

Co-Design Model for Neuromorphic Technology Development in Rolling Element Bearing Condition Monitoring

Daniel Strömbergsson¹, Ashwani Kumar², Pär Marklund³ and Fredrik Sandin⁴

^{1,4} *Machine Learning Group - EISLAB, Luleå University of Technology, 97187, Luleå, Sweden*
daniel.strombergsson@ltu.se

^{2,3} *Division of Machine Elements, Luleå University of Technology, Luleå, 97187, Sweden*

ABSTRACT

This paper presents an end-to-end condition monitoring co-design model, from vibration measurement to anomaly detection, where conventional signal processing methods are combined with neuromorphic computing concepts to enable a systematic investigation of potential improvements offered by brain-like information processing technologies.

The high cost of processing digital sensor data from condition monitoring systems implies that only a minor fraction of measured data is analysed. Thereby, an untapped potential exists to improve these systems with embedded machine learning condition monitoring solutions using energy-efficient neuromorphic processors and event-triggered sensing.

The co-design model outlined here is evaluated on two use cases involving rolling element bearing failures. One use case is based on a laboratory environment dataset with known bearing condition, while the other is based on a wind turbine gearbox output shaft bearing failure representing a real-world scenario with stochastic changes in machine state and a high degree of uncertainty in the bearing condition. By adjusting the co-design model parameters, the resulting hybrid conventional/neuromorphic model has a detection performance on the laboratory dataset that is comparable to the state-of-the-art reported in the literature. Similarly, for the wind turbine dataset, a bearing defect detection time comparable to that reported in previous work is obtained.

This is a successful implementation of a hybrid conventional/neuromorphic co-design model for condition monitoring applications, which can be used and further extended to investigate performance trade-offs and efficiency improvements enabled by neuromorphic technologies.

Daniel Strömbergsson et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1. INTRODUCTION

The rapid developments of deep learning seen in the last decade have driven research into applying these techniques in condition monitoring of machine elements, e.g. gears and rolling element bearings, used throughout heavy rotating industries such as machining, marine assets, paper mills, trains and wind turbines (Liu, Yang, Zio, & Chen, 2018; Qiu et al., 2023). However, the post-Moore's law trend in machine learning since 2012 of rapidly increasing computing performance and the associated both monetary and energy costs to train deep learning models has become unsustainable (Mehonic & Kenyon, 2022). Therefore, another complementary approach to implement large machine learning and artificial intelligence models is needed, which can alleviate these drawbacks.

In information processing, moving away from conventional computing based on synchronously switching between discrete states is being driven toward at least two fundamentally different alternatives. In quantum computing, superpositions in the form of entangled quantum states are exploited in a major effort to develop more efficient computers and algorithms for particular problems, such as factorisation etc. Neuromorphic computing requires less exotic hardware and focuses on designing time-encoded signal processing to enable highly energy-efficient, highly parallelisable, and decentralised information processing systems inspired by the structure and function of biological neural networks. Industrial applications of neuromorphic sensing and computing devices are forecast to be a driving market segment within the next 15 years (Cambou & Tschudi, 2019) and large-scale adoption of neuromorphic devices within the coming five (Nguyen, Jump, & Casey, 2023). Therefore, the introduction of neuromorphic technology in machine element condition monitoring systems is an emerging and interesting research path to follow in the near future.

In neuromorphic computing, spiking neural networks (SNNs) that model analogue, neurosynaptic processes, as well as the

event-based communication (spikes) between neurons in the brain, more accurately than conventional deep learning models are used (Tavanaei, Ghodrati, Kheradpisheh, Masquelier, & Maida, 2019). The computation is event-driven, where spikes carry information and neurosynaptic spatial and temporal dynamics are used to imitate a biologically plausible computational model, which can be energy-efficiently implemented using standard semiconductor technology. Thereby, new opportunities to efficiently solve problems such as constrained optimisation, graph algorithms, kernels for composition, and signal processing have been opened by neuromorphic engineering (Aimone et al., 2022). Most relevant in this case is the application in edge computing, where neuromorphic systems can provide real-time processing on board an Internet-of-Things-based sensor. Here, the constraints on available energy for processing and transmission of data with limited latency can favourably be solved with neuromorphic systems. In the context of industrial condition monitoring, the computational demands of conventional deep learning together with increased data generation from resource-constrained edge devices, calls for new, innovative solutions that can efficiently handle high-volume and high-frequency data. Neuromorphic systems incorporating SNNs have an inherent ability to process this type of high amount of information from the environment in real-time, with abilities to reduce energy consumption. Thereby, embedded machine learning solutions for condition monitoring can be realised through their capacity for event-triggered sensing and low-power processing. This as, using mixed-signal neuromorphic processors, the cost for information processing and learning can be reduced to enable the move towards an edge device that can be placed within the machine, where the available power is limited (Nilsson, Liwicki, & Sandin, 2022). Notably, this would enable comprehensive analysis closer to the source, reducing the amount of data to transmit while also having the potential to improve defect detection and diagnosis speed (del Campo & Sandin, 2017). Thereby, the exploration and adoption of neuromorphic principles could unlock previously untapped potential in machine learning when applied in machine element condition monitoring.

