Scalable processing of Bosch Welding Data with Graph Massivizer

Mikel Mendibe[1], Gad-Elrab Mohamed[1] and Evgeny Kharlamov[1,2]

[1]Bosch Center for Artificial Intelligence, Germany
[2]University of Oslo, Norway

Abstract

The transition to Industry 4.0 is driving the adoption of IoT technologies in manufacturing, transforming processes like Resistance Spot Welding (RSW) at companies such as Bosch. RSW is essential in automotive production but presents significant challenges in quality assurance due to its complexity and the limitations of traditional offline inspection methods. While data-driven techniques using dynamic electrical features offer some improvements, they often lack the ability to fully capture the intricacies of the RSW process or require expensive hardware.

This paper introduces a hybrid approach from the Graph Massivizer project, combining sensor data with domain-specific knowledge in a knowledge graph. The proposed method leverages existing ontologies as inductive biases to enhance prediction accuracy and adaptability across various welding scenarios. By integrating time-series data with knowledge graph representations, the approach offers a more holistic understanding of the welding process. This paper discusses the methodology, the challenges of merging these data types, and the potential to improve RSW quality monitoring by bridging the gap between conventional data-driven methods and expert insights, resulting in a more robust and efficient monitoring framework.

Keywords

Industry 4.0, resistance spot welding, knowledge graphs, Graph Massivizer

1. Introduction

With the advent of the fourth industrial revolution, the manufacturing environments are evolving towards digital ecosystems. Propelled by the improvement in sensing and communication technologies, the Internet of Things (IoT) has taken root in traditional industries, and this has opened the world of Big Data, where the volume, velocity and variety of data is rapidly rising [1]. This has been no different for the case of Bosch, currently undergoing a transformation from a traditional manufacturing company to an IoT one[2].

One of the most critical manufacturing processes at Bosch is Resistance Spot Welding (RSW), primarily employed for joining car body components. This process involves passing a high-intensity electrical current through two electrodes and the metal sheets positioned between them, generating localized heat that fuses the sheets together at specific points, creating a weld spot [3]. Even if it is considered a conventional method, this is still a very complex yet critical procedure [4], as the failure of a single spot can halt an entire production line [3]. In the industry, quality assurance for RSW is typically achieved via offline destructive inspection methods on samples selected by predefined empirical rules, but due to the heavy class imbalance on the sets, detecting faulty spots is usually a very rare occurrence. Consequently, data-driven approaches have been proposed in recent years, with the objective of detecting defective samples more efficiently and reducing scrap and downtime [4]. Some of these approaches focus on image processing, and they are oriented towards assessing the quality of the spots in an online fashion with the help of cameras mounted on the welding devices Prediction of the Weld Qualities Using Surface Appearance Image in Resistance Spot Welding, Monitoring of Resistance Spot Welding Process). Others, mainly directed by the work of Chen et al., developed systems for weld quality estimation based on infrared thermography [5] [6]. However, implementing these methods is often challenging due to the

First International Workshop on Scaling Knowledge Graphs for Industry, co-located with 20th International Conference on Semantic Systems (SEMANTICS) - Amsterdam, Sept. 17–19, 2024
mikel.mendibeabarrategi@de.bosch.com (M. Mendibe); mohamed.gad-elrab@de.bosch.com (G. Mohamed); evgeny.kharlamov@de.bosch.com (E. Kharlamov)

need to acquire, install, and integrate specialized hardware, which can be costly and time-consuming. Therefore, the research focused on using intrinsic process variables for monitoring is especially valuable. From the early works of Cho et al. [7], many different data-driven methods have been developed, either defined as classification or prediction tasks, depending on whether the intention is to estimate the welding nugget diameter, or to identify defective welds. In this paper, developed under the Graph Massivizer project, we introduce the need for a hybrid approach for welding quality estimation, one that leverages on existing ontologies to act as an inductive bias on top of the available sensor data, with the objective of obtaining a more grounded and efficient algorithm.

2. Use case

Resistance Spot Welding (RSW) is widely used in the automotive industry, as over 90% of assembly work in a car body is completed by RSW [8]. These spots are critical for the structural integrity of the vehicle, and the amount of spot welds in a car are counted in the thousands [9]. Resistance Spot Welding is a process used to join two or more metal sheets together by applying pressure and passing an electric current through the materials. The process generates heat due to the resistance to the electrical current in the contact area between the metal sheets, causing the metals to melt and fuse together. The welding electrodes also provide the necessary force to hold the materials in place during welding [10]. This is a very complicated process involving electromagnetic, thermal, mechanical and metallurgical variables [8]. Even in very controlled environments, defects have been shown to be very difficult to detect, as the welds are hidden between two sheets of metal. As such, in resistance spot welding, quality assurance typically involves the use of offline destructive inspection techniques on selected samples. These methods involve directly measuring the mechanical strength, fatigue resistance, and failure modes of the weld joints. By physically testing these samples to the point of failure, critical information about the integrity and durability of the weld can be obtained. This approach, while effective in assessing joint quality, is limited by its destructive nature and the fact that it only provides insight into the specific samples tested, rather than the entire production run [4].

