 doi:10.20944/preprints202105.0123.v1

Article

Statistical analysis and neural network in detecting steel cord failures in conveyor belts

Dominika Olchówka [1], Aleksandra Rzeszowska [2,*], Leszek Jurdziak [1], Ryszard Błażej [1]

[1] Faculty of Geoengineering Mining and Geology, Wroclaw University of Science and Technology, Na Grobli 15, Wroclaw 50-421, Poland
[2] Faculty of Electronics, Wroclaw University of Science and Technology, Janiszewskiego 11/17, 50-372 Wroclaw, Poland
* Correspondence: aleksandra.rzeszowska@pwr.edu.pl;

Abstract: The paper presents the identification and classification of steel cord failures in the conveyor belt core based on an analysis of a two-dimensional image of magnetic field changes recorded using the Diagbelt system around scanned failures in the test belt. The obtained set of identified changes in images obtained for numerous devices parameters settings were the base for statistical analysis. It makes it possible to determine the Pearson's linear correlation coefficient between the parameters being changed and the image of the failures. In the second stage of the research, artificial intelligence methods were applied to construct a multilayer neural network (MLP) and to teach its appropriate identification of damage. In both methods were used the same data sets, which made it possible to compare methods.

Keywords: conveyor belts; magnetic method; diagnostics; NDT method; belt damage; statistical analysis; neural networks;

1. Introduction

The non-destructive testing (NDT) of conveyor belts gives vast possibilities related to optimization of the belt conveyor maintenance costs, through choosing the right moment for repair, replacement or recondition of the belt based on, among others, detected damage [1] or the rate of change of belt thickness [2]. Belt operating time depends on several factors that were presented in the article [2]. NDT research use, inter alia, analysis of the magnetic field changes generated by damaged or missing cords. Research on this method has been carried out since 1970 [3]. One of the systems in use for studies is a Diagbelt magnetic system, which enables to obtain two-dimensional images, suitable for further analysis [4,5]. The device uses magnetic field changes arising during movement of cord failures beneath the measuring probe (installed across the width of the belt), generating a discredited signal (-1, 0, or 1) corresponding to measured values of the magnetic field.

For the performed comparative analysis the reference conveyor belt containing several artificial cord failures was used. The measurements were carried by modifying system parameters: belt speed, the distance between the measuring probe and the cord and measurement sensitivity. These parameters were selected based on previous studies [6], that have confirmed their impact on received signals. The measurements were performed for many combinations of these parameters, receiving ten measuring cycles of which three were selected and used for the analysis. Preliminary, visual evaluation of the data indicates the relationship between failure detection signal and the above parameters. Some of the damage with the appropriate settings of parameter values generate similar signals. Figure 1 gives an example of this situation (damage 3 and damage 1).

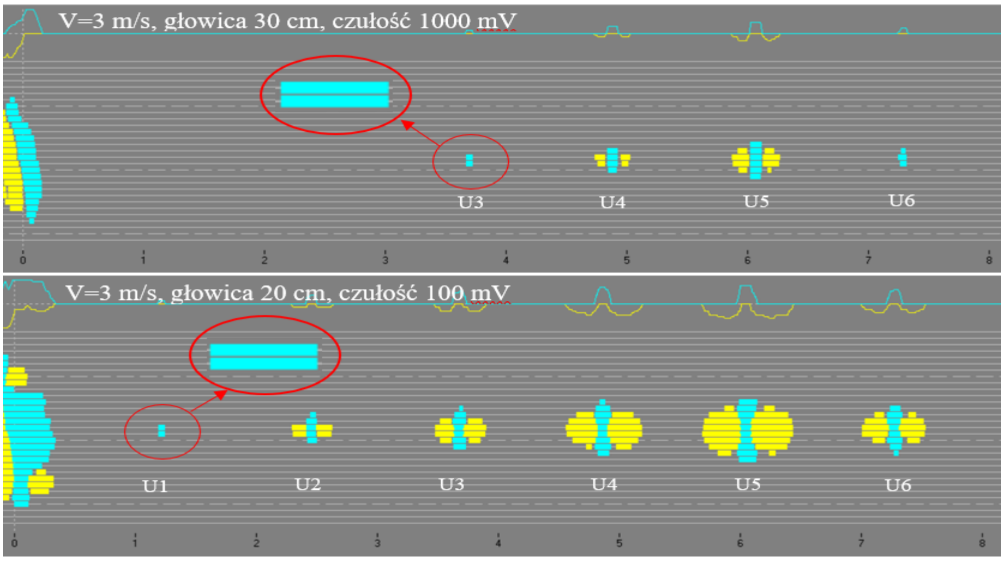


Figure 1. 2D images of the damage for significantly different settings measuring equipment parameters.

