

# A One-Shot Learning Framework to Model Process Systems

Sin Yong Teng<sup>a,\*</sup>, Vítězslav Mášá<sup>a</sup>, Hon Loong Lam<sup>b</sup>, Petr Stehlík<sup>a</sup>

<sup>a</sup>Brno University of Technology, Institute of Process Engineering & NETME Centre, Technická 2896/2, 616 69 Brno, Czech Republic

<sup>b</sup>Department of Chemical and Environmental Engineering, University of Nottingham Malaysia Campus, Jalan Broga, 43500 Semenyih, Selangor, Malaysia  
 Sin.Yong.Teng@vut.cz

In the era of Big Data, the utilization of data-driven analytics for process engineering systems is rising exponentially. The abundance of data from industrial sensors and various documentation logs have served as a strong basis for such analysis. Nevertheless, there are some critical data in an industry that simply rare and uncommon due to certain processing constraints or confidentiality. Such constraints may include economic costs for data acquisition, the complexity for data collection, the needs for qualified personnel and many other unforeseeable problems. Due to conventional data-driven approach requiring a large volume of data, such rare but critical data cannot be properly utilized. For this aspect, we proposed a one-shot learning framework to model process systems. The novel framework utilizes prior knowledge from multi-sourced data to learn the conditional relationships of critical variables within the process. By utilizing prior generic knowledge of the system, one-shot learning can provide a better representation of the prediction space when acting as a data-driven black-box model. A combined heat and power (CHP) system is used as the case study for one-shot learning modelling which a mean squared error of 0.00616 was achieved. The efficient use of data within this framework is expected to be beneficial when modelling under high-priority and low data availability.

## 1. Introduction

In advancing sustainable and efficient processing and manufacturing, the needs of data-driven engineering analytics is essential for the transition into the Industry 4.0. The data acquisition, modelling, simulation and optimization of processing systems within small and medium (SME) industries (Máša et al., 2018) are essential for companies to move towards a digitalized future. The use of data for industrial, manufacturing and business analytics has been consolidated to provide data-enabled growth (Ritter and Pedersen, 2020). Brynjolfsson and McElheran (2016) discussed that US manufacturing is transitioning towards a data-driven paradigm for gain productivity in managerial decision-making, tracking performance and communicating of the production process. The concept of data-driven smart manufacturing (Tao et al., 2018) has accelerated the global transition of manufacturing lifecycles towards an age of big data high, giving potential to improve manufacturing performances, process understanding and industrial management. Computational intelligence and process optimization have high potential in improving product quality and efficiency (Yin et al., 2020).

The challenge of implementing data-driven analytics in the industry is on the difficulty of obtaining reliable data sources (Máša et al., 2018). The identification and collection of reliable data within manufacturing systems is a great challenge for the actual implementation of process simulation models within manufacturing industries. Based on the author's industrial experience, carrying out optimization on such simulation models are often tricky in which there is a great dilemma on whether to: (i) optimize the model to be more accurate or (ii) optimize the solution to give better objectives under a controlled number of samples. In some aspect, this concept is related to the well-known exploration-exploitation dilemma (Berger-Tal et al., 2014) in machine learning and optimization. In such sample-critical cases, Bayesian methods are often used to improve sampling efficiency (Baheri and Vermillion, 2017) for "expensive-to-evaluate" problems. The sampling efficiency in such methods is improved by relying on both prior and posterior, gaining statistical significance from prior data (Ghosh and

Dunson, 2009). Deep learning approaches can also be used to effectively process important industrial data for modelling and optimization (Teng et al., 2019c).

