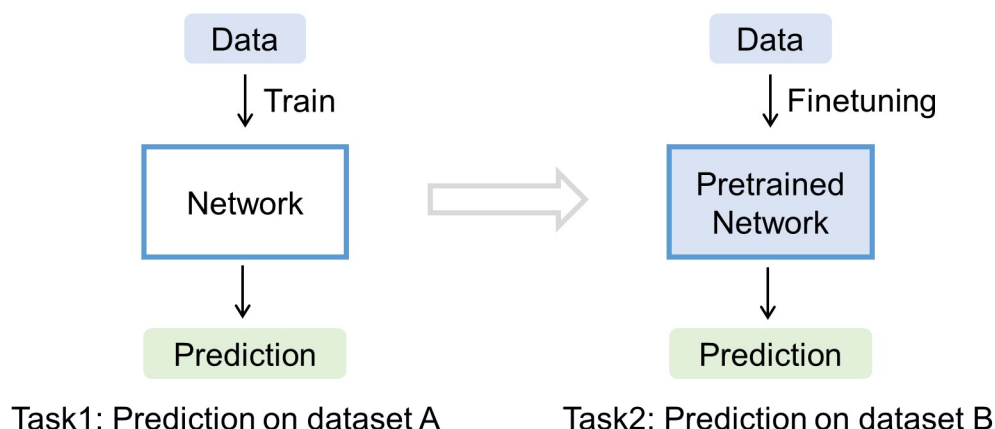


Figure 4. Transfer learning. The network pretrained on dataset A is used as the initialization of the training process on dataset B.

To remove the influence of random initialization, each of the runs of the numerical experiment is repeated five times, and the final loss is calculated as the mean loss of five repetitions.

4. Results

To reduce randomness, the final predictions are also calculated as the mean of predictions over five repetitions. The mean square errors of the test set using different input signals are listed in Table 2. For dataset 1, it is observed that the losses of dataset 1 are significantly larger than the other two datasets. The mean square error is a metric describing the absolute error. The value range of dataset 1 is substantially larger compared with the other two datasets. Hence, a relative error is calculated to evaluate the prediction enabling the comparison of data with different scales. For operating points close to zero, very large relative error values might occur, while the corresponding absolute error is small. Hence, only those operating points (H, Q) that satisfy $H > 0.1H_{max}$ and $Q > 0.1Q_{max}$ are included into the statistics, where H_{max} and Q_{max} are the maximal head and volume flow rate within the dataset.

Table 2. Mean square error using different signals as input (lowest errors of subset marked in bold).

Input Signal	Mean Square Error					
	Dataset 1 P0	Dataset 2 P1	Dataset 2 P3	Dataset 3 P1	Dataset 3 P2	Dataset 3 P3
Axial acceleration	106.51	4.89	3.21	2.59	4.37	4.75
Vertical acceleration	74.71	6.97	4.60	6.87	3.09	3.27
Horizontal acceleration	44.66	21.13	2.14	3.19	3.95	3.21
Sound pressure 1	247.43	7.47	20.90	2.94	3.25	3.75
Sound pressure 2		2.81	1.47	3.02	2.18	3.93

The relative error of the head as a percentage value is calculated as follows:

$$e_r\% = \frac{1}{N_T} \sum_{i=1}^{N_T} \left(\left| \frac{H_i - \hat{H}_i}{H_i} \right| \times 100\% \right), \hat{H}_i = \text{mean}(\{\hat{H}_{i,1}, \dots, \hat{H}_{i,5}\}^T), \quad (4)$$

where N_T is the number of samples in the test set after filtering points near zero, H_i and \hat{H}_i are the true head and mean predicted head of the i -th operating point. The relative error of volume flow rate is similarly calculated.

Table 3 shows the relative error of head and volume flow rate using different signals as input on three datasets. For dataset 1, the horizontal acceleration exhibits the best performance compared with other input signals. The model trained with sound pressure as input shows the largest relative error. A possible reason for this might be that the test bench for dataset 1 is not located in an anechoic room, which means that noise from other equipment is present during the measurement. The next stage of the research will focus on the methods with noisy input signals, such as filter, data augmentation methods and so on.

