Chapter 5
Enhanced Bearing Fault Diagnosis Using Convolutional Neural Networks and Signal Preprocessing Techniques on the CWRU Dataset

M. Conti, G. Battista, S. Pavoni, and M. Vanali

Abstract Convolutional Neural Networks (CNN) are increasingly being used in the diagnostics of industrial machines, rotors, gears and bearings. Originally adopted for image recognition, they have been appreciated for their ability to extract important features from data. Both model-based and data-driven methods are extensively present in the literature for bearing diagnostics. In this paper, the symbiosis of the two to enhance the classification of faults is proposed. As deep learning method, LeNet-5, which is a kind of CNN architecture, is used and improved in this paper to obtain better accuracy results. Raw data provided by accelerometers has also been pre-processed with different mathematical models aimed at extracting the content linked to the fault from the signal. Envelope Analysis is the main technique employed for this target and can be paired with preprocessing models as Autoregressive Filter or Cepstrum Pre-Whitening. Another solution has been the use of a Median Filter paired with differentiation of the second order of the acceleration signal, to highlight impulses due to damage. Case Western Reserve University dataset has been taken as a reference to analyze the performance of the improved LeNet-5 network in classifying bearing faults in different conditions, such as speed and fault size, with different preprocessing. Two different fault sizes have been compared in order to evaluate the net ability to recognize the same fault at different sizes.

Keywords Convolutional Neural Networks · bearings · diagnostic · data preprocessing.

Introduction

Bearings are crucial components within rotary machines. They perform a dual function of supporting and enabling mutual rolling between the rotating parts. They consist of three objects: the inner race, the outer race and the rolling bodies, which can be balls or rollers. Each of the two races can be subjected to surface wear that manifests itself in small cracks. A third type of damage is called brinelling, which occurs when excessive loads on the bearing cause plastic deformation of the contact area between the rolling element and either the outer or inner race. When one of these faults happens, pulses are generated. These specific pulses can be detected in the vibration at particular frequencies that depend on the geometric characteristics of the bearing and the speed of the shaft on which it is mounted.

In recent years, in addition to model-based methods, data-driven machine learning and even more advanced deep learning methods have gained popularity to extract features from increasingly complex signals and map these features to bearing fault types. The employment of Neural Networks (NN) has been proven to be particularly useful in classifying the type of damage [1]. One of the most promising techniques is the Convolutional Neural Networks (CNN), which has shown strong effectiveness in this field. This type of NN is a class of deep learning based on convolutional layers, fully connected layers and a classifier [2]. Input given to CNN can be raw vibration signals or a pre-processed one.

In this paper, the use of various preprocessing strategies for improved classification of bearing faults is investigated. The techniques that, according to the literature, have shown the most promising results in clearly identifying the characteristic frequencies of bearing damage have been investigated. These techniques exploit either single processing approaches or the

M. Conti · G. Battista · S. Pavoni · M. Vanali
Università di Parma – Parco Area Delle Scienze,181/A – 43124 Parma
e-mail: marcello.vanali@unipr.it; gianmarco.battista@unipr.it; stefano.pavoni@unipr.it; marcello.vanali@unipr.it

Kristin B. Zimmerman, PhD (eds.), *Special Topics & Experimental Techniques in Structural Dynamics, 2026, Vol. 5*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.1007/978-8-743-81444-3_5

Table 1 Characteristics bearing fault frequencies for each velocity employed in the investigated dataset

	Fault Frequency [Hz]			
	1725 rpm	1750 rpm	1772 rpm	1797 rpm
BPFO	103.36	104.56	105.87	107.36
BPFI	156.13	157.94	159.93	162.19
BSF	135.91	137.48	139.21	141.16

combination of multiple methods, such as the envelope analysis [3], the Auto Regressive (AR) model [4], the Cepstrum Pre-Whitening (CPW) [5] and the Median Filter (MF) [6]. For comparison reasons, the same CNN network model is adopted, inspired by the LeNet-5 architecture with a few improvements [7]. The data employed in this work is the Case Western Reserve University (CWRU) dataset [8], which has been extensively treated in Bearing Diagnostics [9]. To better understand the improvement of the considered preprocessing techniques, LeNet5 models were initially trained using simple input data, such as the raw signal in the time domain and the power spectrum in the frequency domain. Then, the same strategy has been employed by providing pre-processed data as input, using common methods to highlight the characteristic components of bearing damage in the spectrum.

Results are compared in terms of accuracy in classifying the bearing faults. The different types of preprocessing are briefly explained in the following sections.

Experimental Data and Theory Background

When a fault occurs in a bearing, it generates a series of impulsive excitations caused by the interaction between the rolling elements and the localized defect. These excitations propagate through the structure and become evident in the vibration response. These distinctive behaviours manifest in the vibration signal at specific frequencies, which are determined by the bearing's geometry and the rotational speed of the shaft.