The tested failures were divided into six categories: partial core damage (20% (U1) and 50% (U2) loss one cord), complete cut of one steel cord (U3), cut off three (U4) and six (U5) cords, resection of one cord in the length of 20 mm.

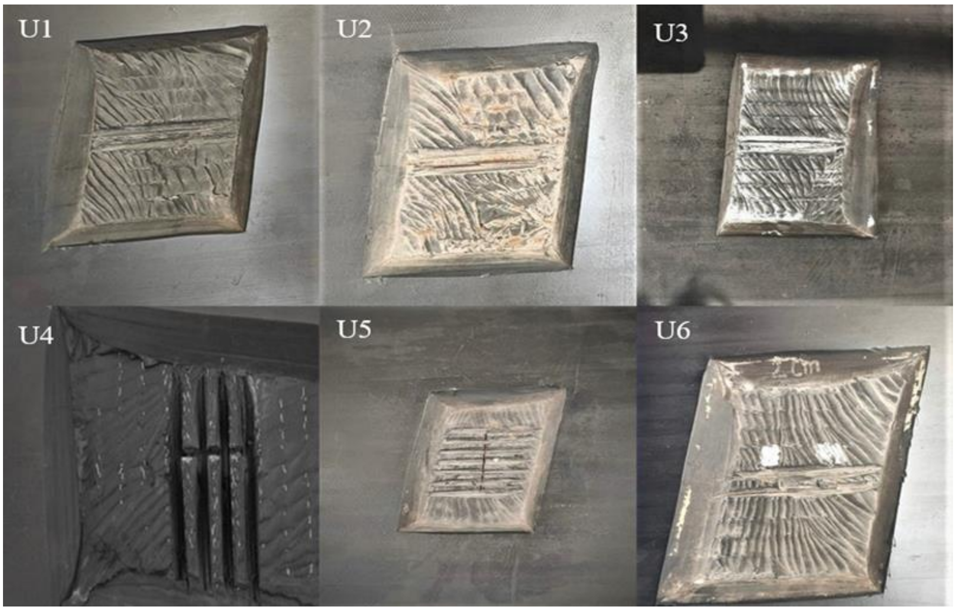


Figure 2. The actual appearance of the damage.

Each one of the cord failures described above generates the magnetic signal, and data consolidated into 12 values, 4 describing each of sub-areas: magnetic signal surface areas, number of channels on which signal be detected, width and length of the signals. The method of calculating the size of signal for exemplary damage is shown in Figure 3 and described with equations 1-3.

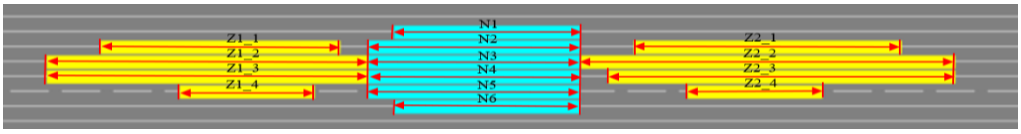


Figure 3. The method of calculating values describing selected damage.

 doi:10.20944/preprints202105.0123.v1

$$Z1sum = Z1\_1 + Z1\_2 + Z1\_3 + Z1\_4 \tag{1}$$

$$Nsum = N1 + N2 + N3 + N4 + N5 + N6 \tag{2}$$

$$Z2sum = Z2\_1 + Z2\_2 + Z2\_3 + Z2\_4 \tag{3}$$

During the measurements the belt speed (V) was increased from 2 up to 5 m/s (for every 1 m/s), the distance between the measuring probe and the cord (g) was changed within the range of 20-50 mm (for every 10 mm) and the following sensitivity levels (c) were applied: 100 mV, 150 mV, 200 mV, 250 mV, 300 mV, 400 mV, 500 mV, 600 mV, 700 mV, 1000 mV. The selection of those parameters was predicated on technical capabilities (speed of the test conveyor) and also the observation of system behaviour and its settings in numerous, previous studies [6]. The above parameter values were shown on the axis in Figure 4.

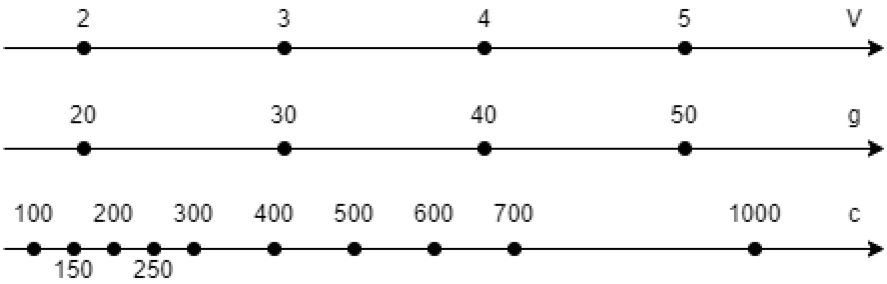


Figure 4. The distribution of measurement parameters.

