APEM journal
Advances in Production Engineering & Management **ISSN 1854-6250**
Volume 18 | Number 2 | June 2023 | pp 237–249 Journal home: apem-journal.org
https://doi.org/10.14743/apem2023.2.470 Original scientific paper

Enhancing automated defect detection through sequential clustering and classification: An industrial case study using the Sine-Cosine Algorithm, Possibilistic Fuzzy *c*-means, and Artificial Neural Network

Nguyen, T.P.Q.[a,*], Yang, C.L.[b], Le, M.D.[c], Nguyen, T.T.[d], Luu, M.T.[e]

[a]Faculty of Project Management, The University of Danang – University of Science and Technology, Danang, Vietnam
[b]Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan (ROC)
[c]Faculty of Transportation Mechanical Engineering, The University of Danang – University of Science and Technology, Danang, Vietnam
[d]Faculty of Computer Science and Engineering, University of Technology, Vietnam
[e]Nghia Nippers Corporation, Vietnam

A B S T R A C T

Most existing inspection models solely classify defects as either good or bad, focusing primarily on separating flaws from perfect ones. The sequential clustering and classification technique (SCC) is used in this work to not only identify and categorize the defects but also investigate their root causes. Conventional clustering techniques like *k*-means, fuzzy *c*-means, and self-organizing map are employed in the first stage to find the defects in the finished products. Then, a novel clustering method, that combines a sine-cosine algorithm and possibilistic fuzzy *c*-means (SCA-PFCM), is proposed to classify the detected defects into multiple groups to identify the defect categories and analyze the root causes of failures. In the second stage, the ground truth labels taken from the clustering technique are used to construct an automated inspection system using back propagation neural networks (BPNN). The proposed approach is applicable for detecting and identifying the causes of errors in manufacturing industry. This study applies a case study in nipper manufacture. The SCA-PFCM algorithm can detect 97 % of defects and classify them into four types while BPNN shows a predicted accuracy of up to 96 %. Additionally, an automated inspection system is developed to reduce the time and cost of the inspection process.

A R T I C L E I N F O

Keywords:
Back Propagation Neural Network;
Clustering;
Classification;
Combined SCA-PFCM;
Defect detection;
Nipper manufacturing;
Possibilistic Fuzzy *c*-means (PFCM);
Root cause analysis;
Sine-Cosine Algorithm (SCA)

**Corresponding author:*
ntpquyen@dut.udn.vn
(Nguyen, T.P.Q)

Article history:
Received 24 April 2023
Revised 12 June 2023
Accepted 23 June 2023

References

[1] Yun, J.P., Shin, W.C., Koo, G., Kim, M.S., Lee, C., Lee, S.J. (2020). Automated defect inspection system for metal surfaces based on deep learning and data augmentation, *Journal of Manufacturing Systems*, Vol. 55, 317-324, doi: 10.1016/j.jmsy.2020.03.009.