The characteristics frequencies associated to the bearings components are the following

$$BPFO = \frac{nf_r}{2}\left(1 - \frac{d}{D}\cos\theta\right) \tag{1}$$

$$BPFI = \frac{nf_r}{2}\left(1 + \frac{d}{D}\cos\theta\right) \tag{2}$$

$$BSF = \frac{Df_r}{2d}\left(1 - \left(\frac{d}{D}\cos\theta\right)^2\right) \tag{3}$$

where $BPFO$ is Ball Pass Frequency Outer, $BPFI$ the same for inner race and BSF Ball Spin Frequency, n is the number of rolling bodies, f_r is the shaft frequency, d is the diameter of the rolling body, D is the primitive diameter of the bearing and θ is the contact angle between race and rolling body [10].

The experimental data are taken from the CWRU, which is an open-source dataset commonly used as a reference for developing new bearing faults detection techniques. The dataset is sampled at 48 kHz with different types of bearing damage (inner race, outer race faults and brinelling) at multiple damage levels (7 mils, 14 mils and 21 mils), operating at four various speeds (about 1725, 1750, 1772 and 1797 rpm). The characteristic fault frequencies for each case are shown in Table 1.

In the following section, the data processing techniques employed in this work to enhance the bearing fault diagnosis using a LeNet5 architecture are introduced.

LeNet5 Input Data

The experimental data used in this work are taken from the CWRU bearing dataset. It consists of the measurements of two accelerometers placed on the fan end and drive end bearing of the electrical motor benchmark. The vibration signals are sampled at a frequency of 48 kHz and each recording lasts for some seconds. Since the duration of the measurements is relatively short (5/10 seconds), in order to obtain a sufficiently large dataset to reliably train the networks, a compromise must be reached between the temporal length of the signal in the time domain and the frequency range in the spectrum. Different dimensions of the input data were tested to evaluate the best possible combination, resulting in the following outcome:

- **Time domain**: Signal segments of 1024 samples, corresponding to approximately 0.02 seconds, were selected as the input data length. This duration was considered sufficient to capture in the raw signal the information related to the characteristic bearing frequencies, which in each case are greater than 50 Hz (0.02^{-1} s).
- **Frequency domain**: Power spectra were computed using 1-second signal segments, providing a frequency resolution of 1 Hz. The first 1024 elements of the spectrum (covering the frequency range 0-1023 Hz) were then employed as input data, as this range is sufficient to capture the bearing fault characteristics.

Several studies have proven that the damaged signal is not perfectly periodic but cyclo-stationary. This is due to the fact that the rolling element, upon encountering the crack, tends to slip, changing the temporality of the impulses [11]. For this reason, methods were sought in the literature to isolate the deterministic component from the signal, i.e. linked to the normal periodic operation of the machine. Below are listed and explained the procedures employed to obtain the input data of the neural networks. All the methods proposed, except for the raw signal and the simple power spectrum, aim to better identify the characteristic component of the damage.

Raw Data

Segments of the signal were extracted in the time domain by selecting 1024 consecutive samples, with the initial point of each segment chosen randomly. This approach allows for a diverse representation of the signal characteristics while ensuring that each segment contains sufficient information to capture the dynamics and features of the bearing vibrations.

Power Spectrum

The input dataset for the Power Spectrum (PS) was obtained by segmenting the signal into 1-second sections, with the initial point of each segment chosen randomly. For each section, the Fast Fourier Transform (FFT) was computed and multiplied by its complex conjugate as follows

$$S_{xx}(f) = X(f) \cdot X^*(f) \tag{4}$$

where $X(f)$ is the FFT of the segment and $X^*(f)$ its complex conjugate.

Then, the first 1024 elements of the resulting PS were selected as input data.

Simple Envelope Analysis

Envelope analysis aims to demodulate the signal to extract the components related to the fault [12]. Mathematically, demodulation is performed by evaluating the absolute value of the discrete-time analytic signal, exploiting the Hilbert transform of the signal as follows

$$x_e(t) = |x\,(t) + j\tilde{x}(t)| \tag{5}$$

where $x\,(t)$ is the raw signal and $\tilde{x}(t)$ is its Hilbert transform [13].

Then, the PS of $x_e(t)$ is evaluated as in Section 3.2.

Autoregressive (AR) Filter

The AR model, as described in [4], predicts the next value in a time series by using a series of previous observations. Given a signal in the time domain $x\,(n)$ with $n = 1,\ 2, \ldots, N$, the predicted value $x_p(n)$ at the n-th time step is calculated by the following equation

$$x_p\,(n) = -\sum_{k=1}^{p} a\,(k) x\,(n-k) \tag{6}$$

where $x_p(n)$ is a weighted sum of its p previous values in the sequence, called order of the model, $a\,(k)$ is the weighting coefficient, and $x\,(n-k)$ is the n-kth data.