Table 3. Relative error in percent of head and volume flow rate using different signals as input (lowest relative errors of subset marked in bold).

Input Signal	Relative Error in Percent (Head/Volume Flow Rate)					
	Dataset 1 P0	Dataset 2 P1	Dataset 2 P3	Dataset 3 P1	Dataset 3 P2	Dataset 3 P3
Axial acceleration	10.7/12.2	4.0/2.2	8.2/4.8	4.5/7.4	5.4/10.3	4.6/9.4
Vertical acceleration	6.7/13.3	7.9/4.1	6.7/4.5	5.3/12.0	5.1/9.8	3.7/8.1
Horizontal acceleration	8.5/7.3	14.7/8.3	8.4/4.7	4.6/8.1	4.3/10.0	3.0/7.6
Sound pressure 1	15.5/13.7	8.5/5.1	11.4/5.9	4.2/8.8	4.7/7.5	3.4/10.9
Sound pressure 2		3.7/2.4	4.9/2.6	6.4/8.2	3.5/6.2	2.6/8.3

For datasets 2 and 3, sound pressure 1 is from the microphone near the motor (1 m distance), while the microphone of sound pressure 2 is located near the pump (0.3 m distance). In dataset 2, the signal of sound pressure 2 shows the lowest loss in both subsets, outperforming other signals. In dataset 3, the sensors showing the best performance on the three subsets are different. It is also observed that the relative errors of the head are smaller than the relative errors of the volume flow rate. We speculate that this might be due to the fact that the distribution of the head data is clustered around certain values because the characteristic curve presents a flat trend in a large range. It makes the problem of predicting the head shift from a regression problem to a simpler classification task.

Figure 5 shows the results of models with the input signal that achieve lowest relative error for each subset. A pair of the mean prediction (blue) and the corresponding true operating point (red) is connected using a straight line. Obviously, the shorter the line between two points, the more accurate the network's prediction.

Figure 6 presents the result of transfer learning. In the first case, the feasibility of transfer learning between pumps with the same specific speeds under almost the same measurement conditions (dataset 3) is investigated. The network is pretrained with data collected on P1 and P2, then the pretrained network is fine-tuned with the training set of the subset P3.

To obtain a more reasonable comparison, the split of the train set, validation set, and test set on subset P3 is kept the same as the standard training process. The only difference between the transfer learning process and the training without pretraining is that the initialization of the network is changed.

In the second case, transfer learning between pumps with different specific speeds and different measurement conditions is further explored. Dataset 2 is used for pretraining and the performance is evaluated on dataset 1. It is observed that for both cases, the pretrained networks outperform the network trained without pretraining and the improvement is significant in several subsets: 34.7% reduction of loss for axial acceleration in case 1 and 67.1% reduction of loss for horizontal acceleration in case 2 (Figure 6). For the relative error, it is observed that the results are significantly influenced by one or two outliers, so it does not show the expected large reduction. This may be due to the small size of the dataset, which makes the data of different rotational speeds unevenly distributed in the train set, validation set, and test set. Therefore, it is planned to build a larger dataset containing data under more rotational speeds in the following.