Most modern RSW operations are very efficient, and defective welds are a rare yet critical occurrence, as such, defining data-driven methods to target these defects is a priority for the automotive industry, and it has been a very rich subject of research. The main difficulties explored in the literature for the implementation of efficient quality prediction algorithms are the following: (i) limit on labeled data amount, (ii) limit on features, (iii) limit on coverage of relevant situations [11]. Increasing the amount of labeled data is a resource-intensive task, therefore, it is often best avoided. Regarding (ii), some authors have made use of additional sensors (infrared sensors [5], cameras...) for reliable quality prediction, however, this is often difficult to implement and over-complicates the process, which leads to more failures and downtime. Lastly, (iii) is the most pressing challenge, because as quality is good under normal conditions, the datasets are often times heavily imbalanced, and creating more anomalous labels requires expensive laboratory testing.

Therefore, it is specifically interesting to detect RSW defects using the highly available dynamic electrical features captured during the process, namely the current and the voltage, which are used to conform to the dynamic resistance. These methods treat the dynamic curves as time series and aim at identifying the underlying patterns that lead to anomalies in the welds. In [12], a random forest classifier was trained with features derived from this signal and achieved up to 98% accuracy. Similarly, [13] used a radial basis function neural network to predict nugget diameter and weld strength, achieving 98% accuracy by classifying features into different quality levels. In [14], a kernel extreme machine model optimized with a particle swarm algorithm also reached 98% accuracy. The performance of these models is very solid in their specific scenarios, and it leads to the incorrect idea that quality assurance in RSW is a solved topic. However, as mentioned by Stavropoulos et al. [4], RSW is a very heterogeneous process, and the variation of the elements that comprise the RSW ecosystem and process (machines, materials, thickness, position...) heavily affect the final quality of the joint.

Towards this direction, a need for a more holistic approach has been defined, one that takes into account a bigger portion of the welding ecosystem and is able to adapt to the different flavors that RSW can have. It is this approach that Tan et al. followed [15] by framing the quality prediction problem as one of link prediction in a literal-aware RSW welding knowledge graph. By integrating background knowledge on the prediction, the model is able to get a higher-level view of the operation and tackle the heterogeneity of RSW in a more holistic manner. Even if the performance of this approach is notable, in an attempt to frame it under the more traditional embedding methods, it deals with the time-series nature of the curves by aggregating the values of the dynamical electrical features, leading to a great loss of expressivity and decrease in performance, as Dickinson et al. stated in [16].

3. Proposed methodology

It has been shown that while data-driven methods lack the capacity to model the complexities of the RSW ecosystem, knowledge-graph-driven approaches lack the expressivity and granularity afforded by time-series data. Thus, motivated by the work of Jain et al. [17], the objective of Bosch and the consortium partners under the Graph Massivizer project is to merge both approaches and design the next-generation quality monitoring methods that go beyond the more traditional data-driven methods by combining knowledge and sensor measurements.

On the one hand, we refer to sensor measurements as the time series that are produced during the welding operation, which can have the form of currents, voltages, temperatures… These are variables that evolve during each weld and are fundamental to the estimation of the quality (e.g., if an abnormality has been detected in the current that flows through the cathode of the welding machine, the presence of an anomaly will be likely). On the other hand, we refer to knowledge that encompasses diverse and rich prior information about the process, derived from expert knowledge, such as the ontology for a specific RSW machine. This ontology is used as the backbone to create knowledge graphs for each specific operation, similar to the approach that Tan et al. followed in [15], but instead of pointing at the aggregated values of the dynamic features, it points at the entire representation of the time-series, thus maintaining the expressivity in the analysis. This ontology is then used as a structural inductive bias that inherently captures the nature of RSW. Subsequently, a Graph Neural Network (GNN) is trained for prediction, leveraging the ontology-based knowledge graph to enhance its learning capability [18].