The deterministic signal $x_p(n)$ could be seen also in a matrix form as $x_p = -Xa$. The matrix X is an N x p matrix based on the past series values of x. The coefficient $a\,(k)$ can be obtained by solving the Yule-Walker equation [14]

$$\begin{bmatrix} r_{xx}(0) & r_{xx}(-1) & \ldots & r_{xx}(-p+1) \\ r_{xx}(1) & r_{xx}(0) & \ldots & r_{xx}(-p+2) \\ \ldots & \ldots & \ldots & \ldots \\ r_{xx}(p-1) & r_{xx}(p-2) & \ldots & r_{xx}(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \ldots \\ a(p) \end{bmatrix} = - \begin{bmatrix} r_{xx}(1) \\ r_{xx}(2) \\ \ldots \\ r_{xx}(p) \end{bmatrix} \tag{7}$$

where $r_{xx}(k)$ is the autocorrelation function of $x(n)$, which is defined by

$$r_{xx}(k) = \frac{1}{N}\sum_{n=1}^{N} x(n)x(n-k),\ k = 0, \ldots, p-1 \tag{8}$$

The residual signal is estimated using a p order of the AR model. There are several criteria for choosing the optimal solution of order p. To select the proper p order, the Kurtosis is an excellent parameter to indicate the presence of pulses within the signal that are often associated with damage, such as in the case of bearings [15]. In this paper, an order $p = 10$ is selected because the Kurtosis of the residual signal begins to settle.

Once the deterministic signal is estimated by this process, the residual signal is given by subtracting $x_p(n)$ to $x(n)$.

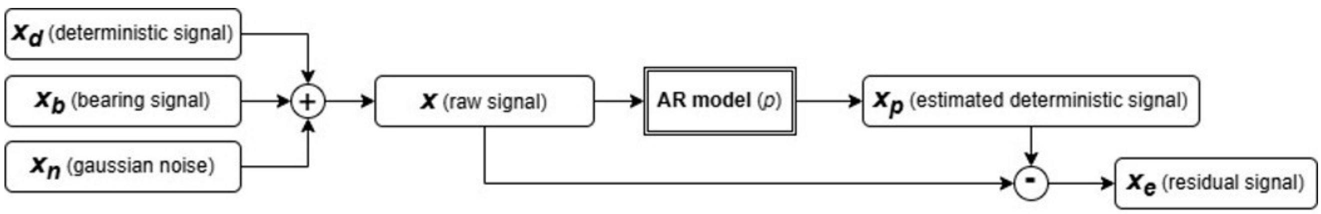

Fig. 1 Scheme of the AR model processing

The resulting residual signal was finally processed with the envelope analysis and the PS using the method presented in Section 3.2. We decided not to use spectral kurtosis to identify a bandpass filter in which the impulse content, and therefore the kurtosis, was maximum because it did not lead to better results.

Cepstrum Pre-Whitening (CPW)

In rotary machines, there are elements such as shafts, gears meshing, and motors that generate periodic but non-sinusoidal signals that are given by the sum of families of multiple harmonics of a fundamental frequency. Using the Cepstrum Transform (CT), it is possible to group these families through a single peak at a given quefrency. The CPW [5] is very easy to implement and very useful to automate bearing diagnostics since it does not require any parameters. It is based on the real cepstrum as follows

$$C(T) = \mathcal{F}^{-1}\{log[|\mathcal{F}(x(t))|]\} \tag{9}$$

where $x(t)$ is the raw signal, $\mathcal{F}$ and $\mathcal{F}^{-1}$ indicate the Fourier Transform and its inverse operation and T is the new domain of the CT that is called Quefrency. The CPW exploits the CT characteristics to separate the ciclostationary component of the signal from the periodic one, highlighting the components represented by the bearing faults. The process consists in setting a zero value for the whole real cepstrum. Then the signal is transformed back in the frequency domain and is recombined with the phase of the original signal and inverse transformed to the time domain. Making this operation the bearing signal that is not periodic but cyclo-stationary of the second order will not be filtered and will remain as a residual signal. The residual signal $x_{cpw}(t)$ obtained with CPW can be implemented by the following equation

$$x_{cpw}(t) = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}(x(t))}{|\mathcal{F}(x(t))|}\right\} \tag{10}$$

where $x(t)$ is the raw signal, $\mathcal{F}$ and $\mathcal{F}^{-1}$ indicate the Fourier transform and its inverse operation. Finally, the resulting signal $x_{cpw}(t)$ was processed with the envelope analysis and the PS using the method presented in Section 3.2.

Median Filter (MF)

The MF processing is an alternative method to the EA. A kind of envelope-squared signal is obtained in two steps. Firstly, the acceleration signal is differentiated: this process enhances the contribution of jerk forces and lifts the high-frequency content of the signal, where impacts caused by sudden jumps due to bearing faults are characterized [16]. The selection of the optimum differentiation order is performed by means of the examination of the maximum Kurtosis in the differentiated signal. In this work, the acceleration signal was processed with a second order derivative, obtaining the so-called Jounce $\left(\frac{m}{s^4}\right)$.