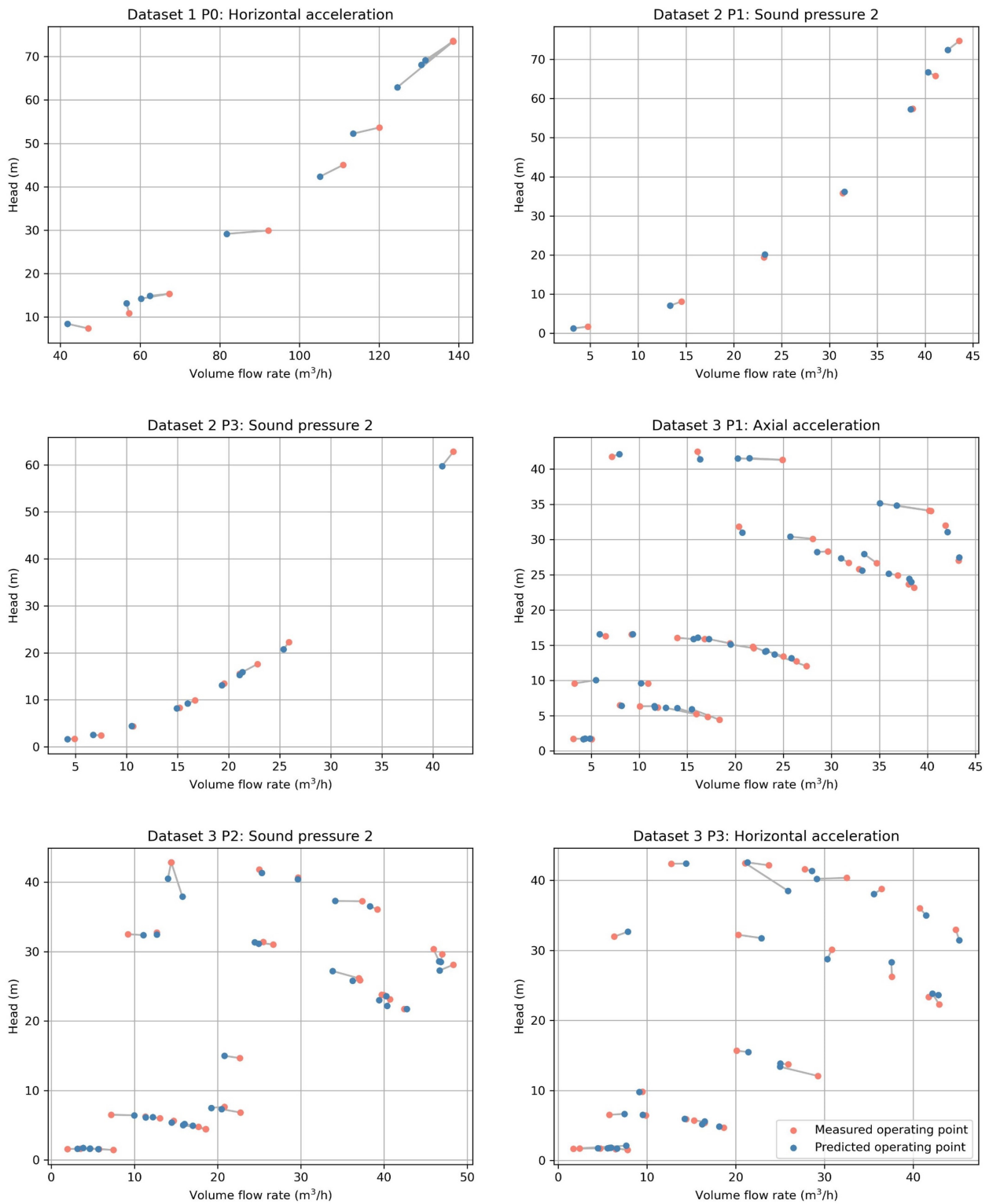


Figure 5. Prediction of operating points on three datasets. The error between the prediction (blue) and the corresponding measured operating point (red) is marked with a gray line.