In the last couple of years, a few instances of neuromorphic machine learning being used in rolling element bearing condition monitoring has been published. Dennler et al. in 2021 presented an anomaly detection pipeline using low-power neuromorphic circuits moving towards edge-component, on-line vibration condition monitoring (Dennler, Haessig, Cartiglia, & Indiveri, 2021). However, the signal decomposition was done by applying a cochlea-inspired Gammatone filter-bank to separate different frequency bands, originated in the bio-inspired scientific field from where neuromorphic technology originates. Instead, an implementation of condition monitoring-based filtering methods from the last five decades can introduce more certainty in comparing to pre-

vious condition monitoring work. Also in 2021, Zou et al. presented an SNN-based defect diagnostics tool for rolling element bearings by creating spiketrains from six statistical moments (Zuo, Zhang, Zhang, Luo, & Liu, 2021). To alleviate the non-stationary and nonlinear properties of the raw measurements, due to variations in the experimental conditions of the public datasets being used, the local mean decomposition method was used to decompose the measurements. The local mean decomposition can be an interesting alternative to use for future investigations in this area. This, together with the statistical moments yielded a comparable accuracy on the public datasets using 128x48x4 SNN. However, the highly idealised situations in the public datasets can create overly distinct differences between the different defects being classified which is fundamentally different in real-world applications where signal-to-noise levels are much lower and the ability to separate different fault sources beyond a single bearing is needed. Therefore, not distilling the information down to singular value statistical moments could be an improvement. The same authors extended their work in 2022 introducing a probabilistic spiking response model and also adding another public dataset to the analysis where the separation between fault classes is less idealised (Zuo, Xu, Zhang, Xiahou, & Liu, 2022). There, the accuracy drops to roughly 75% due to miss-classification between fault classes highlighting the need for a well-thought-out methodology for either frequency-based fault extraction or signal decomposition. Wang et al. used an SNN to improve the diagnosis of defects in aeroplane engine intershaft bearings with a newly developed spiketrain encoding scheme together with adaptations to the backpropagation algorithm enabling it to be used in the SNNs (Wang, Li, Sun, Yan, & Chen, 2022). Here, the raw vibration measurements were converted to time-frequency maps using the short-time Fourier transform, and the resulting time series is fed into an SNN classifier which can perform better than previous generation deep learning methods. While this method can achieve high classification accuracy, there exists some improvements to be investigated. First, the suitability of the short-time Fourier transform, which was not motivated beyond it being simple and requiring the least previous expertise. Secondly, the research direction being towards utilising more processing power of time series information, rather than efficiency for an edge device implementation, leaves a gap in how to set up an energy-efficient model based on previous condition monitoring methods and techniques to detect, and in the future, diagnose bearing defects.

From a neuromorphic point of view, Schuman et al. compiled the needed shift when moving towards a co-design structure to solve associated problems (Schuman et al., 2022). I.e. moving from a bottom-up approach with predefined materials and devices informing the architectures, algorithms and applications sequentially towards a co-design

approach with all design aspects influencing all other components directly. Here, another consideration can be made towards the Machine-learning Optimized Design of Experiments (MODE) collaboration target for an end-to-end optimisation scheme with a fully differentiable pipeline to allow the simultaneous optimisation of all design parameters (Dorigo et al., 2023).

This paper presents a neuromorphic co-design model for vibration monitoring of rolling element bearings, which opens up for an implementation of automatic, sophisticated optimisation scheme. Recent work introduced a neuromorphic engineering approach to rolling element bearing vibration monitoring, by adopting methodology from biological auditory research. In particular, the filter bank considered mimics the cochlea and is different from the filters typically used in condition monitoring. How does the choice of filter bank in combination with event-based sampling influence the performance of the neuromorphic approach. Therefore, the knowledge gaps and problems this paper aims to fill is to extend this previous work using methodology developed in the bearing condition monitoring field over the last decades. Thereby, an investigation has been made of condition monitoring based signal decomposition methods and an analysis of a both a laboratory bearing dataset as well as a field dataset from a wind turbine gearbox output shaft bearing failure. With these datasets, evaluations and comparisons can be made to previously published results using an SNN approach for the laboratory dataset and previous results achieved using conventional condition monitoring methods. The main contributions are:

- Systematic overview of the design parameters to be optimised

- Comparison of condition monitoring based filtering methods used for signal decomposition of vibration measurements
- Spiketrain conversion of signal decomposition channels reducing the data-rate via event-triggered sampling
- Investigation of consequences for a spiking neural network anomaly detector similar to the one used by Denler et al.

2. CO-DESIGN MODEL

A schematic compilation of its constituent component is summarised in Figure 1. One such co-design model includes a number of variables in the pipeline which has to be tuned to optimise the model, thereby the need for a co-design approach.

In this study, these are:

1. Decomposition filter types, bandwidths and limits will have an influence on the results which needs to be investigated.
2. Chosen threshold value on the amplitude change in the event-triggered sampling needs to be tuned, either to a set value or to ensure a specific spikerate.
3. SNN structure, weights as well as neuron response variables need to be tuned to the application and desired result as well.

