# Detection and Level Estimation of Cavitation in Hydraulic Turbines with Convolutional Neural Networks

[1]Andreas Look*; [1]Oliver Kirschner; [1]Stefan Riedelbauch; [2]Jörg Necker

[1]*University of Stuttgart, Institute of Fluid Mechanics and Hydraulic Machinery, Stuttgart, Germany*

[2]*VOITH Hydro, Heidenheim, Germany*

**Abstract**

In this paper a method for detecting and furthermore estimating the intensity of cavitation occurrences in hydraulic turbines is presented. The method relies on analyzing high frequency signals with a convolutional neural network (CNN). The CNN is trained in an adversarial manner in order to get more robust results. After successful training the obtained network is modified in such a way, that it is possible to obtain estimations of the intensity. For evaluation purposes a separate dataset is investigated.

**Keywords**: turbines; monitoring; machine learning

## Introduction

Cavitation in hydraulic turbines can cause damage, and mostly occurs when turbines operate close to the limits of the operating range. Often it is not clearly known whether the machinery is currently operating under cavitation conditions or not. It would be advantageous to detect cavitation while the turbines are running, thus the operator can take actions in order to avoid it. Since visual detection of cavitation in a hydraulic power plant is not an option, acoustic event detection offers an alternative. Existing methods [1, 2, 3] rely on analyzing acoustics signals by designing statistical or handcrafted features, e.g. kurtosis. In this work a method is proposed, which does not rely on handcrafted features, i.e. the features will be automatically designed by the algorithm. Furthermore the method shows only little dependency on the sensor location. In the second part of this paper the method will be extended in order to estimate the cavitation intensity level.

## Data Acquisition and Preprocessing

The radiated cavitation noise is acquired by acoustic emission sensors. The sensors used in this work have an operating range from 0.1MHz up to 1MHz. Working in such a high frequency range comes with the advantage of mostly noise free signals, i.e. no inference from bearings or other mechanical parts. The signals are divided into smaller time intervals and spectrograms are generated. Since spectrograms have high dynamics, dynamic range compression (DCR) [4] is used in order to visualize acoustic events with low energy content. DCR corresponds to the elementwise application of function 1

$$f(x) = \log(1 + C * x) \tag{1}$$

where $C$ is a variable hyperparameter. A high value of $C$ will emphasize more frequencies with lower energy content, though the spectrogram becomes more noisy. After DCR the spectrogram is wrapped by a filter bank, which consists of several overlapping triangular windows. This is used for dimensional reduction in order to facilitate the training process of the CNN, while keeping the most important information. Similar preprocessing is done for obtaining Mel-frequency cepstral coefficients, which are used in speech recognition applications. In the final preprocessing step the spectrograms are normalized. In the process of this work data from several different model and prototype turbines, as well as different turbine types, was collected. Furthermore every turbine was equipped with multiple sensors at diverse positions.

## Detecting Cavitation

In order to detect cavitation a CNN is used. CNNs are typically used for image recognition tasks [5], but have also proven useful for acoustic event detection. A CNN basically consists of several learnable filters, followed by a multilayer perceptron. The learning process is done by minimizing a specified loss function via gradient descent.

*Andreas Look: andreas.look@ihs.uni-stuttgart.de

The network used in this work is inspired by the Visual Geometry Group (VGG)-architecture [6]. This means several alternating convolutional and max pool layers followed by a multilayer perceptron (MLP). Convolutional layers are able to augment the input picture in such a way, that specific shapes or structures become more visible. Max pool layers are interpretable as an infinitely strong prior, that a feature should be present, regardless of its position in the input data. Combining these two layers it is possible to create good and representable features. Based on these features the following MLP is capable to detect cavitation occurrences. For any machine learning algorithm and especially artificial neural networks (ANN) it is crucial to avoid overfitting. Any ANN is capable to obtain a training error of $0\%$ even at completely randomly labeled data [7]. In order to avoid overfitting and to obtain good generalization capability dropout [8] and early stopping is used. Dropout means randomly setting a specified fraction of neuron inputs to $0$, and thus to perturb the training process.