The number of tested triples variants for given settings of the measuring system amounted to 160.

$$p = n_c \cdot n_v \cdot n_g = 4 \cdot 10 \cdot 4 = 160 \tag{4}$$

where: $\boldsymbol{p}$ – quantity of triples parameters variants, $\boldsymbol{n_c}$ – number of settings of sensitivity parameter, $\boldsymbol{n_v}$ – number of settings of belt speed, $\boldsymbol{n_g}$ – number of settings of distance between the measuring probe and the cord.

For each of the three measuring cycles for six, defined types of damage should be obtained 2880 records describing the damage.

$$L_p = p \cdot l_c \cdot l_k = 160 \cdot 3 \cdot 6 = 2880 \tag{5}$$

where: $\boldsymbol{L_p}$– theoretical number of records, $\boldsymbol{l_c}$ – number of measuring cycles taken into account, $\boldsymbol{l_k}$ – number of types of damage.

The actual amount of data was lower (2367) because magnetic field changes have not been detected for less core damage, for certain measurement settings.

Paper [7] defines the most appropriate measuring system parameters, which presented in Table 1. For these ranges apparatus settings (three parameters) number of output data sets decrease to the value:

$$L_p = n_c \cdot n_v \cdot n_g \cdot l_c \cdot l_k = 4 \cdot 1 \cdot (4 + 3 + 3 + 2) \cdot 3 \cdot 6 = 864 \tag{6}$$

In reality, however, number of damage was 693 (some of little defects have not been detected for specific measurement settings).

Table 1. Preferred measurement system parameter settings.

Belt speed [m/s]	Range between the belt core and the measuring probe [mm]	Sensitivity [mV]
2	20-50	200-300
3	20-50	300-400
4	20-50	400-500
5	20-50	600-700

The actual sensitivities of the measuring device are inversely proportional to parameter value called „sensitivity". When the value of this parameter is very low (e.g. 50-100 mV), the measuring system is extremely sensitive for the slightest field changes, however, the signal produced by the device is difficult to interpret. Images of the failures fuse and also appear measuring noise (Fig.5). On the other hand, while the value of this parameter is too large (the system was set to be insensitive), minor damage may not be registered, since they generate slight field changes, which are outside the scope of sensitivity of the equipment.

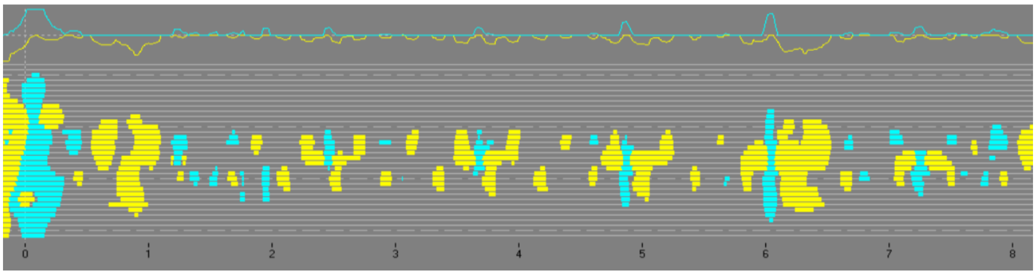


Figure 5. Inconclusive image of the damage for the measurement with high sensitivity of the measuring probe.

2. Statistical analysis

The statistical analysis was performed only for the data obtained with the optimal sets of parameters (Table 2). The statistical analysis was started with the verification of the obtained data to remove gross errors that may appear in the database, resulting, for example, from human oversight (entering incorrect data). In the next step, the correlation between the parameters taken into account in the analysis was examined. There are 13 such analysed values. They include damage number (Nr_U), number of cut cords (LL), area of damage (Pole_R), three parameters connected with the measurement system (belt speed - V, measuring probe distance - G, sensitivity - Czul), the measurement cycle taken into account (Cycle) and the failure description including two values for each of the three damage sub-areas (yellow field before damage - Z1, blue field - N and yellow field behind the damage - Z2). The measured damage parameters are the sum of the lengths of the signals recorded in the measurement channels (Z1sum, Nsum, Z2sum) and the number of channels recording the signal related to a given sub-area (Z1_LK, N_LK, Z2_LK).