First, the measurements from the datasets are filtered and decomposed dependent on the frequency content, where three alternatives have been investigated. Independent of the filtering method, the measurements have been decomposed into

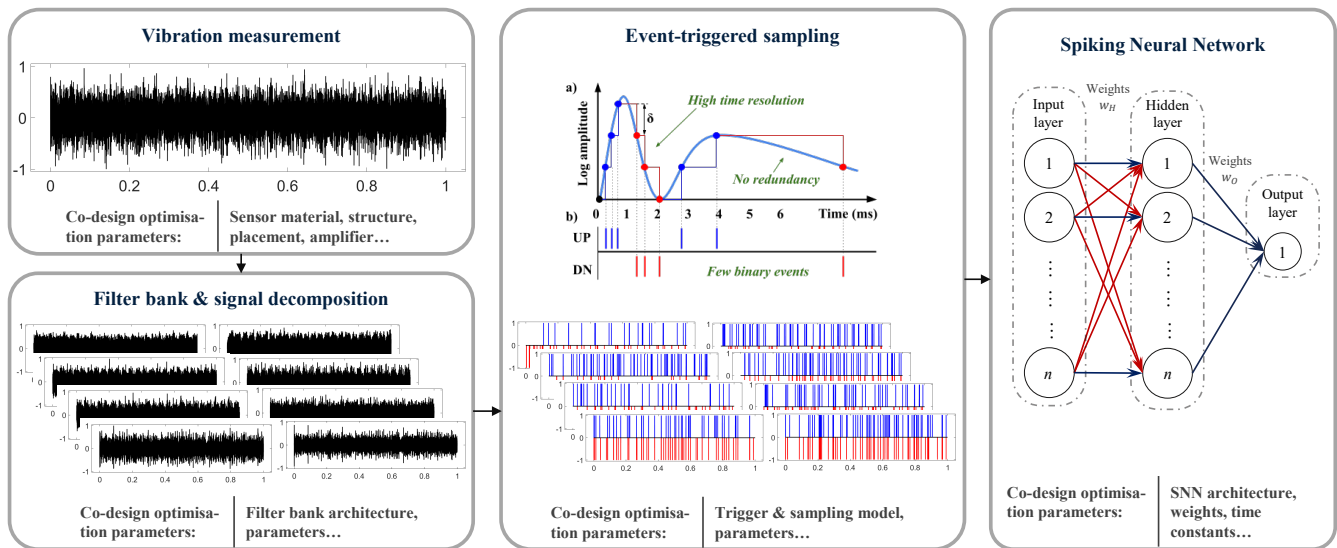


Figure 1. Co-design model with its constituent components together with the parameters in each which has to be tuned in the optimisation.

eight channels. These have then each been passed to an event-triggered sampler to extract spiketrain pairs of up and down spikes for each channel. In total, this creates 16 spiket trains which are fed into a balanced spiking neural network where a single output neuron gives an anomaly detection.

2.1. Datasets

Two different vibration measurement datasets were used for analysis, one public dataset from a laboratory bearing test rig environment which was used for validation against previous work and one dataset of wind turbine gearbox output shaft bearing run-to-failure measurements. This wind turbine dataset is unfortunately not publicly available. However, the same data has been used in previous studies, both as is in this study but also extended with more measurements before and after bearing defect inception. When comparing with individual measurements in the extended dataset, a corresponding measurement number to the shorter version has been identified and presented here.

The public, bearing test rig dataset consists of six vibration measurements in total, three healthy reference measurements and three with a defect on the outer raceway of a small-sized bearing. The measurements were taken over six seconds with a sampling rate of 97.7kHz sampling rate, a shaft speed of 1500rpm and a load of 1200N. The fault frequency of the outer ring at the set speed is 81.1Hz. State-of-the-art results using an SNN implementation with these measurements from this dataset is a perfect confusion matrix, achieved by Dennler et al. (Dennler et al., 2021). Omitted in this study were two further measurement campaigns with an outer and inner raceway defect respectively. Here, the load was varied between 0 and 1333N in 222N increments, i.e. not including the 1200N of the healthy measurements. Thereby, an uncertainty is introduced where differences comparing these two faulty measurement campaigns to the healthy reference cannot only correlate to the presence of a defect. Zuo et al. used this complete dataset in their two studies achieving an accuracy above 99% in both studies, indicating that good accuracy could be obtained without regard to some operational conditions (Zuo et al., 2021, 2022).

The wind turbine dataset consists of 219 measurements taken over a six-month period, i.e. once a day with additions due to e.g. alarms in the turbine drivetrain condition monitoring system triggering further measurements being stored beyond the regular interval. The turbine experienced an inner raceway, spalling defect in an NSK HR30234J tapered roller bearing mounted on the gearbox output shaft. The measurements are 1.28s long and with a 12.8kHz sampling rate. Since the wind speed is constantly varying, the gearbox output shaft speed span 700-1200rpm over the dataset. However, a 30rpm condition on maximum shaft speed variance over the measurement time exists in the monitoring system

to ensure that the stored measurements are not overly non-stationary and enable a more stable frequency analysis. The varying operation conditions of the turbine will induce an inherent variation in vibration measurement amplitude strength which needs to be taken into account by the monitoring methods employed. This dataset has been used in two previous studies involving this paper's corresponding author which has reported different transitions of the bearing from healthy to faulty. First, Saari et al. used an extended version of this dataset to study the performance of a Support Vector Machine on features from enveloped measurement spectra and reported possible signs of incipient inner raceway damage at measurement 51 as well as a spall appearing at measurement 159 (Saari, Strömbergsson, Lundberg, & Thomson, 2019). Strömbergsson et al. investigated the measurement properties' influence on detection and diagnosis performance using similar enveloped measurement spectra features with two moving average alarm levels set on normalised trends of the features (Strömbergsson, Marklund, Berglund, & Larsson, 2021). Here, a conservatively set alarm level, to avoid false alarms, yielded an alarm at measurement 165 while a more fine-tuned alarm level was triggered at measurement 147. Thereby, a separation of this dataset into healthy and faulty measurements cannot be done with a single cutoff point. Instead, the healthy data of this dataset was defined as the first 40 measurements while the faulty data was defined as the last 61, i.e. from measurement 159 forward.