Using these overfitting techniques it is possible to obtain a CNN with a detection accuracy of $94{,}2\%$. The detection accuracy was evaluated on separate turbines, which have not been part of the training dataset, and multiple sensor positions. The test dataset had a $50/50$ split between measurements with and without cavitation.

In order to further increase generalization capability an auxiliary classifier generative adversarial network (AC-GAN) [9] is built. In this network topology two networks compete against each other. The first network (generator, *G*) has the task to synthesize samples, which are ideally not distinguishable from real samples. Furthermore it is possible to control the class of the output of *G* by adding an additional input. The second network (discriminator, *D*) has the task to distinguish between real and fake samples, and also to predict their class (cavitation present or not?).

$$L_S = E[\log P(S = real|X_{real})] + E[\log P(S = fake|X_{fake})] \tag{2}$$

$$L_C = E[\log P(C = c|X_{real})] + E[\log P(C = c|X_{fake})] \tag{3}$$

$$I(X|Y) = \sum_{\mathrm{i,j}} (x_{i,j} \log \frac{x_{i,j}}{y_{i,j}} - x_{i,j} + y_{i,j}) \tag{4}$$

The source loss $L_S$ corresponds to the log likelihood of the correct source (real?) and $L_C$ to the log likelihood of the correct class (cavitation?). The Kullback–Leibler divergence $I(X|Y)$ is a measurement of similarity between two different spectrograms, being commonly used in denoising tasks [10]. The generator *G* is than trained to maximize $L_C + I(X|Y) - L_S$ and *D* to maximize $L_C + L_S$.

Essentially *G* is used as a data augmentation method, in order to get more diverse training samples. After successful training of the whole network, *D* can be used as a predictor in order to detect cavitation. Using this method it is possible to further raise the accuracy to $98{,}2\%$.

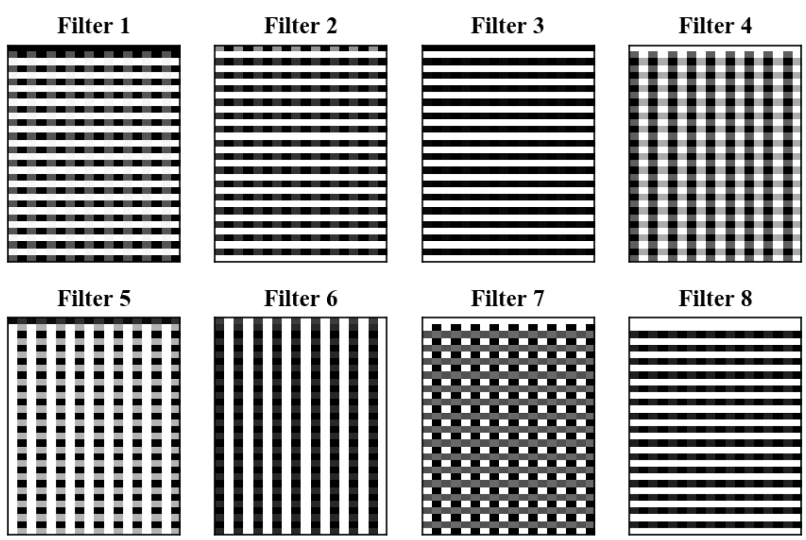


Figure 1: Maximum response inputs for convolutional filters

Figure 1 exemplarily visualizes eight learned convolutional filters in the first convolutional layer. The images have been created by modifying the input image in such a way, that it maximizes the activation of each corresponding filter. Since ANNs are fully differentiable, these images were created by the hill climbing technique. The filters maximally respond to simple geometric forms, like vertical lines, horizontal lines, or checkerboard patterns. Filter in higher layers tend to more intensively respond to inputs, with more detailed textures. Although these forms seem to be very simple, they are able to create good features, when they are stacked.