Once the best differentiation order is identified, the signal is squared and filtered with a median filter. That represents a better alternative to low-pass filtration and can be compared to an envelope process with robust smoothing. The median filter length (order N) is selected in order to have the first 10 harmonics of the characteristic fault frequencies [17]. Therefore, considering a signal with a sampling frequency f_s, the order of the filter N is obtained through the following equation

$$N = \frac{f_s}{10 f_c} \tag{11}$$

where f_c is the characteristic frequency of the bearing fault.

Finally, the resulting filtered signal was processed with the envelope analysis and the PS using the method presented in Section 3.2.

The functions employed for the CNN training are summarized in Figure 2.

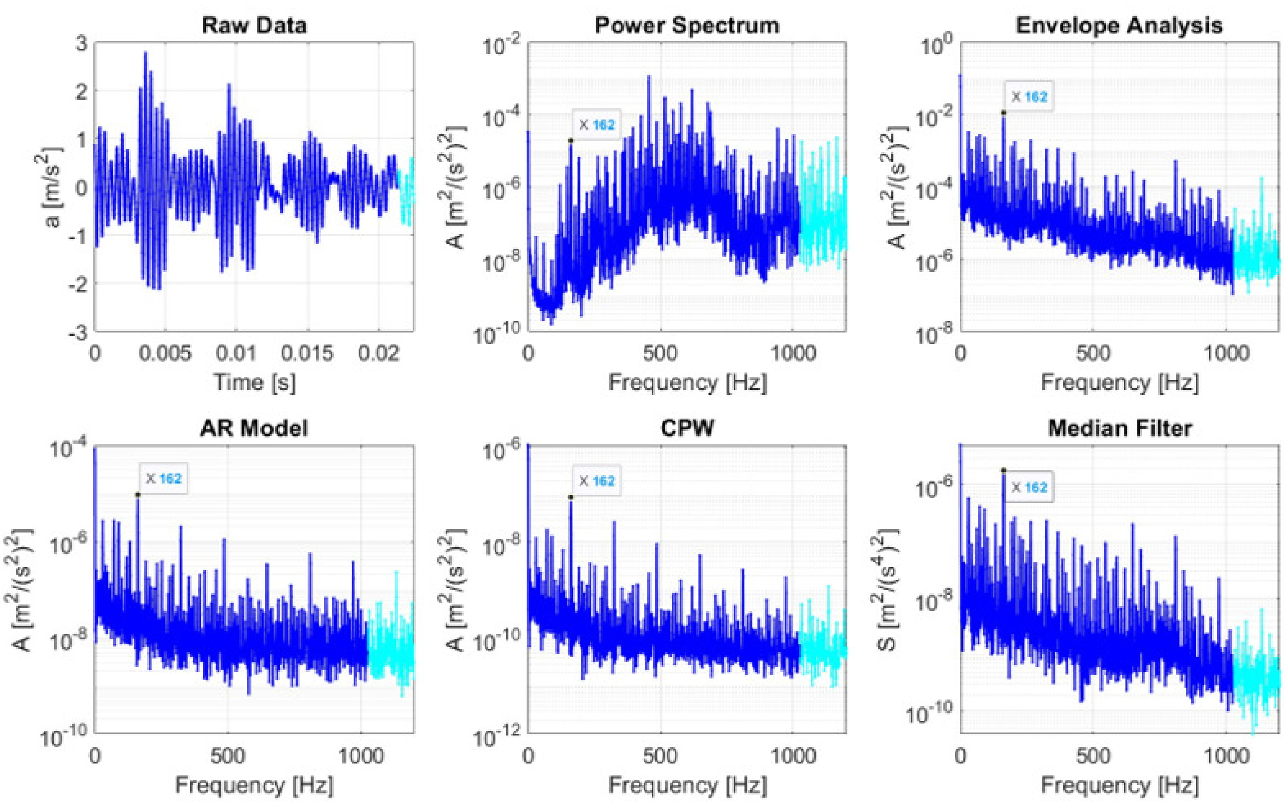

Fig. 2 Representation of the signals employed for the network's training. In blue, the 1024 samples selected as input data for each function

LeNet5 Architecture and Training

The LeNet-5 model is a classic CNN model proposed by [18]. It is widely used in several fields due to its high accuracy in the classification process. Usually, its architecture consists of an input represented by an image converted in grayscale, a series of convolutional layers followed by pooling and fully connected layers, which finally end in a classification output vector. In Figure 3 an example of a LeNet-5 model is shown.