2.2. Signal decomposition

A basis of the different signal decomposition methods investigated in this paper is the enveloping technique, which has become a conventional tool in bearing condition monitoring. A developing defect in a bearing component will emit vibrations in a repeated pattern dependent on the shaft speed and bearing dimensions as the defect repeatedly enters a rolling element/raceway contact, commonly denominated the bearing fault frequencies. However, these repetition vibrations exist at low-frequency area, often at a 10^2 Hz range, where they get drowned in structural noise and other dominant vibration sources. Instead, a developing bearing defect will incite the resonance in the bearing and surrounding components at several thousand Hz, seen in the time signal as repeated burst vibrations. These can be demodulated to separate the repetition pattern of the resonance events and thereby correlated to the specific bearing component experiencing a singular defect. This is done by:

1. Bandpass-filtering over the high-frequency resonance area
2. Rectification
3. Lowpass-filtering to highlight the repetition frequency.

The three investigated filtering and signal decomposition

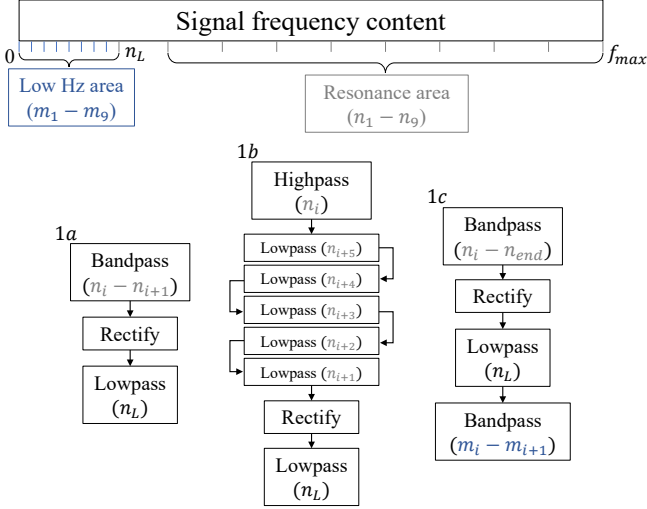


Figure 2. Flowchart of the three different filtering methods 1a-c used with symbolic indications on where in the signal frequency content they have been applied, by n_i and m_i .

methods that were used to decompose the measurements into eight, frequency-band dependent channels were:

- Envelope filters with the bandpass operation equally distributed over the whole frequency content of the measurement, derived from the Nyquist frequency
- Envelope filters with cascaded upper bandpass cutoff frequency, bandpass operation equally distributed over the whole measurement frequency content
- First, an envelope over the resonance frequency area to demodulate the early signs of the defect, then a bandpass decomposition of the lower frequency area to separate different fault frequencies into separate channels

These three different methods are summarised by symbolic flowcharts in Figure 2.

The cascaded envelope alternative has been implemented on the lowpass component in the first operation of the enveloping, i.e. the bandpass. Instead of performing a single bandpass, the input is highpass-filtered before four lowpass operations are performed four with the cutoff frequency being gradually lowered to the corresponding bandpass cutoff. This will theoretically reduce the decomposition overlap as the filter response decays more sharply at the higher cutoff points.

2.3. Event-triggered sampling & balanced SNN

To perform the event-triggered sampling of the decomposed measurements, a Level Crossing Analogue-to-Digital Converter (LC-ADC) has been implemented, shown in the event-triggered sampling box in Figure 1. Here, a set threshold on the signal amplitude change, hereafter denominated as Delta-value, is set and as soon as the amplitude has increased or decreased enough since the last recorded event, an up or down

spike is registered respectively. Thereby, the signal is represented by a limited number of binary events and inconsequential noise in-between is made redundant. Consequently, the digital processing in the following training of the SNN will in theory be computationally easier without tangible decreases in anomaly detection performance if the LC-ADC can be designed to eliminate the noise which otherwise needs to be processed. In our case when applying this to detect bearing fault frequencies, this data reduction can be done quite substantially as these fault frequencies easily becomes dominant in the filtered signals. However, to generalise the co-design model to cover other failure modes and its consequences in data reduction still needs to be investigated, as the signs of these failures might be more deeply hidden in the noisy signal. It has been shown that these types of circuits can be highly suitable for artificial intelligence engines in Internet-of-Things systems where non-uniform, sparse events need to be recorded as they are driven by the input signal instead of a clock signal (Ye et al., 2021). Drawbacks with this type of circuit exists regarding e.g. noise level of the input signal and in an energy-efficient way achieve high amplitude resolution. Thereby, these factors need to be considered and solved with additional techniques when designing the LC-ADC. However, this is deemed to be outside the scope of this paper and seen as future work when a proof-of-concept hardware implementation is built.