Despite efficient sample utilization, many manufacturing facilities face data collection difficulties related to organisation authority, machine design, information transfer methods, analytical instruments, and cost of measurement. For example, obtaining new operational data for an oil refinery was expensive and difficult (Teng et al., 2019a). Industries such as co-generation plants (Leong et al., 2019) have many organisational and data confidentiality challenges in providing real-time data. While other energy management systems for SME generally has inadequate data-acquisition systems (Máša et al., 2016). Yong et al. (2016) also demonstrated that data reconciliation is essential for successful total site integration projects. While installing new data acquisition systems such as microcontrollers (Zhang and Chen, 2008) and carrying out carefully designed experiments in laboratories (Teng et al., 2019b) can provide high-quality data, it is generally expensive to do as such. Some processing equipment requires measurement devices that are expensive to operate (Hamacher et al., 2003) and manufacturing companies often turn them on only when required. This situation gives rise to many SME companies with only one or a few critical operational data, making process modelling difficult.

One-shot learning is a field of machine learning which uses one or a few samples (sometimes known as few-shot learning) to carry out inference instead of using hundreds or thousands of samples. This concept already existed in the 1990s (Yip and Sussman, 1997), while works from Fei-Fei et al. (2006) and Miller et al. (2000) popularized the use for object detection in images. Throughout the years, the applications of one-shot learning have shown successes for face verification (Guo et al., 2011), representing human gesture (Yang Yang et al., 2013), and other mobile authentication applications. One of the most successful implementations of by using a twin neural network, which is commonly called Siamese neural network to learn the similarity of data within the prior dataset (Koch et al., 2015). The strategy of using a Siamese neural network was particularly successful even for difficult tasks such as dynamic object tracking (Guo et al., 2017) and sentence plagiarism checking (Mueller and Thyagarajan, 2016).

As such, conceptually even one or a few samples can be effectively used to model process systems by one-shot learning techniques. This work presents a novel framework to model a process system with data acquisition problems using only one or a few samples via one-shot learning techniques. The novelty of this work is that the one-shot learning technique is adapted for process manufacturing and industrial data as an alternative to the conventional field of image and sequence classification.

## 2. Method and conceptual framework

The concept of one-shot learning framework is to learn the representation of the process model from its performance database (see Figure 1). This performance database can be obtained from multiple manufacturers that provide a similar type of unit or even similar operating units from other facilities. The most important constraint is that the sampled unit must have the same functionality as all the units in the database, but units in the database can be of a different design model. This preserves the representation of the unit functionality. The expected result from this framework is to obtain a relatively accurate data-driven model of the desired performance of the sampled unit by lending knowledge from the aforementioned database.

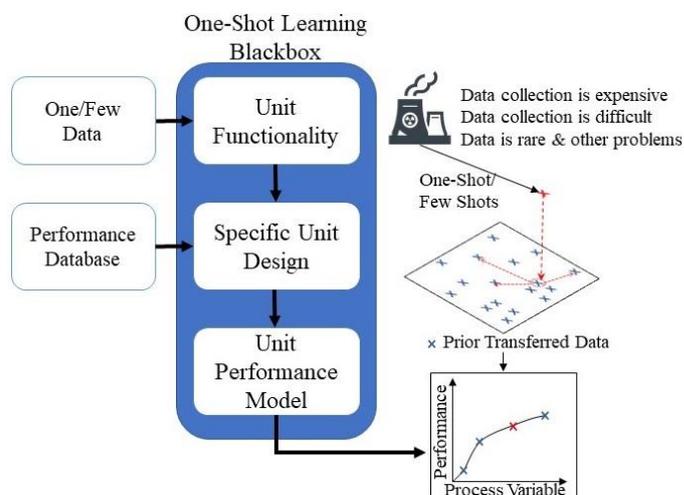


Figure 1: Conceptual diagram of one-shot learning for process modelling

Referring to Figure 2., the first step to start the one-shot learning workflow is to confirm that a data-driven modelling approach is required. It is required to verify if the use of one-shot learning is suitable by evaluating data availability. If there is flexibility for data sampling or data quantity, the use of an alternative data-driven method such as principal component analysis statistical process optimization (Teng et al., 2019a), Monte-Carlo simulation (Ngan et al., 2020), or adaptive analytical approaches (Leong et al., 2019) can be deployed. For the next step, it is required to identify the process unit functionality and obtain one or a few samples (which includes the performance variable of interest). Based on this unit functionality, prior data from other process units with similar functionality should be obtained from sources such as multiple manufacturer's databases, other facilities, and commercial software. As an example regarding "similar functionality", during the modelling of a 1,2-pass heat exchanger, one could use the data from a 2,2,-pass heat exchanger in similar conditions as prior data due to its similar functionality within the process. This means that the data from units of a different design but fulfils the same processing purpose can be used as a knowledge basis for this one-shot learning framework. The knowledge basis should have a statistically significant amount of data to give a good representation of the unit functionality.