**Level estimation**

In order to estimate the intensity of the detected cavitation an additional network is built. Since there is less data available with reliable cavitation intensity estimation, transfer learning [5] is used for building a network for intensity estimation. In transfer learning, learned rules/features from one task shall be transferred for solving another, but still similar, problem. In this case the predictor, which was obtained for detecting cavitation, is very powerful, so the learned features principally can also be used for other cavitation related problems. For this application the convolutional part of $D$ is used only, the following part (MLP) is removed. Instead, a new and randomly initialized MLP is used instead of the old MLP. The new MLP is practically a clone of the old MLP, but with only one output neuron (intensity). During the training process the convolutional part is not allowed to be changed, in order to avoid overfitting.
The intensity is said to be a number $\in [0,9]$, whereas $0$ corresponds to no cavitation, and $9$ to maximum cavitation. In order to force the network to predict numbers in the given range a stretched sigmoid function is used as the activation function for the output neuron. In order to increase generalization capacity the training data is simply augmented by making the target labels (intensity) noisy. Therefore a random number $\in U(0,1)$ is added to target labels with intensity less than five and a random number $\in U(-1,0)$ is added to target labels with intensity greater than five. Furthermore aggressive dropout $(50\%)$ is used for maximum input variance.

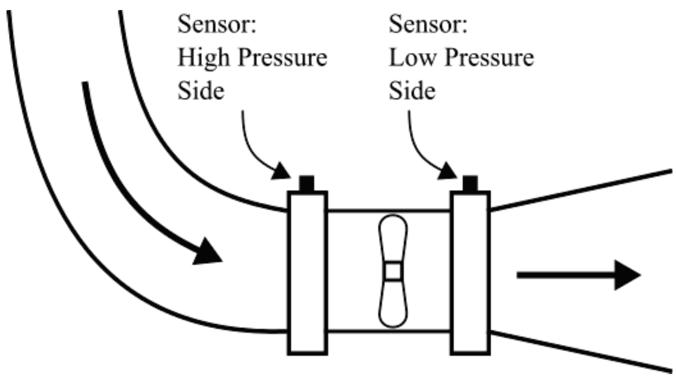


Figure 3: Principle sketch of investigated turbine

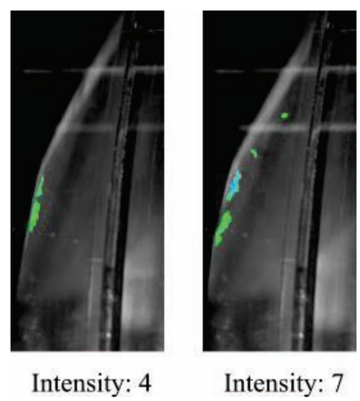


Figure 2: Intensity ranking

The algorithm for intensity estimation was evaluated on a model bulb turbine, which was not part of the training dataset. Figure 2 shows a principle sketch of the test turbine with two mounted sensors. One sensor was mounted on the high pressure side upstream the runner, the other one was mounted downstream the runner on the low pressure side. Two different measurement series have been conducted with different rotational speed. The first/second series has started with no cavitation and went on with increasing cavitation intensity on the suction/pressure side. Both series have been recorded by a high speed camera. Figure 3 shows photographs of cavitation with two different intensity levels.
Figure 4 shows the predicted and visually classified intensity. For both measurement series the RMSE is between $0.6$ and $1.2$ depending on the input channel and measurement. Although the signals have been normalized, the sensor near the cavitation occurrence tends to estimate higher cavitation intensity compared to the sensor, which is further away. There are different signal-to-noise-ratios, which cannot be easily adjusted to different positions by normalizing the signal. Nevertheless both sensor show, despite their different positions, the same tendency and predict intensity values in approximately the same range. The absolute values of $0$ and $9$ are very unlikely to be

predicted, since they have never been target labels, caused by the used data augmentation technique. Therefore the predicted intensity seems to float around 0.5 when there is no cavitation present.