SNN used for the anomaly detection of the channel spiketrains consisted of Leaky-Integrate-and-Fire neurons (Gerstner, Kistler, Naud, & Paninski, 2014). These can be seen as a capacitor with a non-zero resting potential, u_{rest} , a diffusion leakage and an input current, $I(t)$, coming from the synapses, i.e. spikes stimulating the neuron. At $I(t) = 0$, the neuron membrane potential will decay towards the resting potential and when enough incoming spikes under a short enough time increases the membrane potential above a set threshold, the neuron sends out an outgoing spike and the membrane potential rapidly falls to a reset potential, often the same as resting potential. The time-dependent membrane potential, $u(t)$, of the neuron can be calculated by:

$$\frac{du}{dt} = \frac{u_{rest} - u(t) + RI(t)}{\tau_m}, \quad (1)$$

where the time constant, or refractory period

$$\tau_m = RC. \quad (2)$$

To ensure an adequately low time constant, the time between each spike in all channel spiketrains from a single healthy measurement was calculated and presented in an interspike interval histogram in Figure 3. With the interspike interval histogram in mind, the time constant was set to 0.5ms for all neurons in the SNN going forwards.

The SNN was set up in a 16x16x1 structure with the input,

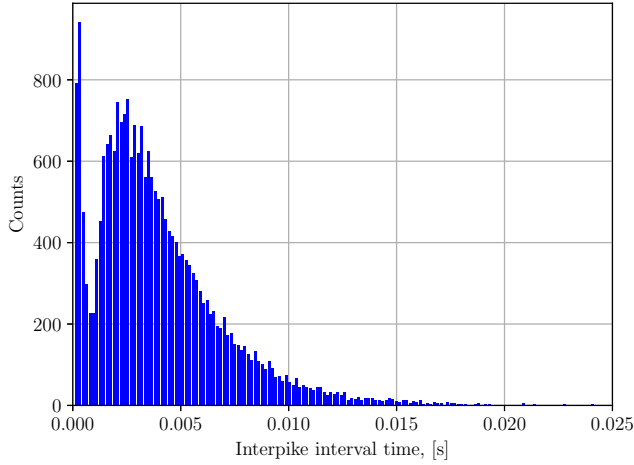


Figure 3. Interspike interval histogram for all spiketrains from a single healthy measurement from the test rig dataset.

hidden and output layers respectively. This structure can detect channel-wise changes in spikerate as this balanced structure, inspired by Efficient Balance Networks, can achieve rapid reaction times in a faulty state and sparse spiketrains through the network at a healthy state (Bourdoukan, Barrett, Deneve, & Machens, 2012). The synapse weights, used to augment the current I passed between neurons, in between the input and hidden layers, w_{Hj} , were excitatory horizontally across from neuron $n_{Ii} \xrightarrow{i=j} n_{Hj}$, while inhibitory towards the remaining 15, $n_{Ii} \xrightarrow{i \neq j} n_{Hj}$. However, the excitatory weights are scaled higher than the inhibitory by the relation $\alpha(N-1)$, where the scaling parameter $\alpha = 1.25$ and N is the number of neurons in the layer. The weights from the hidden layer neurons to the output neuron were all excitatory in nature.

2.4. Training and testing procedure

The training and testing procedure, used for both datasets individually, can be summarised as:

1. Filter and decompose each measurement with one of the following methods:
 - (a) Envelope filters with the bandpass operation equally distributed over the whole frequency content.
 - (b) Envelope filters with cascaded high bandpass cutoff frequency, bandpass operation equally distributed over the whole frequency content.
 - (c) Envelope filter over the whole resonance frequency area followed by bandpass decomposition of the lower frequency area.
2. LC-ADC on each filter channel of all training set measurements, defined as the healthy measurements for each dataset. Here, the generated spikerate was used in a feedback loop to incrementally change the amplitude Delta-

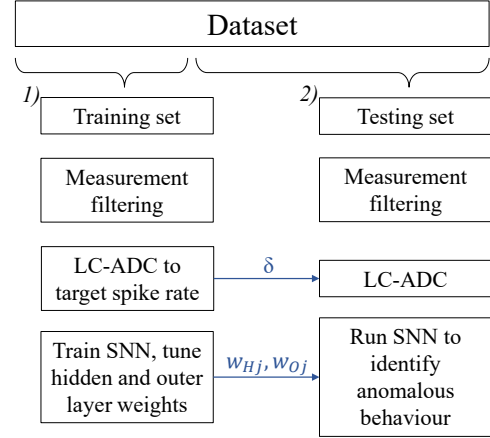


Figure 4. Procedure flowchart with indicated parameters which are being tuned in the training phase 1) and transferred to the testing phase 2).

value to achieve a 500Hz channel spikerate, i.e. each up and down spikerate pair. The target spikerate was chosen well above the fault frequencies to not act as a further lowpass filter. The final Delta-value of each channel for each measurement was stored for later use on the test set.

3. Run training set spiketrains in the SNN and tune the weights w_{Hj} and w_{Oj} until the outer layer neuron membrane potential is just below the spiking threshold, repeat for each training set measurement and store the lowest weight network.
4. LC-ADC on each filter channel of all measurements, i.e. the testing set, using the mean channel Delta-value from the training set.
5. Run testing set spiketrains in the stored SNN and define anomalies if the output layer neuron n_O spikes.

This procedure, also summarised in Figure 4, has been implemented and simulated in Python for both datasets with all three filtering decomposition methods, resulting in six experiments to run and then analyse the anomaly detection results.