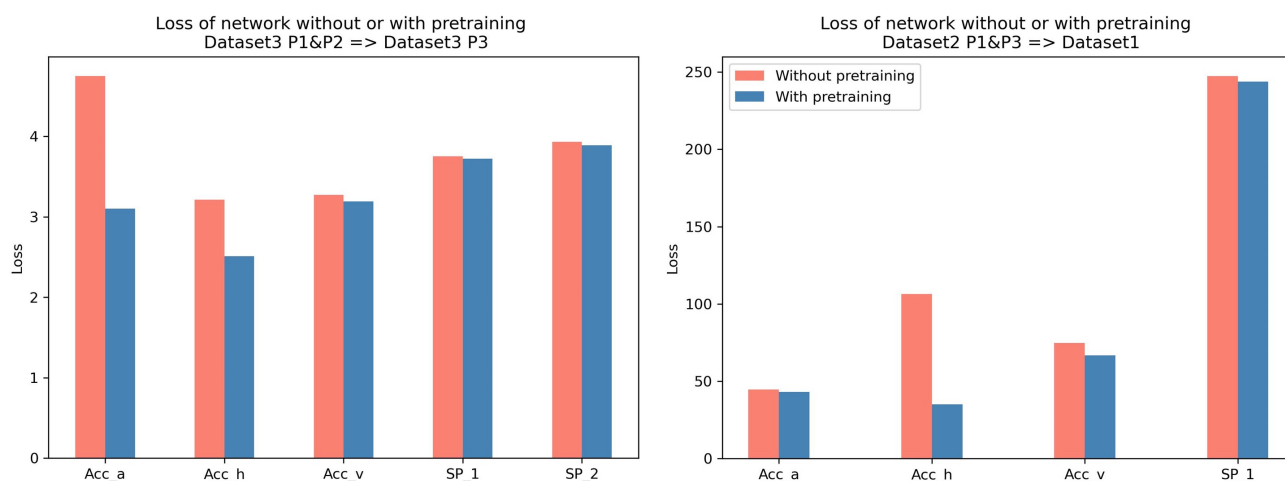


Figure 6. Comparison of the loss of network without or with pretraining. (Left) Case 1: Transfer learning in dataset 3 (Right) Case 2: Transfer learning between dataset 1 and dataset 2.

5. Conclusions and Outlook

This work offers new insights into the monitoring of pumping units operating points. Vibration and acoustic signal are used to predict the operating point of the pump. The core findings are as follows:

- The results show that the proposed method works reasonably well on different datasets. Without complex modeling, it is proven that convolutional neural networks are able to extract pump operating state features hidden in the vibration and acoustic signals.
- For the datasets measured in anechoic chamber, the best model of each subset achieves relative errors smaller than 4.9% for the prediction of head, in one case, even 3.0% only. For the prediction of the volume flow rate, the relative errors are less than 7.6%. In one case, the relative error is even 2.4% only.
- By comparing the results without and with pretraining, transfer learning achieves further reduction in loss in both cases. It reveals a promising potential in real cases when only a small amount of data is usually available. The data collected from other pumps with same specific speed (case 1) or even only under similar measurement setup (case 2) can be brought in as ‘background knowledge’ to the training of the current dataset.

For future research, to further explore performance improvements, it is meaningful to perform a systematic analysis and comparison of individual input signals or signal combinations. For example, the structure-borne sound signal with an extremely high sampling rate and the acoustic signal measured by more cost-efficient microphones is taken into consideration. It is planned to measure more data on the pump, not only for the improvement of the prediction accuracy, but also for a generalized model, which guarantees satisfying performance on different pumps. In the future project phase, the test rig will be placed outside the anechoic chamber to more closely approximate the actual industrial environment. Subsequently, data preprocessing for noisy signals and methods for data augmentation are to be investigated. Additionally, there is a long-term goal to extend this method to achieve fault diagnosis, cavitation detection and maintenance decision.

Author Contributions: Conceptualization, H.M., L.G. and S.R.; methodology, H.M. and L.G.; validation, H.M.; investigation, H.M. and L.G.; data curation, H.M.; writing—original draft preparation, H.M.; writing—review and editing, H.M., L.G. and S.R.; visualization, H.M.; supervision, S.R.; project administration, S.R.; funding acquisition, S.R. All authors have read and agreed to the published version of the manuscript.

Funding: This publication was funded by KSB Foundation (Grant number: 1.1368.2021.1) and German Research Foundation (DFG) grant “Open Access Publication Funding/2023-2024/University of Stuttgart” (512689491).

Data Availability Statement: The data presented in this study are not publicly available due to confidentiality agreements.

Acknowledgments: The authors are very grateful for the financial support provided by the KSB Foundation and German Research Foundation (DFG), to conduct this research project at the Institute of Fluid Mechanics and Hydraulic Machinery at the University of Stuttgart.