The dimensions of the input images in 2D are obtained from the 1D signal from segments of 1024 samples transformed into greyscale images of 32x32 pixels, where each sample is normalised according to the equation

$$p_i^{'} = \frac{p_i - p_{min}}{p_{max} - p_{min}} \cdot 255 \tag{12}$$

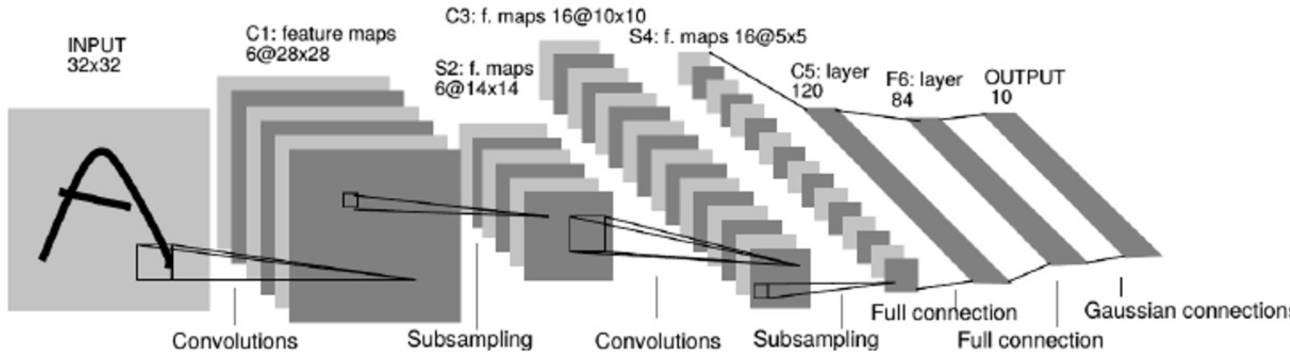

Fig. 3 Generic LeNet5 architecture

where p_i is the single value of the function to normalize and p_{min} and p_{max} are respectively its minimum and maximum. In Figure 4n, an example of the .png files employed as input data for the LeNet5.

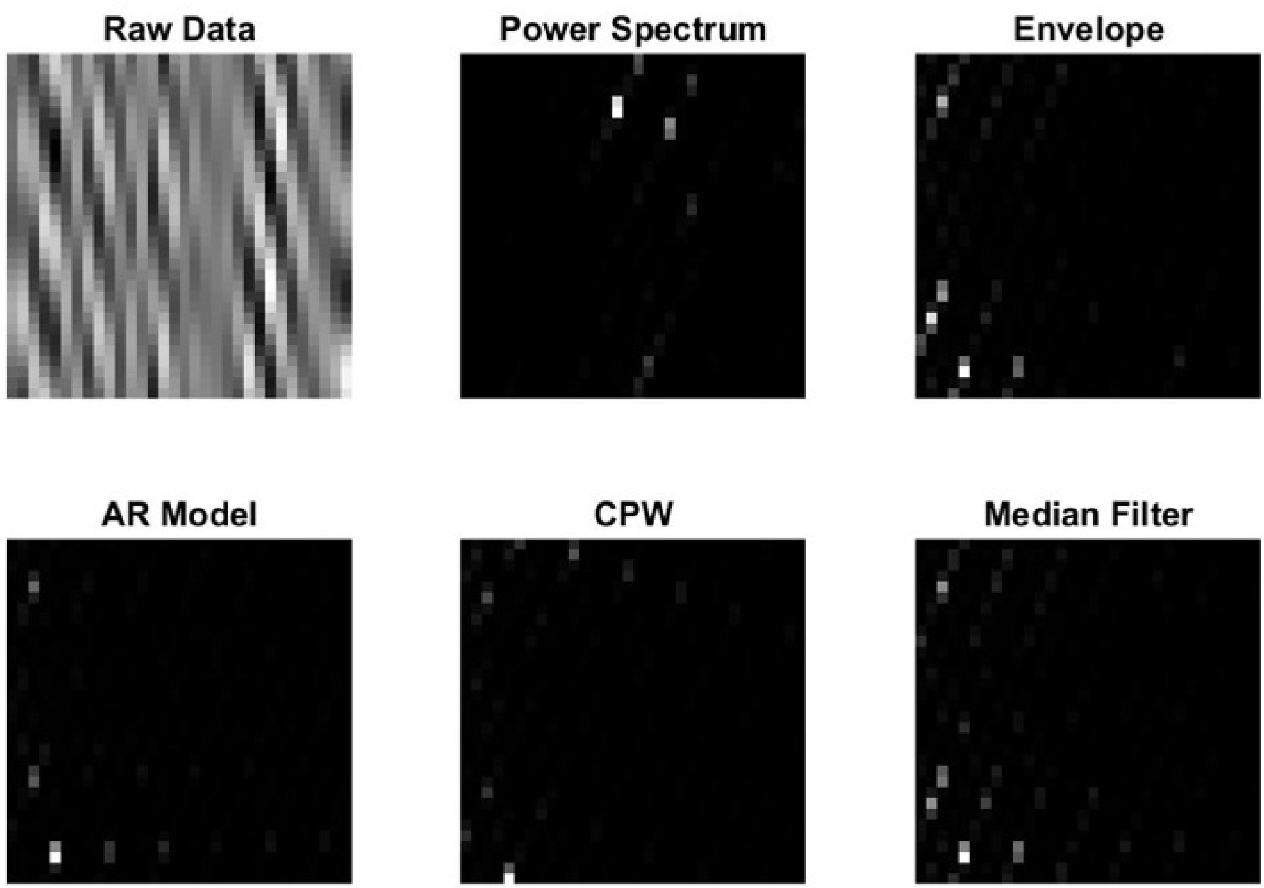

Fig. 4 Images obtained through the preprocessing techniques employed for the networks as input