Figure 6 presents charts showing the values of six measured parameters for damage depending on the class to which the damage belongs. A visual evaluation of the data helps to decide whether a given parameter affects the class differentiation or is irrelevant and can be removed.

 doi:10.20944/preprints202105.0123.v1

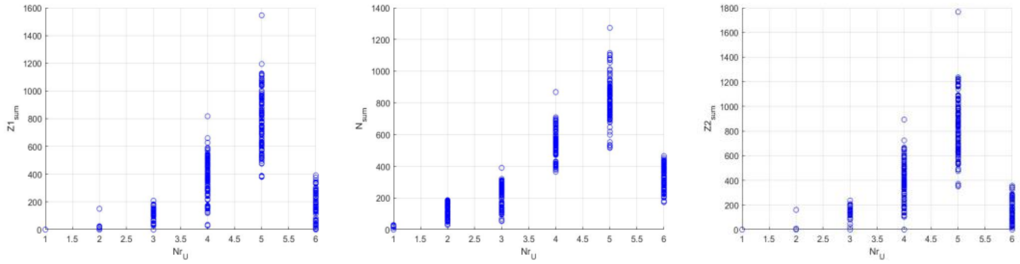


Figure 6. Distribution of the values of Z1sum, Nsum, Z2sum depending on the damage class.

The first part of the statistical analysis was determined confidence intervals for the mean for each of the analysed measurements [8]. Confidence intervals for the mean are given as a formula:

$$\bar{x} \pm \Delta = \bar{x} \pm 1.96\frac{\sigma}{\sqrt{N}} \tag{7}$$

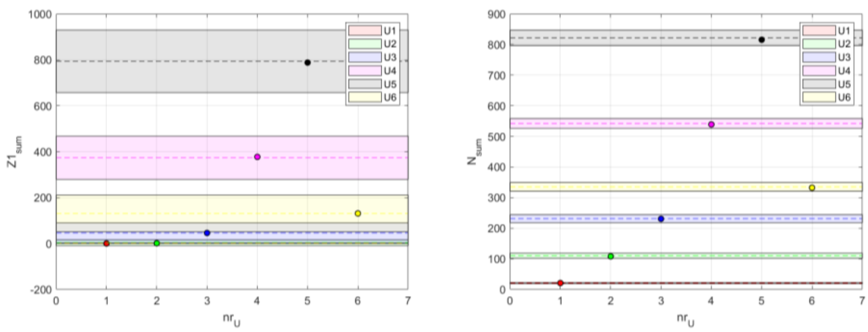
The input base is divided into two parts - training and a test set. Every third value went to the test set, while the remaining samples were left in the training set. The size of the training set is 462 samples, and the size of the test set is 231.

Table 2 summarizes the calculated values that allow to determine the confidence interval for each of the analysed data sets.

Table 2. Confidence intervals for the mean of the measured parameters.

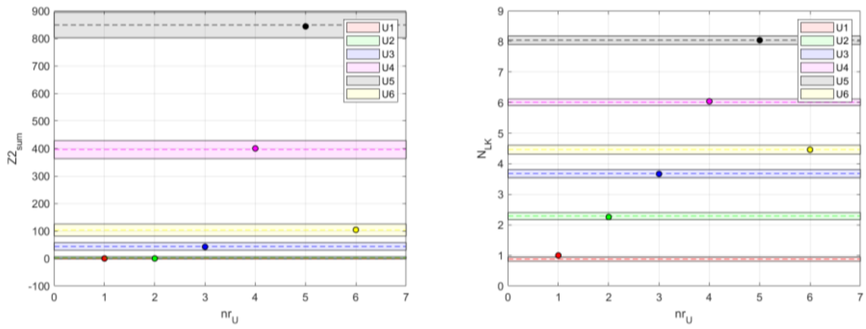
		U1	**U2**	**U3**	**U4**	**U5**	**U6**
Z1sum	$\bar{x}$	0.00	3.09	45.72	373.48	793.65	130.91
	Δ	0.00	12.67	44.37	94.09	136.12	80.08
Z1_LK	$\bar{x}$	0.00	0.09	0.79	3.40	5.17	1.61
	Δ	0.00	0.08	0.23	0.17	0.21	0.27
Nsum	$\bar{x}$	20.50	110.20	231.10	542.26	821.79	334.82
	Δ	2.07	8.29	13.12	16.09	25.25	14.44
N_LK	$\bar{x}$	0.88	2.29	3.68	6.01	8.04	4.46
	Δ	0.07	0.12	0.14	0.11	0.14	0.15
Z2sum	$\bar{x}$	0.00	2.40	43.95	395.60	848.97	103.40
	Δ	0.00	3.83	14.06	32.92	45.95	21.77
Z2_LK	$\bar{x}$	0.00	0.04	0.61	3.24	4.89	1.24
	Δ	0.00	0.05	0.18	0.19	0.18	0.23