[2] Wang, T., Chen, Y., Qiao, M., Snoussi, H. (2018). A fast and robust convolutional neural network-based defect detection model in product quality control, *International Journal of Advanced Manufacturing Technology*, Vol. 94, No. 9, 3465-3471, doi: 10.1007/s00170-017-0882-0.
[3] Ko, T., Lee, J.H., Cho, H., Cho, S., Lee, W., Lee, M. (2017). Machine learning-based anomaly detection via integration of manufacturing, inspection and after-sales service data, *Industrial Management & Data Systems*, Vol. 117, No. 5, 927-945, doi: 10.1108/IMDS-06-2016-0195.
[4] Yang, Z.J., Du, X.J., Chen, F., Chen, C.H., Tian, H.L., He, J.L. (2018). Change-point estimation for repairable systems combining bootstrap control charts and clustering analysis: Performance analysis and a case study, *Advances in Production Engineering & Management*, Vol. 13, No. 3, 307-320, doi: 10.14743/apem2018.3.292.
[5] Božič, J., Tabernik, D., Skočaj, D. (2021). Mixed supervision for surface-defect detection: From weakly to fully supervised learning, *Computers in Industry*, Vol. 129, Article No. 103459, doi: 10.1016/j.compind.2021.103459.
[6] Chongwatpol, J. (2015). Prognostic analysis of defects in manufacturing, *Industrial Management & Data Systems*, Vol. 115, No. 1, 64-87, doi: 10.1108/IMDS-05-2014-0158.
[7] Yang, C.-L., Quyen, N.T.P. (2018). Data analysis framework of sequential clustering and classification using non-dominated sorting genetic algorithm, *Applied Soft Computing*, Vol. 69, 704-718, doi: 10.1016/j.asoc.2017.12.019.
[8] Yang, C.-L., Nguyen, T.P.Q. (2022). Sequential clustering and classification approach to analyze sales performance of retail stores based on point-of-sale data, *International Journal of Information Technology & Decision Making,* Vol. 21, No. 3, 885-910, doi: 10.1142/S0219622022500079.
[9] Kuo, R.J., Setiawan, M.R., Nguyen, T.P.Q. (2022). Sequential clustering and classification using deep learning technique and multi-objective sine-cosine algorithm, *Computers & Industrial Engineering*, Vol. 173, Article No. 108695, doi: 10.1016/j.cie.2022.108695.
[10] Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 24, No. 7, 881-892, doi: 10.1109/TPAMI.2002.1017616.
[11] Bezdek, J.C., Ehrlich, R., Full, W. (1984). FCM: The fuzzy *c*-means clustering algorithm, *Computers & Geosciences*, Vol. 10, No. 2-3, 191-203, doi: 10.1016/0098-3004(84)90020-7.
[12] Vesanto, J., Alhoniemi, E. (2000). Clustering of the self-organizing map, *IEEE Transactions on Neural Networks,* Vol. 11, No. 3, 586-600, doi: 10.1109/72.846731.
[13] Ienco, D., Bordogna, G. (2018). Fuzzy extensions of the DBScan clustering algorithm, *Soft Computing*, Vol. 22, No. 5, 1719-1730, doi: 10.1007/s00500-016-2435-0.
[14] Goh, A.T.C. (1995). Back-propagation neural networks for modeling complex systems*, Artificial Intelligence in Engineering*, Vol. 9, No. 3, 143-151, doi: 10.1016/0954-1810(94)00011-S.
[15] Nguyen, T.P.Q., Kuo, R.J., Le, M.D., Nguyen, T.C., Le, T.H.A. (2022). Local search genetic algorithm-based possibilistic weighted fuzzy *c*-means for clustering mixed numerical and categorical data, *Neural Computing and Applications*, Vol. 34, No. 20, 18059-18074, doi: 10.1007/s00521-022-07411-1.
[16] Tan, P.N., Steinbach, M., Kumar, V. (2016). *Introduction to data mining,* Pearson Education, New Delhi, India.
[17] Zhang, Z.L., Wang, Y.F., Li, Y. (2019). Inventory control model based on multi-attribute material classification: An integrated grey-rough set and probabilistic neural network approach, *Advances in Production Engineering & Management*, Vol. 14, No. 1, 93-111, doi: 10.14743/apem2019.1.314.
[18] Zhang, Y.D., Liao, L., Yu, Q., Ma, W.G., Li, K.H. (2021). Using the gradient boosting decision tree (GBDT) algorithm for a train delay prediction model considering the delay propagation feature, *Advances in Production Engineering & Management*, Vol. 16, No. 3, 285-296, doi: 10.14743/apem2021.3.400.
[19] Tao, X., Zhang, D., Ma, W., Liu, X., Xu, D. (2018). Automatic metallic surface defect detection and recognition with convolutional neural networks, *Applied Sciences,* Vol. 8, No. 9, Article No. 1575, doi: 10.3390/app8091575.
[20] Keller, J.M., Gray, M.R., Givens, J.A. (1985). A fuzzy K-nearest neighbor algorithm*, IEEE Transactions on Systems, Man, and Cybernetics,* Vol. 15, No. 4, 580-585, doi: 10.1109/TSMC.1985.6313426.
[21] Gao, H.N., Shen, D.H., Yu, L., Zhang, W.C. (2020). Identification of cutting chatter through deep learning and classification, *International Journal of Simulation Modelling*, Vol. 19, No. 4, 667-677, doi: 10.2507/IJSIMM19-4-CO16.
[22] Wang, Y.D., Lu, X.C., Shen, J.R. (2021). Improved Genetic Algorithm (VNS-GA) using polar coordinate classification for workload balanced multiple Traveling Salesman Problem (mTSP), *Advances in Production Engineering & Management*, Vol. 16, No. 2, 173-184, doi: 10.14743/apem2021.2.392.
[23] Zhang, L., Ma, J., Liu, X., Zhang, M., Duan, X., Wang, Z. (2022). A novel support vector machine model of traffic state identification of urban expressway integrating parallel genetic and *c*-means clustering algorithm, *Tehnički Vjesnik – Technical Gazette,* Vol. 29, No. 3, 731-741, doi: 10.17559/TV-20211201014622.
[24] Tawhid, M.A., Savsani, V. (2019). Multi-objective sine-cosine algorithm (MO-SCA) for multi-objective engineering design problems*, Neural Computing and Applications*, Vol. 31, No. 2, 915-929, doi: 10.1007/s00521-017-3049-x.
[25] Chaouni Benabdellah, A., Benghabrit, A., Bouhaddou, I. (2019). A survey of clustering algorithms for an industrial context, *Procedia Computer Science,* Vol. 148, 291-302, doi: 10.1016/j.procs.2019.01.022.
[26] Gocken, T., Yaktubay, M. (2019). Comparison of different clustering algorithms via genetic algorithm for VRPTW, *International Journal of Simulation Modelling*, Vol. 18, No. 4, 574-585, doi: 10.2507/IJSIMM18(4)485.
[27] Kuo, R.J., Zheng, Y.R., Nguyen, T.P.Q. (2021). Metaheuristic-based possibilistic fuzzy k-modes algorithms for categorical data clustering, *Information Sciences*, Vol. 557, 1-15, doi: 10.1016/j.ins2020.12.051.
[28] Nguyen, T.P.Q., Kuo, R.J. (2019). Automatic fuzzy clustering using non-dominated sorting particle swarm optimization algorithm for categorical data, *IEEE Access*, Vol. 7, 99721-99734, doi: 10.1109/ACCESS.2019.2927593.
[29] Mirjalili, S. (2016). SCA: A sine cosine algorithm for solving optimization problems, *Knowledge-Based Systems*, Vol. 96, 120-133, doi: 10.1016/j.knosys.2015.12.022.
[30] Geng, Y.A., Li, Q., Zheng, R., Zhuang, F., He, R., Xiong, N. (2018). RECOME: A new density-based clustering algorithm using relative KNN kernel density, *Information Sciences*, Vol. 436-437, 13-30, doi: 10.1016/j.ins.2018.01.013.

