UniCas for Industry

A. Miele[1,†], H. Mustafa[1,†], M. Vitelli[1,2,*,†], A. Bria[1,†], C. De Stefano[1,†], F. Fontanella[1,†], C. Marrocco[1,†], M. Molinaria[1,†] and A. Scotto di Freca[1,†]

[1]University of Cassino and Southern Lazio, Cassino, 03043, Italy
[2]Sensichips s.r.l., 04011 Aprilia, Italy

Abstract

Artificial Intelligence (AI) is transforming industries, particularly through Industry 4.0, by integrating technologies such as the Internet of Things (IoT) to optimize production processes and resource management. It addresses challenges such as reducing environmental impact while fulfilling consumer demands. Innovative sensors enable real-time data collection for environmental monitoring. Adopting advanced technologies such as energy cells, particularly lithium-ion batteries, is crucial for sustainable mobility and reducing environmental impact in the automotive industry. It is vital to understand the key parameters of energy cells, including range, energy density, and durability, and implement them while embracing the principles of Second Life effectively. For example, machine learning (ML) algorithms are utilized in industrial contexts to identify air and water pollutants and estimate the State of Charge (SoC) for automotive applications. These methodologies improve efficiency, sustainability, and innovation in various industrial sectors.

Keywords

Artificial Intelligence, Industry 4.0, AI on the Edge, Smart Sensors, Pollutants Identification, State of Charge estimation

1. Introduction

Artificial Intelligence (AI) is revolutionizing various sectors including healthcare, finance, education, transportation, and notably, industry. Its capacity to analyze vast data sets in real-time and generate precise predictive insights is reshaping production processes, enhancing resource allocation, and boosting operational efficiency in industries worldwide. Industry 4.0 [1] represents a crucial turning point in the evolution of the industrial landscape, characterized by the integration of advanced technologies and widespread digitization of production processes. Inspired by the notion of the "smart factory," it emphasizes the interconnection of machines, systems, and people via IoT, AI, big data, cloud computing, and advanced robotics[2, 3, 4]. The concept of Industry 4.0 is based on the idea of automated and connected production, where machines and systems communicate with each other in real-time to optimize processes and decision-making. In this context, innovative sensors [5, 6] play a crucial role by enabling the collection of detailed, real-time data on various environmental and operational parameters. Using artificial intelligence to analyze and interpret sensor data offers numerous benefits. By utilizing sophisticated algorithms and predictive models, it is possible to identify pollutants accurately, continuously monitor air and water quality, and optimize industrial processes to reduce environmental impacts. However, the applications of artificial intelligence in industry are not limited to the environmental sphere. Nowadays, the automotive industry is facing one of the most significant challenges in its history: to provide sustainable mobility and reduce the environmental footprint of transportation on a global scale. One of the key pillars of this transformation is the energy cell[7]. Energy cells, notably lithium-ion batteries, are crucial in revolutionizing vehicle energy usage towards zero-emission transportation, combating air pollution, and mitigating climate change. However, to fully realize it is essential to have a precise comprehension of key parameters like range, energy density, charging time, and durability. Accurate estimation of these parameters is critical for developing large-scale zero-emission vehicles and ensuring proper disposal and reuse, aligning with Second Life principles[8].

The following sections highlight the application of Machine Learning (ML) in industrial challenges, focusing on the detection of pollutants in air (2) and water (3), and on State of Charge estimation in automotive applications (4).

2. Pollutant Identification in Air

Our recent study proposes a novel system integrating sensor technology and machine learning to detect and classify air contaminants effectively and affordably. Current monitoring solutions face size, cost, and complexity

Ital-IA 2024: 4th National Conference on Artificial Intelligence, organized by CINI, May 29-30, 2024, Naples, Italy

*Corresponding author.
†These authors contributed equally.
alessio.miele@unicas.it (A. Miele); hamza.mustafa@unicas.it (H. Mustafa); michele.vitelli@unicas.it (M. Vitelli); a.bria@unicas.it (A. Bria); destefano@unicas.it (C. De Stefano); fontanella@unicas.it (F. Fontanella); c.marrocco@unicas.it (C. Marrocco); m.molinara@unicas.it (M. Molinaria); a.scotto@unicas.it (A. Scotto di Freca)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

issues[9], prompting the development of a more accessible solution. Challenges like low sensitivity and selectivity in miniaturized, low-cost smart solutions are addressed, with comparisons made to CEN (European Committee for Standardization) reference instruments, noting lower accuracy and stability but highlighting their value in data aggregation[10]. Spatial analysis techniques aid in evaluating pollutant sources, while the limitations of chemical micro-sensors are extensively discussed in the literature, along with methodologies to improve their performance.