For the data from the test, the set was determined the mean value of each of the test sets prepared in this way. These values were placed on the graph, which also marked the widths of the determined confidence intervals (Fig. 7). Table 3 summarizes the results obtained with a given test group. In the table, the values that fall outside the designated confidence interval are marked in red.



 doi:10.20944/preprints202105.0123.v1

(a) (b)



(c) (d)

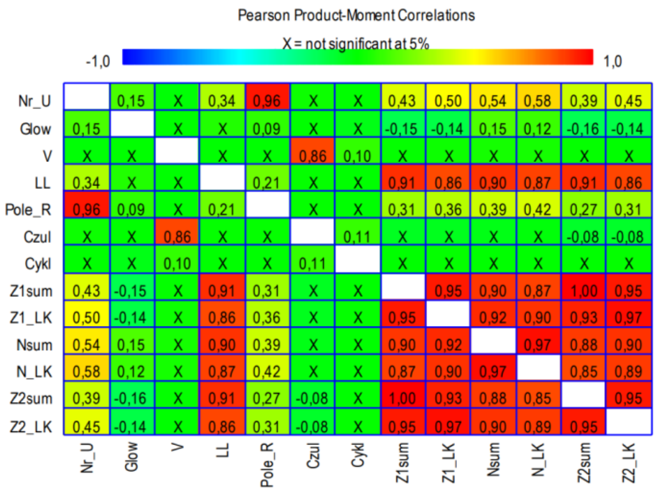
Figure 7. Confidence intervals - test set, (a) – Z1sum (overlapping intervals), (b) – Nsum, (c) – Z2sum, (d) – N_LK (data outside of assigned value).

Table 3. The mean values of the test sets.

	U1	U2	U3	U4	U5	U6
Z1sum	0.00	1.03	45.79	377.27	788.06	131.00
Z1_LK	0.00	0.06	0.75	3.40	5.17	1.63
Nsum	20.75	107.94	230.48	538.98	815.54	332.23
N_LK	1.00	2.26	3.67	6.04	8.04	4.46
Z2sum	0.00	0.00	42.33	400.04	843.58	104.23
Z2_LK	0.00	0.00	0.60	3.19	4.92	1.21

Therefore, it can be noticed that the problem with recognition appears only in the case of data concerning the first type of damage - the number of channels in the test sample was on mean 1.00, and in the training sample 0.88 ± 0.07.

To check the influence of the analysed values on each other, a statistical analysis was performed which determined the Pearson's linear correlation coefficients. Figure 8 shows a table that collects the values of these coefficients.



	Nr_U	Glow	V	LL	Pole_R	Czul	Cykl	Z1sum	Z1_LK	Nsum	N_LK	Z2sum	Z2_LK
Nr_U		0,15	X	0,34	0,96	X	X	0,43	0,50	0,54	0,58	0,39	0,45
Glow	0,15		X	X	0,09	X	X	-0,15	-0,14	0,15	0,12	-0,16	-0,14
V	X	X		X	X	0,86	0,10	X	X	X	X	X	X
LL	0,34	X	X		0,21	X	X	0,91	0,86	0,90	0,87	0,91	0,86
Pole_R	0,96	0,09	X	0,21		X	X	0,31	0,36	0,39	0,42	0,27	0,31
Czul	X	X	0,86	X	X		0,11	X	X	X	X	-0,08	-0,08
Cykl	X	X	0,10	X	X	0,11		X	X	X	X	X	X
Z1sum	0,43	-0,15	X	0,91	0,31	X	X		0,95	0,90	0,87	1,00	0,95
Z1_LK	0,50	-0,14	X	0,86	0,36	X	X	0,95		0,92	0,90	0,93	0,97
Nsum	0,54	0,15	X	0,90	0,39	X	X	0,90	0,92		0,97	0,88	0,90
N_LK	0,58	0,12	X	0,87	0,42	X	X	0,87	0,90	0,97		0,85	0,89
Z2sum	0,39	-0,16	X	0,91	0,27	-0,08	X	1,00	0,93	0,88	0,85		0,95
Z2_LK	0,45	-0,14	X	0,86	0,31	-0,08	X	0,95	0,97	0,90	0,89	0,95	

Figure 8. Pearson's linear correlation coefficients.

The data in the table above are displayed using both colours and numerical values. Data marked with "X" are statistically insignificant. Coefficients define a linear relationship between two different variables - the greater the value of the correlation parameter, the greater the degree of interconnection of the pairs of variables. It is worth noting that the presented table shows that the selected parameters of the measuring system do not significantly affect the type of damage - no correlation was found between the belt speed and the measurement results (statistically insignificant correlation), there is a low correlation between the measuring head distance and the measurement results (negative or positive, depending on the area, within the range [-0.16,0.15]) and no significant correlation or weak negative correlation (-0.08) of the measurement results with the sensitivity of the device.