An improved LeNet-5 is employed, starting from 32 x 32 images with three convolutional layers of 32 kernels 3x3 followed by a ReLu layer and a batch normalization layer to improve the stability and effectiveness of the net and three pooling layers 2x2, two fully connected layers (100 and 4 neurons), a softmax and a classification layer. The architecture of the improved LeNet-5 is shown in Figure 5.

To compare the performance of LeNet-5 models with different processing methods, the same optimization procedure was applied. The input image size was fixed at 32 x 32 pixels. The resulting 2D dataset was split into 70% for training and

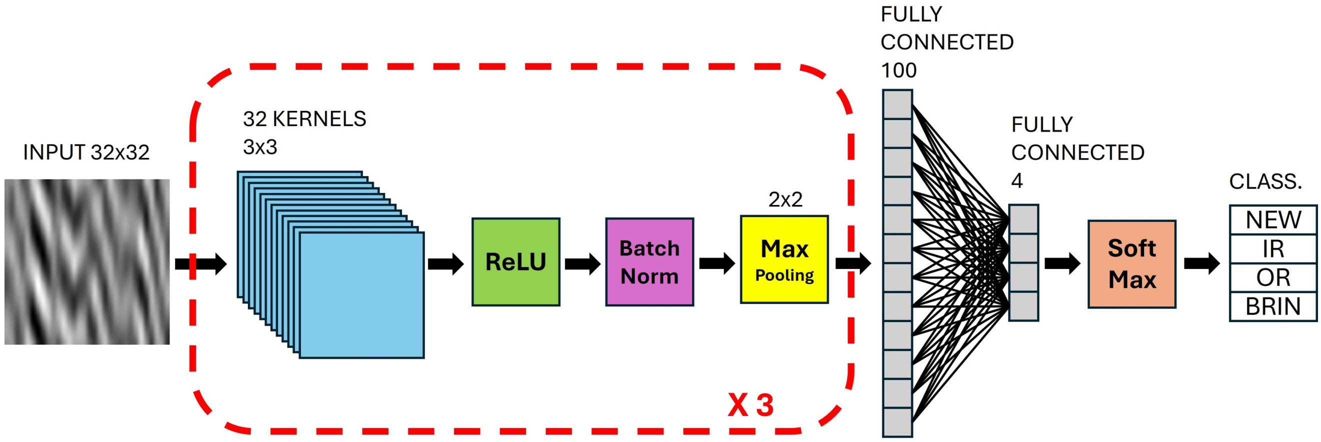

Fig. 5 Improved LeNet5 architecture

30% for testing. The training set was first fed into the LeNet5 network to train the model. Then, the testing set was used to evaluate the model's performance by analyzing the results.

The models were trained and validated using data from the medium wear state (14 mils), while the test sets included the worst damage level (21 mils). The training and validation sets consisted of 2000 observations (1400 training and 600 validation), covering all available shaft velocities. The test set was composed of 400 observations for each case (100 for each velocity). An exception was made for the healthy state and the outer race condition. In fact, the healthy condition presents only two velocities; therefore, the observations considered in the test set were only 200. Regarding the race case, damage was induced in three different positions, resulting in a total of 1200 observations for that case. The number of test observations employed for each condition and shaft velocity is reported in Table 2.

Table 2 Number of observation employed for each case to validate the improved LeNet5s accuracy

	1725 rpm	1750 rpm	1772 rpm	1797 rpm
Brinelling	100	100	100	100
Inner Race	100	100	100	100
New	0	100	100	0
Outer Race	300	300	300	300

Results are presented in the following section.

Results

The six improved LeNet5 models were trained and validated using the different input data mentioned previously:

- Raw Data
- Power Spectrum
- Simple Envelope Analysis + Power Spectrum
- AR Processing + Envelope Analysis + Power Spectrum
- Cepstrum Pre-Whitening + Envelope Analysis + Power Spectrum
- Median Filter + Power Spectrum

For each model and processing, ten different networks have been trained and tested to evaluate the overall efficiency.

Results are reported in terms of the average of the classifications among the ten networks trained for each processing strategy in Figure 6.