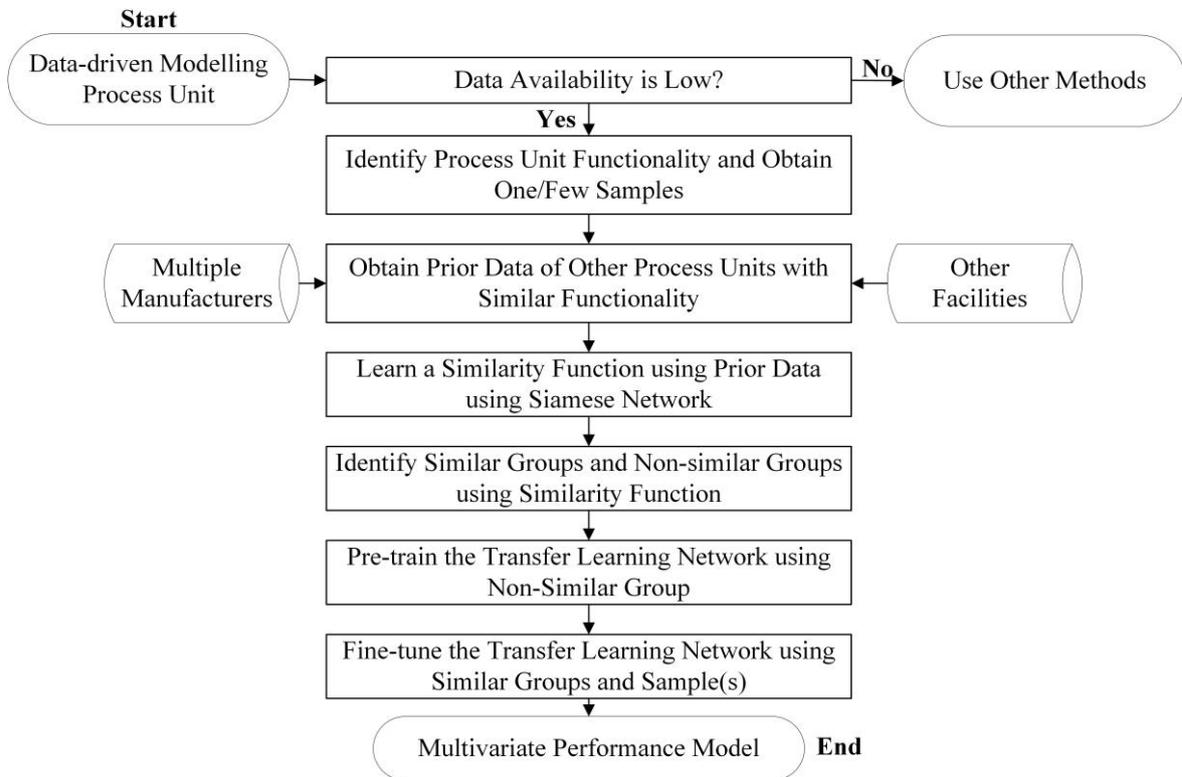


Figure 2: Workflow of the novel one-shot learning framework for process system modelling