It is also worth noting that all measurement results are strongly, positively correlated with each other, and the correlation between the values describing the yellow fields (Z1sum and Z2sum and Z1_LK and Z2_LK) is 1.00 and 0.97, which allows maintaining the hypothesis about their symmetry [1].

Analogical statistical analysis for full data set was described in detail in [9], but the results obtained there turned out to be less satisfactory than the results obtained for specific parameters of the measurement system. The distribution of the data from the complete set is shown in Figure 9 (these graphs show the values of Z1sum, Nsum and N_LK). These results largely overlap and it is impossible to clearly define the boundaries of clusters [10, 11].

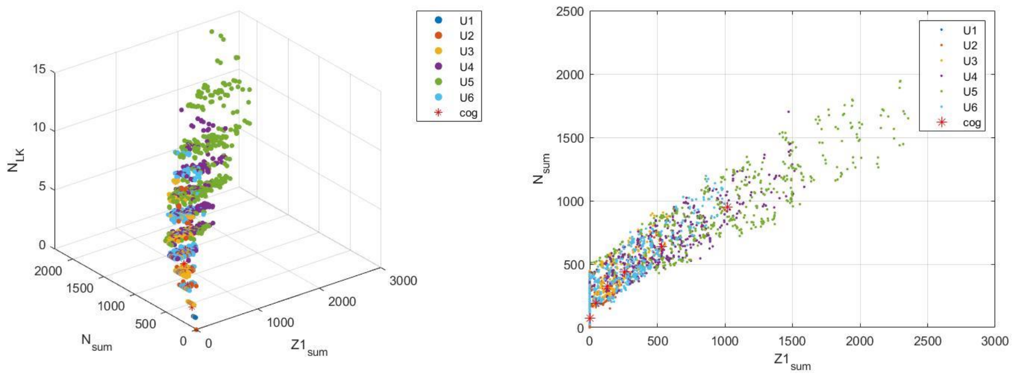


Figure 9. Cluster analysis - full data set, dependence Z1sum, Nsum, N_LK.

Similar graphs were also plotted for the set of parameters of the measurement system tested in this analysis. It can be noticed that in this case, it is possible to limit the obtained data with a certain curve marking the boundary of a given cluster, although there are still areas where the belonging of the measurement result to a given cluster is ambiguous.

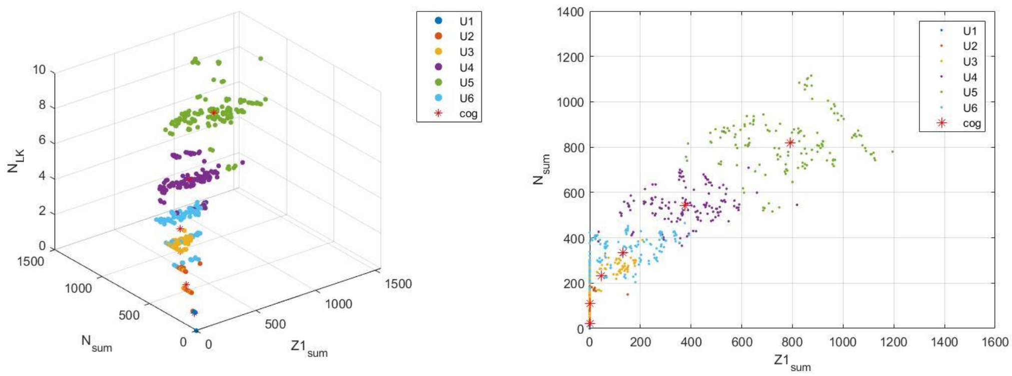


Figure 10. Cluster analysis - limited data set, dependence Z1sum, Nsum, N_LK.

Optical analysis of the obtained charts shows that the measurement data is highly probable to classify the type of damage, but there are areas where the classification may fail because clusters overlap [11]. Due to this fact, another analysis was carried out using artificial neural networks.

3. Analysis with the use of neural networks

The analysis of the selection of the structure and parameters of the neural network as well as the idea of its operation has been widely described in the literature. The studies [12-15] describe in detail the rationale behind the selection of specific parameters used in this research. The MATLAB software with the Deep Learning Toolbox installed was used for the learning process of neural networks. The way of using this toolbox is described in studies [16,17].