At the core of our proposed system is a sensor array including aluminum oxide for broad-spectrum volatile organic compound detection, a commercial capacitive humidity sensor, and graphene-functionalized sensor for pollutant sensitivity. These selections aim for versatility and sensitivity across contaminant types. Integration with the SENSIPLUS platform facilitates precise electrical impedance measurements, crucial for accurate air quality assessment. The proposed integrated system is shown in Figure 1 and is mainly composed of the following: (1) SENSIPLUS Chip (henceforth SPC): a microelectronic measurement device with on-chip sensing capabilities, jointly developed by Sensichips s.r.l.[11] and the Department of Information Engineering at the University of Pisa. Equipped with a versatile analog front end and various internal and external ports, it enables electrical impedance measurements on both internal and external sensors. It has already been adopted in other works, as in [12, 13, 14]. (2) SENSIPLUS Deep Machine (SDM): a hardware/software module designed for data acquisition, processing, and analysis. The block diagram in Figure 2 illustrates the logical flow of operations, highlighting the software and hardware components utilized for each task. Data acquisition is enabled by the SPC API, a software library in C or Java, operating respectively on Micro Controller Units (MCUs) and multiple hosts like Linux/Windows/Android, depending on application needs. Classification tasks utilize ML techniques like MLP, CNN, or LSTM, adaptable to run on MCU or more powerful devices like PCs, depending on computational requirements.

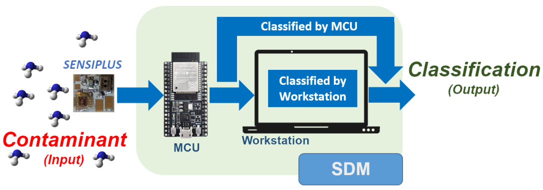

Figure 1: The proposed integrated system. SDM stands for SENSIPLUS Deep Machine.

Our methodology involves a structured measurement protocol to simulate various indoor air quality conditions.

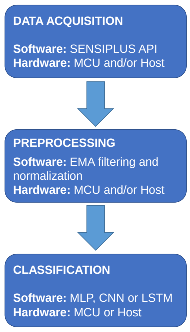

Figure 2: The proposed integrated system. SDM stands for SENSIPLUS Deep Machine.

This includes phases of baseline air exposure, controlled introduction of contaminants, and subsequent recovery, designed to capture the dynamic nature of indoor air quality. This methodological approach is crucial for producing comprehensive sensor data that reflects the complexities of real-world indoor environments.

For the analytical component of our study, we implemented several machine learning models, including Multi-Layer Perceptrons (MLP), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks. These models were trained on datasets collected from our sensor array, to achieve high accuracy in classifying different air contaminants. The contaminants included in our study encompass a range of substances commonly found in indoor settings, such as acetone, alcohol, ammonia, bleach, and various volatile organic compounds (VOCs), along with controls like water vapor and clean air to facilitate accurate classification between polluted and unpolluted conditions.

Our findings indicate that the system can classify air contaminants with an average accuracy surpassing 75%, showcasing its potential effectiveness in indoor air quality assessment. However, classification accuracy varied among different contaminants, with notable challenges in distinguishing similar substances like acetone and alcohol. This variation underscores the complexities of air quality monitoring and identifies avenues for future enhancement.

In evaluating the system's operational efficiency, we prioritized minimizing data acquisition times and energy consumption, optimizing for low-power operations ideal for IoT applications. This focus ensures the effectiveness and practicality of our solution for real-world deployment, highlighting the importance of efficiency in environmental monitoring technologies.