From this prior data, a Siamese network is used to learn the similarity function between the previous grouped data. In this paper, the contrastive loss (Hadsell et al., 2006) is used as the similarity function (see Figure 3). The similar groups with regards to the one/few samples (group encoding is represented as  $C$ ) are then selected using the equation:

$$C = y(\operatorname{argmin} f_c(x, x_s)) \quad \forall x \in X \quad (1)$$

Where  $x$  is the single data from the prior,  $X$  is the full prior dataset,  $x_s$  is the one-shot sample, and  $y$  is the group classification of the data. Other data with a different group than the similar group are all separated into a non-similar group. Next, we propose the use of a multi-layer perceptron (MLP) with distinct losses for transfer learning. During the pre-train, the non-similar group are split at an 80:10:10 training, validation and testing ratio with a modified Pearson's correlation coefficient as the loss to learn the general shape of the performance space. Next, the similar group which contains the one/few samples are used to finetune the features of the pre-trained neural network by using a few extension layers. The loss function for the fine-tuning step is set to be the mean squared error function to give an aggressive fine-tuning result, giving high accuracy. The performance characteristics of the processing unit can be predicted by the transfer learned network with the knowledge represented from the space of prior data of process unit models of similar functionality.

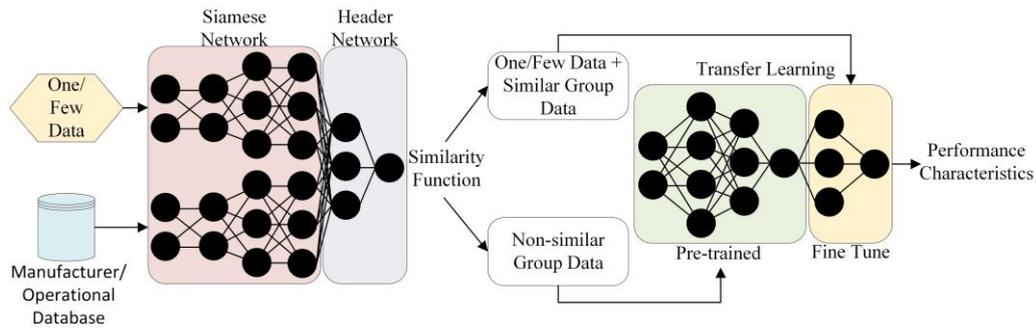


Figure 3: Siamese network and transfer learned MLP for one-shot learning in process system modelling

### 3. Case study problem

The modelling of a combined heat and power (CHP) unit (ECOMAX 44 NGS 1.1 HW model) is studied. CHP units commonly operate at a fix steady-state point, and changing the operational point is generally costly in terms of process economics. One sample data was collected from this CHP model. Prior data were prepared using an in-house collection of CHP performance database consisting of 613 datasets from 64 units. The sampled unit is also cross-checked with the database, and the database does not contain data from the specific CHP unit model. For learning the similarity function, all possible 15 variables were input, such as total efficiency, carbon emission, and power generation. Due to modelling requirement, the performance characteristics only require three inputs to predict the CHP’s thermal efficiency and overall efficiency. These three variables are power utilization percentage, temperature and fuel consumption. Finally, for validation purposes, an extra 7 data samples were obtained from the studied CHP from its operational history.

### 4. Results

Using the single data from the studied CHP unit, the Siamese network was able to learn the similarity function and allowed for a transfer learning MLP network learn the representation of unit functionality and fine-tuned on similar data groups. By comparing the predicted output using the one-shot learning framework and the seven ground truth data (as well as the 1 data sample), it is possible to achieve a very promising mean absolute error (MAE) of 0.02, mean squared error (MSE) of 0.000616 and  $R^2$  of 0.9992. These results demonstrate that the framework was able to model the CHP unit only with 1 sample with acceptable error (See Figure 4).

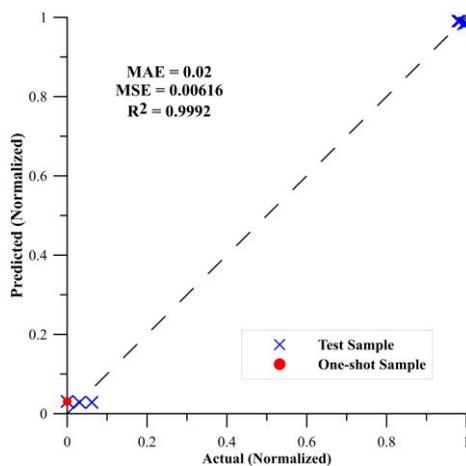


Figure 4: Predicted against the actual plot of test samples and one-shot sample with their overall error.