3. RESULTS

3.1. Public test rig dataset results

Using the laboratory environment dataset in the procedure described in sec. 2.4, a perfect confusion matrix could be achieved when filtering and decomposing the measurements with the 1a baseline envelope and 1b cascaded envelope filters. I.e. all healthy measurements being predicted as healthy and all faulty measurements as faulty. This follows the state-of-the-art accuracy presented in sec. 2.1. To explain as to why the network spikes and indicates anomalies, the variation of the spikerate between the healthy reference and faulty measurements was summarised, i.e. the total number of gen-

Table 1. Anomaly detection for envelope baseline filter (1a), envelope cascade filter (1b) and single envelope filter followed by bandpass decomposition of low frequencies (1c), as well as total spike counts over all channels from one healthy, 1a-filtered measurement.

Measurement	Spike count	Anomaly, Yes/No
Healthy 1	23953	No: 1a-c
Healthy 2	23392	No: 1a-c
Healthy 3	23891	No: 1a-c
Faulty 1	20189	Yes: 1a-c
Faulty 2	27296	Yes: 1a-b, No: 1c
Faulty 3	34071	Yes: 1a-b, No: 1c

erated up/down spiketrains over all eight channels. These are compiled in Table 1 together with the anomaly detection results for the 1a baseline envelope and 1b cascaded envelope filters. Here, the healthy measurements have a total of roughly 24000 spikes for all spiketrains passed to the SNN, which can be derived from the 500Hz target spikerate over the six seconds measurement time for all eight channels. For the faulty measurements, the total spikerate varies from this value significantly. For the first measurement, the spikerate has decreased which has still triggered as an anomaly. Thereby, the model is not purely an anomaly detector on an increase on the number of generated spikes at a certain point in time but can also give indications for other phenomena. In the worst case, the spikerate has increased up to 42% which from an energy-efficiency perspective requires a certain level of increased processing power. However, compared to the amount of data from the >560kSamples of the raw measurements, this is still a substantial decrease of 94%, whereas the healthy measurement decrease constitutes 96%. Thereby, the cost for downstream processing can be reduced using (mixed-signal) neuromorphic processors to enable the move towards an edge device that can be placed within the machine, where the available power is limited. Looking further into the generated spiketrains of the faulty measurements gives more understanding of what is occurring in the SNN to register anomalies in all three faulty measurements. In Figure 5, a summation of the channel-wise number of spikes, i.e. one Up/Down pair, is presented for all faulty measurements together with the channel-wise mean for all healthy reference measurements, as they represent three occurrences of the same situation. Here, an interesting pattern appears where the first faulty measurement generates fewer spikes at the lower frequency bands compared to the healthy measurements while channel 4 generates more. This indicates the resonance frequency band of the test rig, which gets excited at the incipient defect and starts to repeatedly enter a roller/outer raceway contact, exists in the corresponding frequency bands of this channel. The consequence as the spikes propagate through the SNN is that the excitation of the corresponding hidden layer neurons, n_{H7-8} , to this channel increases while

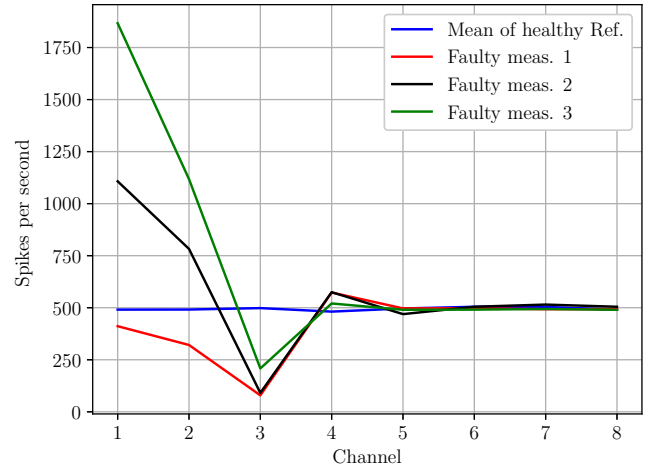


Figure 5. Channel-wise summation of spiketrains for all faulty measurements and channel-wise mean for all healthy measurements

the inhibition from the lower channels decreases, causing the output neuron to spike. For the two following faulty measurements, the spikerate of the resonance channels does not continue to increase while the lower frequency band channels increase well above the healthy measurements. This indicates towards an increasing size of the defect does not necessarily increase the amplitude and duration of the repeated bursts in the resonance frequency bands, but only affects the vibrational amplitude at comparatively lower frequency bands.

For the last filtering method, however, a miss-classification was made for two faulty measurements with a bandpass decomposing the low-frequency area of an already enveloped signal. Thereby, a simplistic bandpass to separate different fault frequencies for a future implementation where several output neurons would correspond to different bearing components experiencing a defect need a different approach in future work.

3.2. Wind Turbine case

The conclusion surrounding the 1c filtering method for the laboratory test rig dataset is found using the wind turbine rig dataset as well. Here, only a few anomalies are detected where conventional vibration monitoring techniques has been able to clearly identify the presence of a defect on the inner raceway of the failing bearing.