Looking forward, we anticipate several potential enhancements to our system. These include integrating additional sensor types to expand the range of detectable contaminants, exploring advanced machine learning

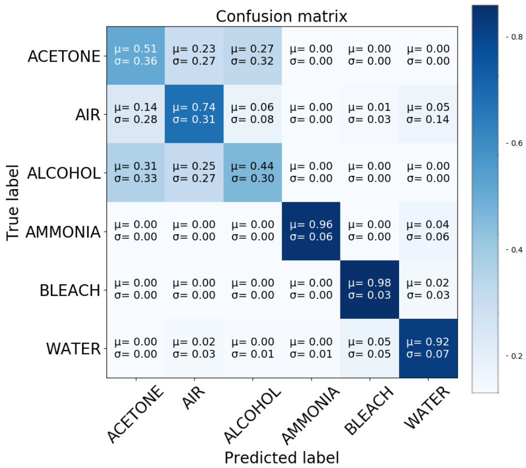

Figure 3: CNN Global Confusion Matrix.

models to enhance classification accuracy, and developing expanded real-time monitoring capabilities. These efforts aim to further improve the comprehensiveness and usability of our indoor air quality monitoring system.

In conclusion, our work contributes to environmental monitoring efforts by demonstrating the feasibility of a sensor-based and machine-learning-integrated system for indoor air quality assessment. While promising, our results also highlight the challenges in air quality monitoring and the necessity for continued innovation in this field. Our study represents a step toward achieving more accessible, efficient, and accurate air quality monitoring solutions.

3. Pollutant Identification in Water

Detecting illegal pollutants in wastewater is crucial for public health and security. An End-to-End IoT-ready node is proposed for sensing, processing, and transmitting wastewater pollutant data. Utilizing Smart Cable Water with SENSIPLUS chip sensors, the system employs impedance spectroscopy to distinguish pollutants from other substances. Data processing, on a low-cost Micro Control Unit, involves anomaly detection, classification, and false positive reduction through machine learning algorithms.

3.1. Metodology

The identification system, depicted in Figure 4, utilizes the Smart Cable Water (SCW), an IoT-ready smart sensor system developed by Sensichips s.r.l. The SCW comprises InterDigitated Electrodes (IDEs) and is based on SENSIPLUS [15]. The system's objective is to detect substances in wastewater. However, direct measurements from sewage drains are impractical due to unreliable conditions and health risks. To address this challenge, Synthetic WasteWater (SWW) is created to simulate sewage composition. The recipe used to create SWW is inspired by previous work, and pH adjustments are made to replicate real wastewater conditions. Fourteen substances have been spilled in the SWW background: (1) Acetic Acid; (2) Acetone; (3) Ethanol; (4) Ammonia; (5) Formic Acid; (6) Phosphoric Acid; (7) Sulphuric Acid; (8) Hydrogen Peroxide; (9) Synthetic Waste Water; (10) Sodium Hypochlorite; (11) Sodium Chloride; (12) Dish Wash Detergent; (13) Wash Machine Detergent; (14) Nelsen. These substances are divided into two categories: substances to be identified (group 1) and outlier samples (group 2) to be excluded by the system. This method guarantees a safe environment for dataset creation without any biological risks.

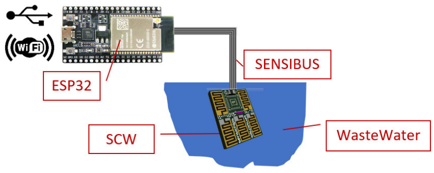

Figure 4: Identification system architecture.

To enhance sensitivity to the substances of interest and the RedOx dynamics, the six IDEs of the SCW were coated with six different metals: Gold (M1), Copper (M2), Silver (M3), Nickel (M4), Palladium (M5), and Platinum (M6). From the resulting sensors, we recorded the resistance measured at a frequency of 78 kHz for the Gold and Platinum IDEs, while Resistance and Capacitance were measured at a frequency of 200 Hz for Gold, Platinum, Silver, and Nickel. This yielded a feature vector comprising ten values: six resistance and four capacitance measurements. Notably, the experimental campaign did not involve the use of Palladium and Copper IDEs.

3.2. Classification

The classification system consists of two phases: Data Preprocessing and Classification. In the Data Preprocessing phase, raw sensor data is normalized before being sorted and evaluated by a Finite State Machine (FSM) shown in Figure 5. This process determines whether the data should proceed to the Classification phase.

The Data Preprocessing phase involves normalizing the raw data from sensors, establishing a robust baseline signal, and determining whether the normalized sample should proceed to the anomaly detector or be directly classified using the FSM.