To use artificial neural networks to classify the conveyor belt damage, it was necessary to generate appropriate sets of training and test sets, train the network on the training set, and then test it on the test set. Since neural networks can divide the classification space non-linearly, in the course of this research, two variants of the selection and division of the input data were distinguished. In each case, the vector of input data consists of 15 elements (and this is the number of neurons in the input layer of the neural network): 3 measurement parameters and 4 values describing each of the three sub-areas. There are 6 neurons in the output layer - one each responsible for belonging to a given class of damage. Between the input and output layers were placed two hidden layers consisting of 31 and 63 neurons - the size of these layers was determined by the Kolmogorov theorem, according to which the number of neurons in the hidden layer should be equal to the number of neurons in the previous layer multiplied by 2 and increased by one [18]. The diagram of the neural network used in this study is shown in Figure 11.

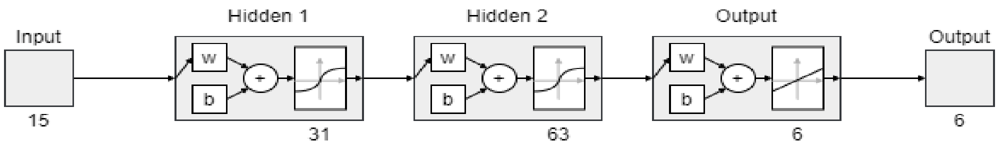


Figure 11. Diagram of the neural network used in the research.

Two variants were used for the process of learning neural networks:
- variant 1: analysis of all available data, 2 measurement cycles of each parameter set go to the training set, one cycle to the test set. The number of training and test sets is 1578 and 789.;
- variant 2: analysis of data obtained in the measurements with the best possible sets of parameters (Table 1), two measurement cycles of each of the included parameter sets go to the training set, one cycle to the test set. The number of training and test sets is 462 and 231.

Each of the network learning processes in a given variant was performed three times, and the results are presented in Table 4.

Table 4. Measurement results of neural networks analysis.

variant	no	recognition effectiveness [%]						
		U1	U2	U3	U4	U5	U6	total
	1	98.57	99.19	99.67	99.69	100.00	100.00	99.37
1	2	97.14	97.56	98.68	98.13	100.00	100.00	98.86
	3	100.00	98.37	98.03	98.13	100.00	100.00	98.99
	1	100.00	97.14	97.92	100.00	100.00	100.00	99.13
2	2	100.00	97.14	97.92	100.00	100.00	100.00	99.13
	3	100.00	100.00	100.00	100.00	100.00	100.00	100.00

4. Conclusion

Many scans performed using the Diagbelt system have shown that the magnetic measuring system is well suited to obtain detailed information about the technical condition of the belt core. The idea of the system is based on the measurement of magnetic field changes at sites of core damage. The data obtained are presented on two-dimensional pictures that can be easily analysed using proposed known methodologies.

One such method is verification based on the average of the samples. For this purpose from the set of training data, the mean value and 95% confidence interval were calculated for each damage. Then, for the test set containing the specific damage data obtained at different set values (not taken for training), the mean value was determined, which allowed for its comparison with the previously determined confidence intervals. In nearly every instance, testing set data have been included in the relevant confidence interval. The number of samples in training and testing data for selected parameter sets were 8 and 4, which could be a too-small value. It is however worth noting that with such a choice of analysis for automatic recognition of damage, it is necessary to the execution of many damage measurements for multiple sets of parameters, to obtain a sample that comes from the same distribution as the training data. This solution can be cumbersome, and as the study shows, verification by mean of the sample data is not always reliable.

An analysis of the Pearson correlation coefficient allows to assess interdependence of evaluated parameters and therefore initially verify, which parameters are worth analysed with the classification of measuring damage, and which are redundant and have no correlation with the type of damage. Such an analysis does not allow for the designated similar data based on a new sample, but on its basis, it is possible to construct a statistical model necessary for the assessment of future data. Creating such a model which is a response to the analysis presented in this article, is a good direction for future research.

When the full set of data will be limited to data obtained for the most appropriate system settings are achieved better statistical analysis values. The full data set shown in the two-dimensional and three-dimensional plots (Fig. 8-9) indicate that these data cannot be isolated from each other and locked in separate clusters, however, for limited data cluster analysis is possible, as areas of interdepending neighbouring clusters are slight.

The analysis based on neural networks allows the omission of the problem of nonlinearity. The network containing two hidden layers allows to solve almost every problem of classification, provided it has the appropriate input data. In the framework of this elaboration was carried out analysis using neural networks, both on the complete dataset and limited dataset containing results obtained for best possible system settings. The data collected in both of these variants have shown good efficacy (above 98%), following the implementation of the testing process.

The network does not have a problem with classification of last two damage (U5 and U6), however, it makes errors recognising defects 2-4. This may be due to real size of the defect concerned. The neural network analysis compared to statistical analysis, allows for quick action of the entire system while maintaining high efficiency.