For the other two filtering and decomposition methods, however, anomalies were registered reliably in the time period where the certainty of a present defect is high. Figure 6 show the root-mean-squared (RMS)-values of the enveloped measurements together with the anomaly detection for the enveloped decomposition baseline in Figure 6b and the cascaded envelope decomposition in Figure 6c. Evident in this

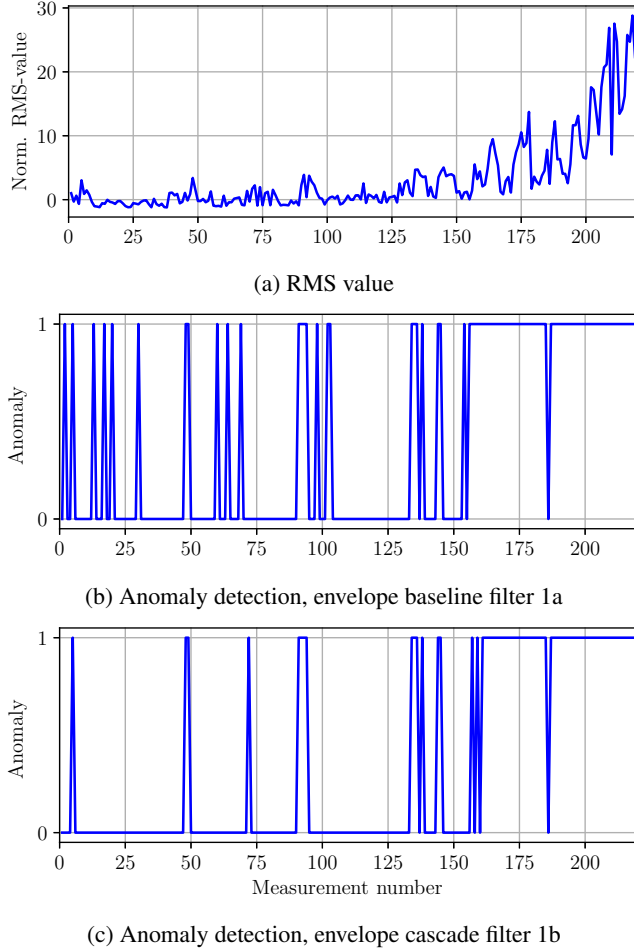


Figure 6. Anomaly detection for the Anomaly detection, envelope baseline filter 1a and envelope cascade filter 1b in (b) and (c) respectively, compared to the RMS-value of the dataset measurements after a normal envelope without decomposition.

comparison is the presence of several false negatives using the baseline envelope method, while the addition of a cascade is able to reduce this to only one such occurrence. As previously mentioned in 2.1, the separation into healthy and faulty data in this dataset was done by the first 40 measurements being designated healthy and the last 61, i.e. from measurement 159 forward, being designated as faulty. Using this separation, the true negative; TN , true positives; TP , false negative; FN , and false positives; FP , could be extracted and an F1-score calculated by:

$$F1 = \frac{2 * TP}{2 * TP + FP + FN}. \quad (3)$$

This information is compiled in Table 2 together with a summation of the designated anomalies and non-anomalies in the time period in between where a high degree of uncertainty ex-

Table 2. True/False Negatives/Positives together with the F1 score for both envelope baseline filter 1a and envelope cascade filter 1b, as well as anomaly detection in the uncertain bearing state time period.

	Filter 1a	Filter 1b
True negative	34	39
True positives	60	59
False negative	6	1
False positive	1	2
F1-score	0.945	0.975
Uncertain anomaly	22	14
Uncertain non-anomaly	96	104

ist on the presence of a defect, where more registered anomalies is found at an early stage for the 1a filtered data.

The presence of more false negatives compared to the 1b filtering method cause further doubt as to the true nature of these anomalies. Both the baseline envelope and cascaded envelope decomposition are able to show a continuous anomaly detection from measurements 156 and 161 forward respectively, which can be compared to the two previous studies results using this dataset where indications of the defect were shown reliably at measurements 159 and 165. Preceding the continuous anomaly detection, both decomposition methods show a period of interspersed anomaly detection from measurement 134 to 145 which correlate closely to the results from previous results with a fine-tuned alarm level at measurement 147 (Strömbergsson et al., 2021), which included a moving average window to be triggered causing a slight latency compared to re results in Figures 6b and 6c. Also, the first occurrences of an anomaly being designated in the uncertain time period between the defined healthy and faulty measurements occurs at measurement 48-49, closely correlating to the first indication of inner raceway damage in the Saari et al. results at measurement 51 (Saari et al., 2019).

4. CONCLUSIONS

In this study, a co-design model for anomaly detection in rolling element bearing vibration monitoring incorporating neuromorphic engineering concepts is presented. Vibration measurements have been filtered and decomposed with three alternatives based on the enveloping technique, a conventional condition monitoring signal processing method used in rolling element bearing research for the last 40 years. Thereafter, the filter output has been converted into spiketrains using Level Crossing Analogue-to-Digital Converters to transform the data into an event-driven domain. Lastly, the spiketrains were used as input to a balanced spiking neural network with a single output neuron used as an anomaly detector. The model has been employed with two different bearing failure datasets, one public dataset from a test rig in a laboratory environment, and one dataset representing a wind

turbine gearbox output shaft bearing failure. Using the described methods we achieve a perfect confusion matrix using the test rig dataset and achieve Using the described methods we achieve a perfect confusion matrix using the test rig dataset and achieve comparable performance in the wind turbine dataset to studies found in literature.

One energy-efficiency aspect of the neuromorphic system can be quantified in this case by comparing the number of spikes passed to the network for one healthy and one faulty measurement to the number of samples in the original measurement. For the laboratory test rig dataset, the six-second measurements contain more than half a million samples while the spiketrains only consist of 24- and 34kSamples, a sparsification of 96- and 94% respectively. This way the fidelity of the vibration features of a faulty bearing can be improved, and the processing can be performed on board an edge device before transmitting an eventual indication of the bearing health.