This method is not only expected to improve the model's accuracy in predicting outcomes but also enable it to generalize better across different RSW operations by integrating domain-specific knowledge. Additionally, the use of GNNs allows for the incorporation of complex relationships and dependencies between various process parameters, providing a more holistic and detailed understanding of the RSW process dynamics. This approach demonstrates the potential to bridge the gap between data-driven methods and expert knowledge, resulting in a more robust and comprehensive modeling framework.

4. Drawbacks

Despite the high expectations for the project and the computational soundness of the method, we are currently encountering some challenges.

- A significant challenge lies in effectively representing time-series data within a graph-based framework. While time-series data provide rich, dynamic information about the welding process, converting this data into a graph structure that preserves its temporal properties and contextual relevance is complex. One possible solution is to use visibility graphs, which transform time-series data into networks where each data point (or "node") is connected to other nodes based on specific visibility criteria. Although visibility graphs can capture the intrinsic patterns and relationships within the time series, they may struggle with scalability when dealing with large datasets or high-dimensional data. Furthermore, it remains unclear how to best integrate these graph-based representations with existing domain knowledge in a cohesive manner that enhances the model's predictive capability.

- Combining knowledge graph data with time-series data in a meaningful way remains an open question. For example, a potential approach could involve appending embeddings of events generated by a BERT-like model [19] to embeddings derived from a feature tokenizer of the time-series data, as proposed by Jain et al. [17]. However, this strategy may not be effective in our use case due to the scarcity of events, which limits the potential information gain from such embeddings.
- Using ontologies as structural inductive biases, while theoretically sound, raises questions about scalability. For instance, one could use a BERT-like model to create embeddings for materials, static conditions, and other domain-specific knowledge, and then train a Graph Neural Network (GNN) with this data. However, it is uncertain how well this approach would scale to larger datasets or more complex welding scenarios. Moreover, techniques like node2vec [20], which are often used for node embeddings in graphs, may not be suitable for capturing the nuanced, multi-relational nature of the knowledge graph in this context.

5. Conclusion

In conclusion, this paper highlights the need for a hybrid approach to quality assurance in Resistance Spot Welding (RSW) by combining knowledge graphs and time-series data to address the limitations of traditional data-driven and knowledge-graph-based methods. The proposed methodology, leveraging Graph Neural Networks (GNNs) and domain-specific ontologies, offers a more comprehensive and adaptive modeling framework that captures the complex dependencies and dynamics of the RSW process, leading to improved prediction accuracy and robustness. However, challenges remain in effectively integrating time-series data within a graph-based framework and ensuring scalability across diverse welding scenarios.

Acknowledgments

The Graph Massiviser (GA 101093202) EU project supported this work.

References

[1] D. Mourtzis, E. Vlachou, N. Milas, Industrial Big Data as a Result of IoT Adoption in Manufacturing, Procedia CIRP 55 (2016) 290–295. URL: https://www.sciencedirect.com/science/article/pii/S2212827116307880. doi:10.1016/j.procir.2016.07.038.

[2] A.-K. Leiting, L. De Cuyper, C. Kauffmann, The Internet of Things and the case of Bosch: Changing business models while staying true to yourself, Technovation 118 (2022) 102497. URL: https://www.sciencedirect.com/science/article/pii/S016649722200044X. doi:10.1016/j.technovation.2022.102497.

[3] B. Zhou, Y. Svetashova, S. Byeon, T. Pychynski, R. Mikut, E. Kharlamov, Predicting Quality of Automated Welding with Machine Learning and Semantics: A Bosch Case Study, in: Proceedings of the 29th ACM International Conference on Information & Knowledge Management, CIKM '20, Association for Computing Machinery, New York, NY, USA, 2020, pp. 2933–2940. URL: https://doi.org/10.1145/3340531.3412737. doi:10.1145/3340531.3412737.

[4] P. Stavropoulos, K. Sabatakakis, Quality Assurance in Resistance Spot Welding: State of Practice, State of the Art, and Prospects, Metals 14 (2024) 185. URL: https://www.mdpi.com/2075-4701/14/2/185. doi:10.3390/met14020185, number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

[5] J. Chen, Z. Feng, Online resistance spot weld NDE using infrared thermography 10169 (2017) 101690K. URL: https://ui.adsabs.harvard.edu/abs/2017SPIE10169E..0KC. doi:10.1117/12.2260896, conference Name: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series ADS Bibcode: 2017SPIE10169E..0KC.

[6] J. Chen, Z. Feng, IR-based spot weld NDT in automotive applications, in: Thermosense: Thermal Infrared Applications XXXVII, volume 9485, SPIE, 2015, pp. 283–288. URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9485/948513/IR-based-spot-weld-NDT-in-automotive-applications/10.1117/12.2177124.full. doi:10.1117/12.2177124.