It is worth noting that while analysis of statistical method has already been used to classification of belts damage, cluster analysis and analysis using the neural networks have so far been rarely discussed and its results presented. Developers of the diagnostic systems, in many cases, prefer to remain ability to the interpretation of measurement results keep it to yourself, that their services will not unnecessary. In the Industry 4.0 era [19] the automatic interpretation of the diagnostic signal is a necessary to cope with data processing for an ever-increasing amount of data. Test results discussed in this paper are promising and they show the direction of further action authors as part of a research project „Integrated mobile system of automatic testing and continuous diagnostics of the condition of conveyor belts" (project number: POIR.01.01.01-00-1194/19) [20].

Author Contributions: Conceptualization, R.B. and L.J.; methodology, D.O., A.R. and L.J.; formal analysis, A.R. and L.J.; resources, D.O.; data curation, D.O, A.R.; writing—original draft preparation,

D.O. and A.R.; writing—review and editing, R.B and L.J.; supervision, R.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Jurdziak L., Błażej R., Burduk A., Bajda M., Kirjanów-Błażej A., Kozłowski T., Olchówka D., Optimization of the operating costs of conveyor beltsin a mine in various strategies of their replacement and diagnostics, Transport przemysłowy i maszyny robocze, 4(50)/2020, 14-22.
2. Webb C., Sikorska J., Khan R. N., Hodkiewicz M., Developing and evaluating predictive conveyor belt wear models. Data-Centric Engineering (2020), 1:e3.
3. Barry, B., 2007. Multi-Channel Conveyor Belt Condition Monitoring, US7275637B2, United States.
4. Błażej, R., Jurdziak, L., Kirjanów, A., and Kozłowski, T. (2015). Evaluation of the quality of steel cord belt splices based on belt condition examination using magnetic techniques. Diagnostyka, 16(3), pp.59-64.
5. Błażej, R., Jurdziak, L., Kozłowski, T., Kirjanów A. (2018). The use of magnetic sensors in monitoring the condition of the core in steel cord conveyor belts – Tests of the measuring probe and the design of the DiagBelt system. Measurement, 123 (2018), 48-53.
6. Kozłowski T., Błażej R., 2018, , Influence of selected parameters on the magnetic testing results in a DiagBelt system, Transport Przemysłowy i Maszyny Robocze, 4/2018, 6-9;
7. Olchówka D., 2020, Selection of measurement parameters using the DiagBelt magnetic system on the test conveyor, Engineer's Thesis, Wrocław 2020.
8. Hazra A., Using the confidence interval confidently. Journal of Thoracic Disease. 9. 4124-4129. 10.21037/jtd.2017.09.14, 2017.
9. Olchówka D., Jurdziak L., Błażej R., Kozłowski T., Examination of the steel cord belts core damage condition using the DiagBelt system Part 1. Determining the number of steel cord cuts in damage., Transport Przemysłowy i Maszyny Robocze, 1(51)/2021, 2-10.
10. Omran M., Engelbrecht A., Salman A., An overview of clustering methods. Intell. Data Anal.. 11. 583-605. 10.3233/IDA-2007-11602, 2007.
11. Saxena A., Prasad M., Gupta A., Bharill N., Patel op, Tiwari A., Er M., Lin C., A Review of Clustering Techniques and Developments. Neurocomputing. 267. 10.1016/j.neucom.2017.06.053, 2017.
12. Hornik K., Stinchcombe M., White H., Multilayer feedforward networks are universal approximators. Neural Networks, vol. 2, 1989.
13. Kirjanów-Błażej A., Rzeszowska A., 2021, Conveyor Belt Damage Detection with the Use of a Two-Layer Neural Network, Procedia Manufacturing
14. Osowski S., Neural networks for information processing, Warszawa: Oficyna Wydawnicza Politechniki Warszawskiej, 2020 (in Polish).
15. Phil Kim, MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence, Seoul, Soul-t'ukpyolsi, Korea, 2017.
16. Anghel D., Ene A., Belu N., A Matlab Neural Network Application for the Study of Working Conditions, 2013.
17. Demuth H., Beale M., Neural Network Toolbox User's Guide, MathWorks, Inc, 2014.
18. Osowski S., Neural networks in algorithmic use, Warszawa: WNT, 1996 (in Polish).
19. Jurdziak L., Błażej R., Bajda M.: Conveyor belt 4.0, Intelligent systems in production engineering and maintenance., Advances in Intelligent Systems and Computing, ISSN 2194-5357; vol. 835. https://doi.org/10.1007/978-3-319-97490-3_61, Springer, cop. 2019, s. 645-654.
20. https://mapadotacji.gov.pl/projekty/1179135/, Integrated mobile system of automatic testing and continuous diagnostics of the condition of conveyor belts.