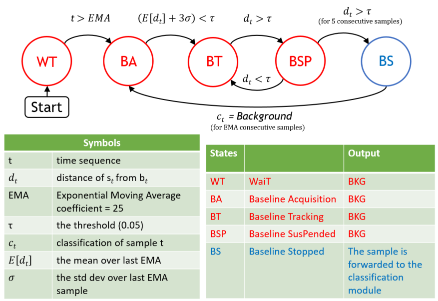

Figure 5: Identification system architecture.

In real scenarios, distinguishing between substances of interest and others in the sewerage network is crucial. The primary aim is to determine if the substance being investigated is of interest, minimizing subjective evaluations unless specified.

The identification phase involves anomaly detection and multiclass classification for precise substance identification. Anomaly detection excludes common substances, focusing on outliers, while classification employs optimized KNN models trained solely on samples of interest. Grid search methods enhance the accuracy of both anomaly detection and multiclass classification models.

3.3. Results

The study combined anomaly detection with a multiclass classifier for the final test, as illustrated in Figure 6. However, the multiclass classifier incorrectly identified some outlier substances, leading to false positive alarms. To mitigate this, the anomaly detection system was integrated before the multiclass classifier. Consequently, most outlier samples were accurately classified as 'UNKNOWN,' achieving an accuracy rate of 79.4%. Notably, 20.6% of outlier samples, primarily sodium hypochlorite, were frequently misclassified as hydrogen peroxide.

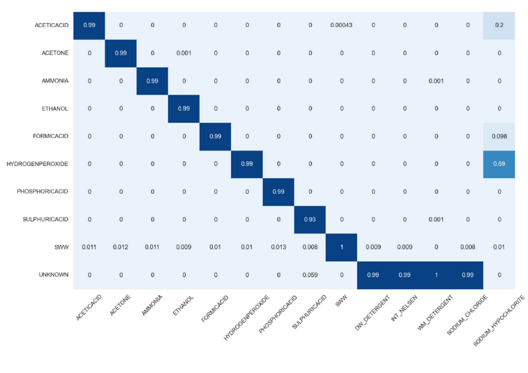

Figure 6: Entire system results shown as Confusion Matrix.

4. Optimization of Battery State of Charge Estimation

Accurate monitoring of State of Charge (SoC) is crucial for tasks like battery life estimation and temperature control. Existing techniques like Coulomb counting and Open Circuit Voltage (OCV) face challenges such as measurement errors and the flat relationship between voltage and SoC in certain battery types like Lithium Iron Phosphate (LFP). Electrochemical Impedance Spectroscopy (EIS) emerges as a promising alternative but suffers from long measurement times. This work proposes a method to minimize measurement time while ensuring accurate SoC estimation, particularly with EIS and knowledge-based SoC classes.

The proposed approach follows the framework shown in Figure 1. It starts with the identifying design parameters and constraints, which include: (1) Resolution of SoC performance estimation, (2) Target measurement time, (3) Target Accuracy, (4) Battery type, and (5) Classifier. The second step is to characterize the device under test,

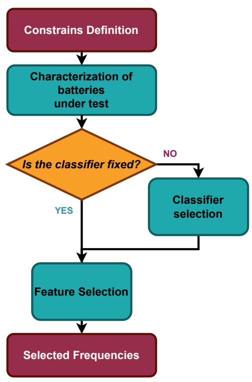

Figure 7: The proposed method workflow.

focusing on achieving the most stringent parameters possible. Then the appropriate classifier from the previous dataset is evaluated. The chosen classifier, demonstrating better accuracy, is then integrated into the feature selection algorithm. The final stage involves feature selection using search algorithms, aimed at minimizing measurement time while preserving accuracy above the specified target.