[7] Y. Cho, S. Rhee, Quality Estimation of Resistance Spot Welding by Using Pattern Recognition With Neural Networks, IEEE Transactions on Instrumentation and Measurement 53 (2004) 330–334. URL: http://ieeexplore.ieee.org/document/1284862/. doi:10.1109/TIM.2003.822713.

[8] Y. Li, Z. Lin, L. Li, G. Chen, Numerical Analysis of Transport Phenomena in Resistance Spot Welding Process, volume 1, 2009. doi:10.1115/FEDSM2009-78210, journal Abbreviation: Proceedings of the ASME Fluids Engineering Division Summer Conference 2009, FEDSM2009 Publication Title: Proceedings of the ASME Fluids Engineering Division Summer Conference 2009, FEDSM2009.

[9] K. Zhou, P. Yao, Overview of recent advances of process analysis and quality control in resistance spot welding, Mechanical Systems and Signal Processing 124 (2019) 170–198. URL: https://www.sciencedirect.com/science/article/pii/S0888327019300573. doi:10.1016/j.ymssp.2019.01.041.

[10] R. L. O'Brien, Welding handbook. Volume 2, Welding processes, 8th ed ed., American Welding Society, Miami, Fla., 1991. OCLC: 664579333.

[11] B. Zhou, T. Pychynski, M. Reischl, R. Mikut, Comparison of Machine Learning Approaches for Time-series-based Quality Monitoring of Resistance Spot Welding (RSW) (2018). doi:10.5445/KSP/1000087327/13.

[12] B. Xing, Y. Xiao, Q. H. Qin, H. Cui, Quality assessment of resistance spot welding process based on dynamic resistance signal and random forest based, The International Journal of Advanced Manufacturing Technology 94 (2018) 327–339. URL: https://doi.org/10.1007/s00170-017-0889-6. doi:10.1007/s00170-017-0889-6.

[13] H. Zhang, Y. Hou, T. Yang, Q. Zhang, J. Zhao, Welding quality evaluation of resistance spot welding using the time-varying inductive reactance signal, Measurement Science and Technology 29 (2018) 055601. URL: https://dx.doi.org/10.1088/1361-6501/aaa830. doi:10.1088/1361-6501/aaa830, publisher: IOP Publishing.

[14] H. Sun, J. Yang, L. Wang, Resistance spot welding quality identification with particle swarm optimization and a kernel extreme learning machine model, The International Journal of Advanced Manufacturing Technology 91 (2017) 1879–1887. URL: https://doi.org/10.1007/s00170-016-9944-y. doi:10.1007/s00170-016-9944-y.

[15] Z. Tan, B. Zhou, Z. Zheng, O. Savkovic, Z. Huang, I.-G. Gonzalez, A. Soylu, E. Kharlamov, Literal-Aware Knowledge Graph Embedding for Welding Quality Monitoring: A Bosch Case, in: T. R. Payne, V. Presutti, G. Qi, M. Poveda-Villalón, G. Stoilos, L. Hollink, Z. Kaoudi, G. Cheng, J. Li (Eds.), The Semantic Web – ISWC 2023, Springer Nature Switzerland, Cham, 2023, pp. 453–471. doi:10.1007/978-3-031-47243-5_25.

[16] D. W. Dickinson, Characterization of Spot Welding Behavior by Dynamic Electrical Parameter Monitoring (????).

[17] S. Jain, M. Burger, G. Rätsch, R. Kuznetsova, Knowledge Graph Representations to enhance Intensive Care Time-Series Predictions, 2023. URL: http://arxiv.org/abs/2311.07180, arXiv:2311.07180 [cs].

[18] M. Trębacz, Z. Shams, M. Jamnik, P. Scherer, N. Simidjievski, H. A. Terre, P. Liò, Using ontology embeddings for structural inductive bias in gene expression data analysis, 2020. URL: http://arxiv.org/abs/2011.10998. doi:10.48550/arXiv.2011.10998, arXiv:2011.10998 [cs, q-bio].

[19] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, in: J. Burstein, C. Doran, T. Solorio (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 4171–4186. URL: https://aclanthology.org/N19-1423. doi:10.18653/v1/N19-1423.

[20] A. Grover, J. Leskovec, node2vec: Scalable Feature Learning for Networks, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, Association for Computing Machinery, New York, NY, USA, 2016, pp. 855–864. URL: https://doi.org/10.1145/2939672.2939754. doi:10.1145/2939672.2939754.