For further analysis, the surface plot of the generated performance model was plotted in Figure 5. It is identified that power utilization in percentage was the main factor of the efficiencies in the CHP model. Fuel consumption and temperature also slightly affects CHP efficiencies. It is also observed that increase in power utilization gives a steep increase in thermal efficiency, however creates a plateau for the overall efficiency. This implies that at over 80 % power utilization, the total energy output is approximately the same. The ratio of power and heat

energy gradually increases when further increasing the power utilization. The temperature and fuel flowrate are providing small effects on the thermal and overall efficiencies.

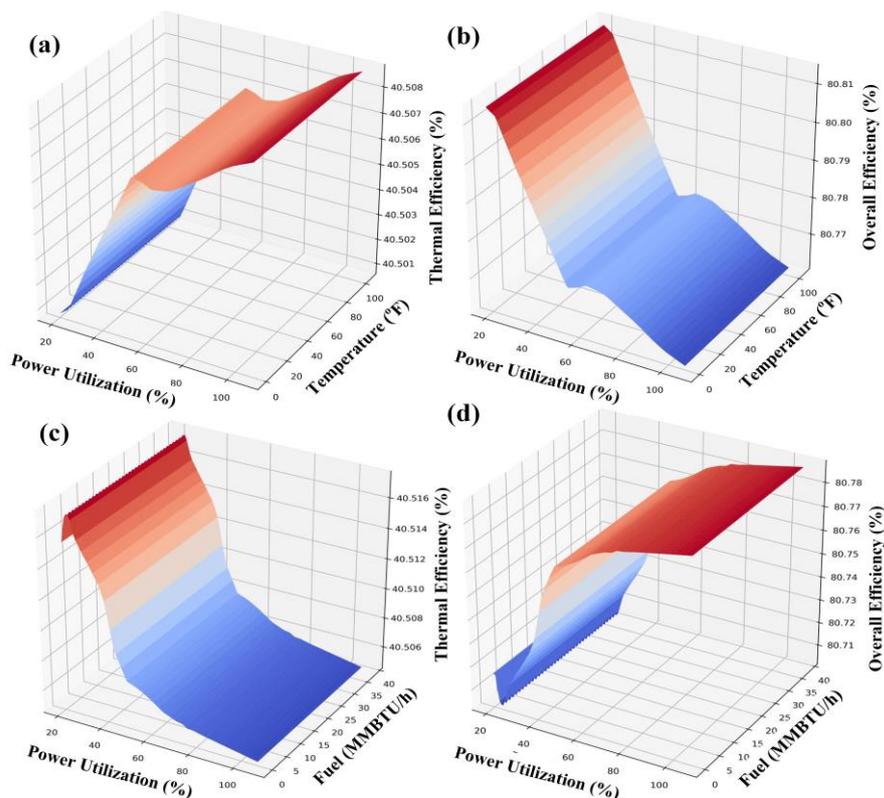


Figure 5: Surface plot of (a) Power utilization (PU), temperature (T) and thermal efficiency (TE) (b) PU, T, Overall Efficiency (OE) (c) PU, Fuel, TE, (d) PU, Fuel, OE

## 5. Conclusions

This paper proposes the use of novel one-shot learning for the application of modelling process systems under low data availability problems. A novel two-step approach is proposed. First, a Siamese network is introduced to learn the similarity function within prior datasets that are from different units but share the same functionality with the studied unit. Next, these datasets are categorized as a non-similar group and similar groups to be sequentially transferred learn by a multi-layer perceptron (MLP) neural network. Using a few extra samples of the studied system that were never shown to the networks, the model was tested. A convincing mean squared error of 0.00616 was achieved only from the original one sample from the studied combined heat and power (CHP) unit. A smooth prediction space was observed from the surface plot, showing smooth continuity in the model. This demonstrates that this approach is stable and able to model process units with high accuracy and only a single sample similarly. Additionally, the one-shot learning framework can be applied to any other units. Our future work will focus on the multi-objective optimization of a one-shot learned process model unit.