4.1. Metodology

In this example, the State of Charge (SoC) estimation problem was addressed using 10-class classification models where each class represents a 10% interval of the SoC. The initial dataset comprises all available features, including 28 impedances (real and imaginary parts) measured at various frequencies, totaling 56 features collected from 7 different cells. These features represent the Nyquist plots of battery impedances at different SoCs, as illustrated in Figure 8. Performance evaluation metrics used are described by Grandini et al[16]. The experiments

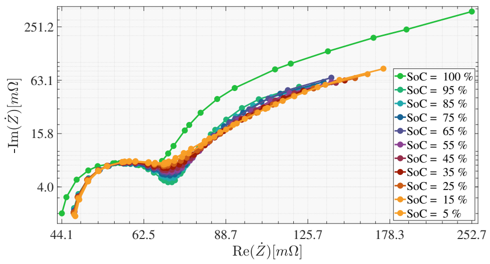

Figure 8: Obtained Nyquist plot of a single battery at different State of Charge.

consistently followed the k-fold method, ensuring optimal dataset utilization by rotating the batteries used. As a result, six trained models were obtained. Evaluation metrics confirm that the Support Vector Machine (SVM) model outperforms others, with a mean accuracy of 0.83 and a standard deviation of 0.04. The resulting confusion matrix shown in Figure 9 illustrates the performance of the SVM model. These preliminary classification tests identify SVM as the most effective ML model among those considered.

The problem of identifying the optimal set of frequencies for impedance measurement via EIS for battery SoC estimation has been addressed using optimization algorithms as search strategies, specifically Particle Swarm Optimization (PSO) [17]. A fitness function is implemented based on a supervised learning model 1, aiming to balance accuracy in SoC estimation and measurement time. The parameter 2 represents the ratio of correct predictions (CP) to total predictions (TP), while parameter 3 is inversely related to measurement duration. Measurement time (T_{meas}) is computed as the sum of selected feature durations, with T_{max} related to the use of all features. Parameters A and B range from 0 to 1, with α serving as a weight coefficient between accuracy and time contributions.

$$S = \alpha \cdot A + (1 - \alpha) \cdot B \tag{1}$$

$$A = \frac{CP}{TP} \tag{2}$$

$$B = 1 - \frac{T_{meas}}{T_{max}} \tag{3}$$

4.2. Results

This case study establishes a target accuracy of 0.95, regardless of measurement time. Figure 10 shows the maximum Accuracy achieved over 50 runs, correlating with the weight coefficient, while considering measurement time. The blue star indicates the solution with the highest Accuracy. The band represents SoC estimation Accuracy considering all features with the SVM classifier,showing a trade-off between accuracy and measurement time optimization, where higher α values prioritize accuracy over time.

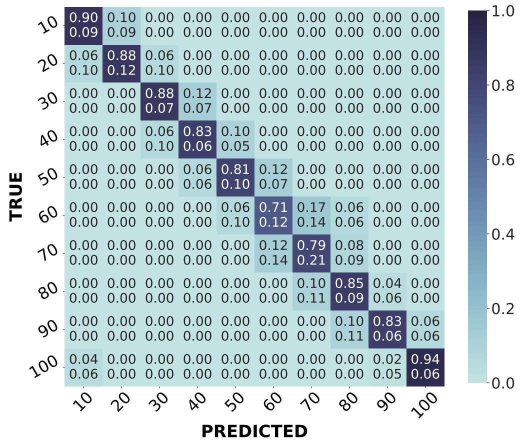

Figure 9: Obtained confusion matrix of the Support Vector Machine model, with mean value (top) and standard deviation (bottom) for each class.

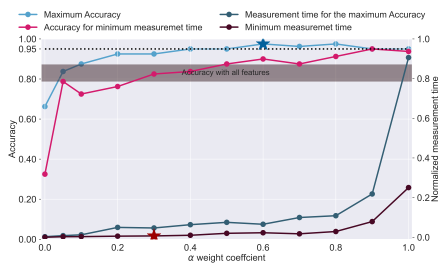

Figure 10: Accuracy and measurement time as a function of weight coefficient for PSO+SVM combination.

Acknowledgments

The EU partially supported this work in the NextGenerationEU plan through MUR Decree n. 1051 23.06.2022 "PNRR Missione 4 Componente 2 Investimento 1.5" - CUP H33C22000420001, partially through Horizon Europe RHE-MEDiation with (GA 101113045).

References

[1] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, J. Barata, Industrial artificial intelligence in industry 4.0 - systematic review, challenges and outlook, IEEE Access 8 (2020) 220121–220139. doi:10.1109/ACCESS.2020.3042874.

[2] J. Zhang, D. Tao, Empowering things with intelligence: A survey of the progress, challenges, and opportunities in artificial intelligence of things, IEEE Internet of Things Journal 8 (2021) 7789–7817. doi:10.1109/JIOT.2020.3039359.