ACKNOWLEDGEMENTS

This work is funded by the Jubilee Fund and Creaternity at the Luleå University of Technology, and the Kempe Foundation, contract numbers SMK21-0046 and JCSMK JF-2303.

REFERENCES

- Aimone, J. B., Date, P., Fonseca-Guerra, G. A., Hamilton, K. E., Henke, K., Kay, B., ... Smith, J. D. (2022, sep). A review of non-cognitive applications for neuromorphic computing. *Neuromorphic Computing and Engineering*, 2(3), 032003. doi: 10.1088/2634-4386/ac889c
- Bourdoukan, R., Barrett, D., Deneve, S., & Machens, C. K. (2012). Learning optimal spike-based representations. In F. Pereira, C. Burges, L. Bottou, & K. Weinberger (Eds.), *Advances in neural information processing systems* (Vol. 25). Curran Associates, Inc.
- Cambou, P., & Tschudi, Y. (2019). *Neuromorphic sensing and computing 2019 - market & technology report*.
- del Campo, S. M., & Sandin, F. (2017). Online feature learning for condition monitoring of rotating machinery. *Engineering Applications of Artificial Intelligence*, 64, 187-196. doi: <https://doi.org/10.1016/j.engappai.2017.06.012>
- Dennler, N., Haessig, G., Cartiglia, M., & Indiveri, G. (2021). Online detection of vibration anomalies using balanced spiking neural networks. In *2021 IEEE 3rd international conference on artificial intelligence circuits and systems (aicas)* (p. 1-4). doi: 10.1109/AICAS51828.2021.9458403
- Dorigo, T., Giammanco, A., Vischia, P., Aehle, M., Bawaj, M., Boldyrev, A., ... Zaraket, H. (2023). Toward the end-to-end optimization of particle physics instruments with differentiable programming. *Reviews in Physics*, 10, 100085. doi: <https://doi.org/10.1016/j.revip.2023.100085>
- Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2014). *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press.
- Liu, R., Yang, B., Zio, E., & Chen, X. (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, 108, 33-47.
- Mehonic, A., & Kenyon, A. J. (2022). Brain-inspired computing needs a master plan. *Nature*, 604(7905), 255-260.
- Nguyen, T., Jump, A., & Casey, D. (2023). *Emerging tech impact radar: 2023*.
- Nilsson, M., Liwicki, F., & Sandin, F. (2022). Spatiotemporal pattern recognition in single mixed-signal VLSI neurons with heterogeneous dynamic synapses. In *Proceedings of the international conference on neuromorphic systems 2022*. New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3546790.3546794
- Qiu, S., Cui, X., Ping, Z., Shan, N., Li, Z., Bao, X., & Xu, X. (2023). Deep learning techniques in intelligent fault diagnosis and prognosis for industrial systems: A review. *Sensors*, 23(3). doi: 10.3390/s23031305
- Saari, J., Strömbergsson, D., Lundberg, J., & Thomson, A. (2019). Detection and identification of windmill bearing faults using a one-class support vector machine (svm). *Measurement*, 137, 287-301. doi: <https://doi.org/10.1016/j.measurement.2019.01.020>
- Schuman, C. D., Kulkarni, S. R., Parsa, M., Mitchell, J. P., Date, P., & Kay, B. (2022). Opportunities for neuromorphic computing algorithms and applications. *Nature Computational Science*, 2(1), 10-19.
- Strömbergsson, D., Marklund, P., Berglund, K., & Larsson, P.-E. (2021). Property requirements of vibration measurements in wind turbine drivetrain bearing condition monitoring. *Insight - Non-Destructive Testing and Condition Monitoring*, 63(11), 667-674. doi: [doi:10.1784/insi.2021.63.11.667](https://doi.org/10.1784/insi.2021.63.11.667)
- Tavanaei, A., Ghodrati, M., Kheradpisheh, S. R., Masquelier, T., & Maida, A. (2019). Deep learning in spiking neural networks. *Neural Networks*, 111, 47-63. doi: <https://doi.org/10.1016/j.neunet.2018.12.002>
- Wang, J., Li, T., Sun, C., Yan, R., & Chen, X. (2022). Improved spiking neural network for intershaft bearing fault diagnosis. *Journal of Manufacturing Systems*, 65, 208-219. doi: <https://doi.org/10.1016/j.jmsy.2022.09.003>
- Ye, L., Wang, Z., Liu, Y., Chen, P., Li, H., Zhang, H., ... Huang, R. (2021). The challenges and emerging technologies for low-power artificial intelligence iot systems. *IEEE Transactions on Circuits and Sys-*

- tems I: Regular Papers*, 68(12), 4821-4834. doi: 10.1109/TCSI.2021.3095622
- Zuo, L., Xu, F., Zhang, C., Xiahou, T., & Liu, Y. (2022). A multi-layer spiking neural network-based approach to bearing fault diagnosis. *Reliability Engineering & System Safety*, 225, 108561. doi: <https://doi.org/10.1016/j.ress.2022.108561>
- Zuo, L., Zhang, L., Zhang, Z.-H., Luo, X.-L., & Liu, Y. (2021). A spiking neural network-based approach to bearing fault diagnosis. *Journal of Manufacturing Systems*, 61, 714-724. doi: <https://doi.org/10.1016/j.jmsy.2020.07.003>