[31] Cui, Y., Liu, H., Wang, Q., Zheng, Z., Wang, H., Yue, Z., Ming, Z., Wen, M., Feng, L., Yao, M. (2022). Investigation on the ignition delay prediction model of multi-component surrogates based on back propagation (BP) neural network, *Combustion and Flame*, Vol. 237, Article No. 111852, doi: 10.1016/j.combustflame.2021.111852.

[32] Dong, Y., Fu, Z., Peng, Y., Zheng, Y., Yan, H., Li, X. (2020). Precision fertilization method of field crops based on the Wavelet-BP neural network in China, *Journal of Cleaner Production*, Vol. 246, Article No. 118735, doi: 10.1016/j.jclepro.2019.118735.

[33] Sadowski, P. (2016). Notes on backpropagation, from http://web.khu.ac.kr/~tskim/MLPR%2025-3%20Back-propagation%20with%20CE%20(Handout).pdf, accessed 5 April 2023.

[34] Wang, K.-J., Hao, F.-J., Lee, Y.-X. (2022). A multiple-stage defect detection model by convolutional neural network, *Computers & Industrial Engineering*, Vol. 168, Article No. 108096, doi: 10.1016/j.cie.2022.108096.

[35] Chicco, D., Warrens, M.J., Jurman, G. (2021). The Matthews correlation coefficient (MCC) is more informative than Cohen's Kappa and Brier score in binary classification assessment, *IEEE Access*, Vol. 9, 78368-78381, doi: 10.1109/ACCESS.2021.3084050.