[3] C. Li, Y. Chen, Y. Shang, A review of industrial big data for decision making in intelligent manufacturing, Engineering Science and Technology, an International Journal 29 (2022) 101021. URL: https://www.sciencedirect.com/science/article/pii/S2215098621001336. doi:https://doi.org/10.1016/j.jestch.2021.06.001.

[4] F. Sherwani, M. M. Asad, B. Ibrahim, Collaborative robots and industrial revolution 4.0 (ir 4.0), in: 2020 International Conference on Emerging Trends in Smart Technologies (ICETST), 2020, pp. 1–5. doi:10.1109/ICETST49965.2020.9080724.

[5] S. C. Mukhopadhyay, S. K. S. Tyagi, N. K. Suryadevara, V. Piuri, F. Scotti, S. Zeadally, Artificial intelligence-based sensors for next generation iot applications: A review, IEEE Sensors Journal 21 (2021) 24920–24932. doi:10.1109/JSEN.2021.3055618.

[6] G. Manfredini, A. Ria, P. Bruschi, L. Gerevini, M. Vitelli, M. Molinara, M. Piotto, An asic-based miniaturized system for online multi-measurand monitoring of lithium-ion batteries, Batteries 7 (2021). doi:10.3390/batteries7030045, cited by: 11; All Open Access, Gold Open Access, Green Open Access.

[7] A. Masias, J. Marcicki, W. A. Paxton, Opportunities and challenges of lithium ion batteries in automotive applications, ACS Energy Letters 6 (2021) 621–630. doi:10.1021/acsenergylett.0c02584.

[8] M. Shahjalal, P. K. Roy, T. Shams, A. Fly, J. I. Chowdhury, M. R. Ahmed, K. Liu, A review on second-life of li-ion batteries: prospects, challenges, and issues, Energy 241 (2022) 122881. URL: https://www.sciencedirect.com/science/article/pii/S0360544221031303. doi:https://doi.org/10.1016/j.energy.2021.122881.

[9] F. Rock, N. Barsan, U. Weimar, Electronic nose: Current status and future trends, Chemical Reviews 108 (2008) 705–725. doi:10.1021/cr068121q, pMID: 18205411.

[10] N. Castell, F. R. Dauge, P. Schneider, M. Vogt, U. Lerner, B. Fishbain, D. Broday, A. Bartonova, Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?, Environment International 99 (2017) 293 – 302. URL: http://www.sciencedirect.com/science/article/pii/S0160412016309989. doi:https://doi.org/10.1016/j.envint.2016.12.007.

[11] Sensichips, ???? URL: www.sensichips.com.

[12] G. Cerro et al., Preliminary realization of a monitoring system of activated carbon filter RLI based on the SENSIPLUS® microsensor platform, in: 2017 IEEE International Workshop on Measurement and Networking (M&N), 2017, pp. 1–5. doi:10.1109/IWMN.2017.8078361.

[13] G. Cerro et al., Metrological characterization of a novel microsensor platform for activated carbon filters monitoring, IEEE Transactions on Instrumentation and Measurement 67 (2018) 2504–2515.

[14] P. Bruschi et al., A novel integrated smart system for indoor air monitoring and gas recognition, in: 2018 IEEE Int. Conf. on Smart Computing (SMARTCOMP), 2018, pp. 470–475. doi:10.1109/SMARTCOMP.2018.00048.

[15] A. Ria, M. Cicalini, G. Manfredini, A. Catania, M. Piotto, P. Bruschi, The sensiplus: A single-chip fully programmable sensor interface, in: S. Saponara, A. De Gloria (Eds.), Applications in Electronics Pervading Industry, Environment and Society, Springer International Publishing, Cham, 2022, pp. 256–261.

[16] M. Grandini, E. Bagli, G. Visani, Metrics for multi-class classification: an overview, arXiv preprint arXiv:2008.05756 (2020).

[17] M. A. Khanesar, M. Teshnehlab, M. A. Shoorehdeli, A novel binary particle swarm optimization, in: 2007 Mediterranean Conference on Control & Automation, 2007, pp. 1–6. doi:10.1109/MED.2007.4433821.

A. Online Resources

Public link for downloading the acquired dataset.