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

Nomenclature

The following symbols and abbreviations are used in this manuscript:

Δp	Static pressure difference
H	Head
Q	Volume flow rate
A_d/A_s	Pipe cross-section area on the discharge/suction side
h	Height difference
n	rotational speed
N_T	Number of samples in test set
CNN	Convolutional neural network
NN	Neural network
SMOTE	Synthetic minority oversampling technique
STFT	Short-time Fourier transform
MSE	Mean square error

References

- Chao, Q.; Gao, H.; Tao, J.; Wang, Y.; Zhou, J.; Liu, C. Adaptive decision-level fusion strategy for the fault diagnosis of axial piston pumps using multiple channels of vibration signals. *Sci. China Technol. Sci.* **2022**, *65*, 470–480. [\[CrossRef\]](#)
- Černetič, J. The use of noise and vibration signals for detecting cavitation in kinetic pumps. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2009**, *223*, 1645–1655. [\[CrossRef\]](#)
- Casoli, P.; Pastori, M.; Scolari, F.; Rundo, M. A vibration signal-based method for fault identification and classification in hydraulic axial piston pumps. *Energies* **2019**, *12*, 953. [\[CrossRef\]](#)
- Feng, J.; Böhle, M. Feasibility study for the application of a neural network for operating condition detection of a centrifugal pump. *J. Phys. Conf. Ser.* **2021**, *1909*, 012071. [\[CrossRef\]](#)
- Nguyen, M.T.; Huang, J.H. Fault detection in water pumps based on sound analysis using a deep learning technique. *Proc. Inst. Mech. Eng. Part E J. Process Mech. Eng.* **2022**, *236*, 298–307. [\[CrossRef\]](#)
- LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [\[CrossRef\]](#)
- Huang, R.; Zhang, Z.; Zhang, W.; Mou, J.; Zhou, P.; Wang, Y. Energy performance prediction of the centrifugal pumps by using a hybrid neural network. *Energy* **2020**, *213*, 119005. [\[CrossRef\]](#)
- Wu, Y.; Wu, D.; Fei, M.; Sørensen, H.; Ren, Y.; Mou, J. Application of GA-BPNN on estimating the flow rate of a centrifugal pump. *Eng. Appl. Artif. Intell.* **2023**, *119*, 105738. [\[CrossRef\]](#)
- He, D.; Li, R.; Zhang, Z.; Sun, S.; Guo, P. Gas–Liquid Two-Phase Flow Pattern Identification of a Centrifugal Pump Based on SMOTE and Artificial Neural Network. *Micromachines* **2021**, *13*, 2. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lei, Y.; Jia, F.; Lin, J.; Xing, S.; Ding, S.X. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Trans. Ind. Electron.* **2016**, *63*, 3137–3147. [\[CrossRef\]](#)
- Pan, H.; He, X.; Tang, S.; Meng, F. An improved bearing fault diagnosis method using one-dimensional CNN and LSTM. *Stroj. Vestnik/Journal Mech. Eng.* **2018**, *64*, 443–452.
- Look, A.; Kirschner, O.; Riedelbauch, S. Building Robust Classifiers with Generative Adversarial Networks for Detecting Cavitation in Hydraulic Turbines. *ICPRAM* **2018**, *2018*, 456–462.
- Chao, Q.; Tao, J.; Wei, X.; Wang, Y.; Meng, L.; Liu, C. Cavitation intensity recognition for high-speed axial piston pumps using 1-D convolutional neural networks with multi-channel inputs of vibration signals. *Alex. Eng. J.* **2020**, *59*, 4463–4473. [\[CrossRef\]](#)
- Harsch, L.; Kirschner, O.; Riedelbauch, S.; Necker, J. Estimation of Cavitation Erosion Damage with Anomaly Detection Neural Networks. In Proceedings of the 11th International Symposium on Cavitation, Daejeon, Republic of Korea, 10–13 May 2021.
- Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

16. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
17. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**. arXiv:1412.6980.
18. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.