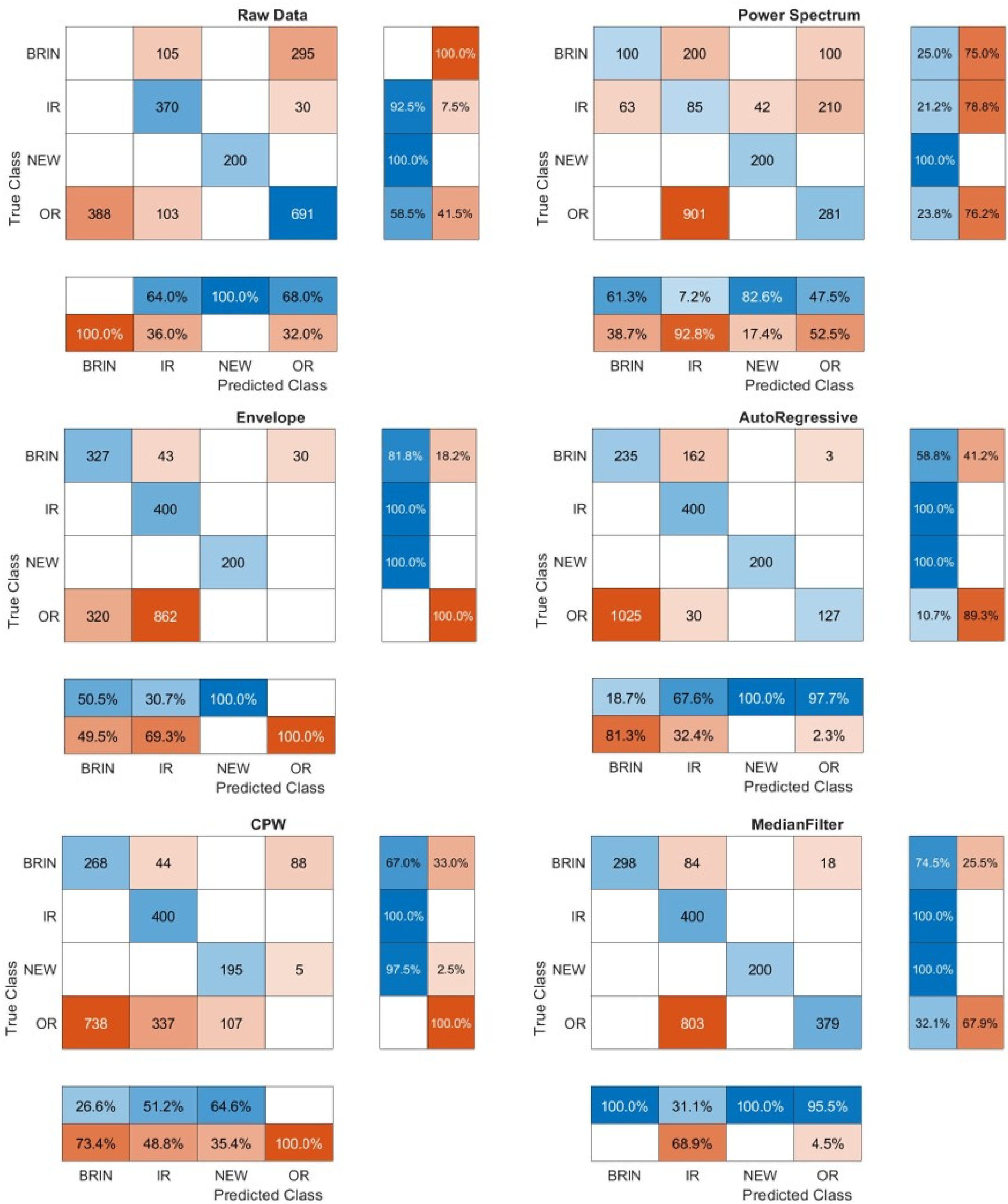

Fig. 6 Confusion matrices obtained through the validation of the trained LeNet5s for each input preprocessing strategy

Results show that all the networks trained can detect the healthy condition properly, with an accuracy of almost 100%. However, some distinctions have to be deepened:

- Raw Data training is unbalanced in the classification, where Brinelling and Outer Race are often confused.
- Power Spectrum performs poorly overall. The damage classes are heavily confused.
- Envelope is slightly better compared to the previous methods. However, the Outer Race condition is still misclassified.
- Auto Regressive improves the classification of the Outer Race.
- Cepstrum Pre-Whitening provides more balanced results. However, Outer Race is still heavily confused, even with the healthy state, which could be problematic in a diagnostic strategy.
- The Median Filter gives the best overall results. Brinelling and Inner Race are classified with high accuracy, while Outer Race achieves a relatively good outcome. Misclassifications are lower and more balanced.

Concluding, among the six preprocessing approaches, the Median Filter achieves the best performance, ensuring high accuracy across all three fault classes and minimizing confusion between damaged and healthy bearings. CPW comes second, with acceptable but less balanced results. Power Spectrum and Envelope are the weakest methods, with a strong bias toward misclassifying cases as OR.