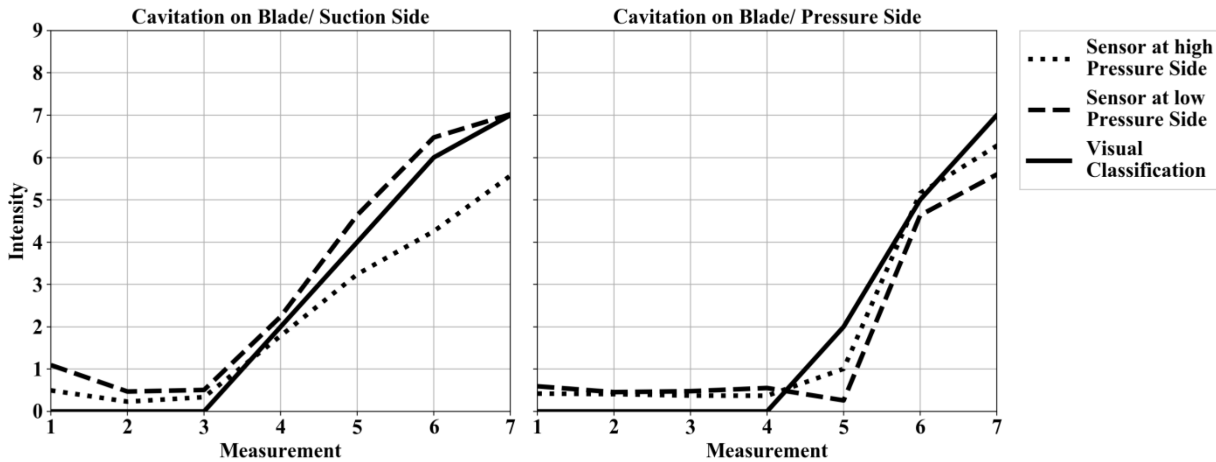


Figure 4: Estimated vs. real Intensity

## Conclusion and Outlook

In this work it was shown that convolutional neural networks, which have been trained with spectrograms, are able to detect cavitation in hydraulic turbines with high accuracy. In order to increase generalization capability of the trained network, an adversarial training method has been used. Using this method it was possible to further boost the detection accuracy beyond 98%. In the second part of this work the network, which was trained for detecting cavitation, was modified in order to estimate the intensity of the cavitation. For this purpose the old network was used as a feature generator, while a new network is used for estimating the intensity based on these features.
In further work it is planned to correlate the cavitation intensity with the erosive potential. This work will be accompanied by erosion tests.

## Acknowledgement

The research leading to the results presented in this paper is part of a highly appreciated common research project of the University of Stuttgart and Voith Hydro Holding GmbH.

## References

[1] Schmidt, H., Kirschner, O., Riedelbauch, S. (2014). *Influence of the vibro-acoustic sensor position on cavitation*. Proceedings of the 27th. Symposium on Hydraulic Machinery and Systems.
[2] Escaler, X., Egusquiza, E., Farhat, M., Avellan, F., Coussirat, M. (2006). *Detection of cavitation in hydraulic turbines*. Mechanical systems and signal processing. 20(4)
[3] Gregg, S., Steele, J., Van Bossuyt, D. (2017). *Machine learning: A tool for predicting cavitation erosion rates on turbine runners*. Hydro Review. 36(3)
[4] Lukic, Y., Vogt, C., Durr, O., Stadelmann, T. (2016). *Speaker identification and clustering using convolutional neural networks* IEEE International Conference on Acoustic, Speech and Signal Processing.
[5] Goodfellow, I., Bengio, Y., Courville, A. (2016). *Deeplearning*. MIT Press.
[6] Simonyan, K., Zisserman, A. (2015). *Very Deep Convolutional Neural Networks for Large-Scale Image Recognition*. International Conference on Learning Representations.
[7] Zhang, C., Bengio, S., Hardt, M., Recht, B., Oriol, V. (2017). *Understanding deep learning requires rethinking generalization*. International Conference on Learning Representations.
[8] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. (2014). *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*. Journal of Machine Learning Research
[9] Odena, A., Olah, C., Shlens, J. (2016). *Conditional image synthesis with auxiliary classifier GANs*. Arxiv Preprint
[10] Wilson, K., Raj, B., Smaragdis, P., Divakaran, A. (2008). *Speech denoising using nonnegative matrix factorization with priors*. IEEE International Conference on Acoustic, Speech and Signal Processing.