Conclusion

Adopting the improved LeNet5 convolutional neural network makes it possible to recognize, with good accuracy, the class of damage of the bearing, also with different entities of fault. The preprocessing of the signal increases the performance of the model, as shown in Figure 6. In particular, several methods (and more traditional), such as Raw Data, Power Spectrum, envelope analysis and the AR model, tended to misclassify specific classes (notably Outer Race). While the Median Filter led to a superior discrimination, reducing the misclassification rate and enhancing overall robustness.

The significance of these outcomes lies in the method's potential applicability for early fault detection in bearings. By minimizing the incorrect labelling of damaged cases as healthy, the Median Filter approach enhances diagnostic reliability.

References

1. B. Li, M.-Y. Chow, Y. Tipsuwan, and J. C. Hung, "Neural-network-based motor rolling bearing fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 1060–1069, 2000, doi: 10.1109/41.873214.
2. V. Pandhare, J. Singh, and J. Lee, "Convolutional Neural Network Based Rolling-Element Bearing Fault Diagnosis for Naturally Occurring and Progressing Defects Using Time-Frequency Domain Features," in *2019 Prognostics and System Health Management Conference (PHM-Paris)*, IEEE, May 2019, pp. 320–326. doi: 10.1109/PHM-Paris.2019.00061.
3. S. A. McInerny and Y. Dai, "Basic vibration signal processing for bearing fault detection," *IEEE Transactions on Education*, vol. 46, no. 1, pp. 149–156, Feb. 2003, doi: 10.1109/TE.2002.808234.
4. S. Kim, D. An, and J.-H. Choi, "Diagnostics 101: A Tutorial for Fault Diagnostics of Rolling Element Bearing Using Envelope Analysis in MATLAB," *Applied Sciences*, vol. 10, no. 20, p. 7302, Oct. 2020, doi: 10.3390/app10207302.
5. P. Borghesani, P. Pennacchi, R. B. Randall, N. Sawalhi, and R. Ricci, "Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions," *Mech Syst Signal Process*, vol. 36, no. 2, pp. 370–384, Apr. 2013, doi: 10.1016/j.ymssp.2012.11.001.
6. W. A. Sethares, A. J. Milne, S. Tiedje, A. Prechtl, and J. Plamondon, "Spectral Tools for Dynamic Tonality and Audio Morphing," *Computer Music Journal*, vol. 33, no. 2, pp. 71–84, Jun. 2009, doi: 10.1162/comj.2009.33.2.71.
7. S. Li, G. Xie, W. Ji, X. Hei, and W. Chen, "Fault Diagnosis of Rolling Bearing Based on Improved LeNet-5 CNN," in *2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS)*, IEEE, Nov. 2020, pp. 117–122. doi: 10.1109/DDCLS49620.2020.9275100.
8. "https://engineering.case.edu/bearingdatacenter."
9. Y. Jin, L. Hou, and Y. Chen, "A new rotating machinery fault diagnosis method based on the Time Series Transformer," Aug. 2021, [Online]. Available: http://arxiv.org/abs/2108.12562
10. R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech Syst Signal Process*, vol. 25, no. 2, pp. 485–520, Feb. 2011, doi: 10.1016/j.ymssp.2010.07.017.
11. R. B. RANDALL, J. ANTONI, and S. CHOBSAARD, "THE RELATIONSHIP BETWEEN SPECTRAL CORRELATION AND ENVELOPE ANALYSIS IN THE DIAGNOSTICS OF BEARING FAULTS AND OTHER CYCLOSTATIONARY MACHINE SIGNALS," *Mech Syst Signal Process*, vol. 15, no. 5, pp. 945–962, Sep. 2001, doi: 10.1006/mssp.2001.1415.
12. D. Hochmann and E. Bechhoefer, "Envelope bearing analysis: theory and practice," in *2005 IEEE Aerospace Conference*, IEEE, 2005, pp. 3658–3666. doi: 10.1109/AERO.2005.1559671.

13. H. Luo, X. Fang, and B. Ertas, “Hilbert Transform and Its Engineering Applications,” *AIAA Journal*, vol. 47, no. 4, pp. 923–932, Apr. 2009, doi: 10.2514/1.37649.
14. R. B. Randall, *Vibration-based Condition Monitoring*. Wiley, 2011. doi: 10.1002/9780470977668.
15. N. Sawalhi, R. B. Randall, and H. Endo, “The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis,” *Mech Syst Signal Process*, vol. 21, no. 6, pp. 2616–2633, Aug. 2007, doi: 10.1016/j.ymssp.2006.12.002.
16. N. Sawalhi, “Rolling element bearings localized fault diagnosis using signal differencing and median filtration,” *Journal of Vibroengineering*, vol. 20, no. 3, pp. 1322–1339, May 2018, doi: 10.21595/jve.2017.18254.
17. J. B. Bednar, “Applications of median filtering to deconvolution, pulse estimation, and statistical editing of seismic data,” *GEOPHYSICS*, vol. 48, no. 12, pp. 1598–1610, Dec. 1983, doi: 10.1190/1.1441442.
18. Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.