## Acknowledgements

The research leading to these results has received funding from the Ministry of Education, Youth and Sports of the Czech Republic under OP RDE grant number CZ.02.1.01/0.0/0.0/16\_026/0008413 "Strategic Partnership for Environmental Technologies and Energy Production".

## References

- Baheri A., Vermillion C., 2017, Altitude optimization of Airborne Wind Energy systems: A Bayesian Optimization approach, in: 2017 American Control Conference (ACC). IEEE, Seattle, WA, 1365–1370 DOI:10.23919/ACC.2017.7963143.
- Berger-Tal O., Nathan J., Meron E., Saltz D., 2014, The Exploration-Exploitation Dilemma: A Multidisciplinary Framework. PLoS One, 9, e95693.

- Brynjolfsson E., McElheran K., 2016, Data in Action: Data-Driven Decision Making in U.S. Manufacturing. SSRN Electron. J., 1–55, DOI:10.2139/ssrn.2722502.
- Ghosh J., Dunson D.B., 2009, Default Prior Distributions and Efficient Posterior Computation in Bayesian Factor Analysis. *J. Comput. Graph. Stat.*, 18, 306–320.
- Guo H., Schwartz W.R., Davis L.S., 2011, Face verification using large feature sets and one shot similarity, in: 2011 International Joint Conference on Biometrics, IJCB 2011, DOI:10.1109/IJCB.2011.6117498.
- Guo Q., Feng W., Zhou C., Huang R., Wan L., Wang S., 2017, Learning Dynamic Siamese Network for Visual Object Tracking, in: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, 1781–1789, DOI:10.1109/ICCV.2017.196.
- Hadsell R., Chopra S., LeCun Y., 2006, Dimensionality Reduction by Learning an Invariant Mapping, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2, 1735–1742.
- Hamacher T., Niess J., Schulze Lammers P., Diekmann B., Boeker P., 2003, Online measurement of odorous gases close to the odour threshold with a QMB sensor system with an integrated preconcentration unit. *Sensors Actuators B Chem.*, 95, 39–45.
- Koch G., Zemel R., Salakhutdinov R., 2015, Siamese Neural Networks for One Shot Image Learning, in: ICML Deep Learning Workshop. JMLR: W&CP, Lille, France, 1–8.
- Leong W.D., Teng S.Y., How B.S., Ngan S.L., Lam H.L., Tan C.P., Ponnambalam S.G., 2019, Adaptive Analytical Approach to Lean and Green Operations. *J. Clean. Prod.*, 235, 190–209.
- Li F.F., Fergus R., Perona P., 2006, One-Shot Learning of Object Categories. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28, 594–611.
- Máša V., Stehlík P., Touš M., Vondra M., 2018, Key Pillars of Successful Energy Saving Projects in Small and Medium Industrial Enterprises. *Energy*, 158, 293–304.
- Máša V., Touš M., Pavlas M., 2016, Using a Utility System Grey-Box Model as a Support Tool for Progressive Energy Management and Automation of Buildings. *Clean Technol. Environ. Policy*, 18, 195–208.
- Miller M.G., Matsakis N.E., Viola P.A., 2000, Learning from one Example through Shared Densities on Transforms, in: Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662). IEEE Comput. Soc, 464–471, DOI:10.1109/CVPR.2000.855856.
- Mueller J., Thyagarajan A., 2016, Siamese Recurrent Architectures for Learning Sentence Similarity, In: 30th AAAI Conference on Artificial Intelligence, AAAI 2016. AAAI Press, Phoenix, Arizona USA, 2786–2792.
- Ngan S.L., How B.S., Teng S.Y., Leong W.D., Loy A.C.M., Yatim P., Promentilla M.A.B., Lam H.L., 2020, A Hybrid Approach to Prioritize Risk Mitigation Strategies for Biomass Polygeneration Systems. *Renew. Sustain. Energy Rev.*, 121, 109679.
- Ritter T., Pedersen C.L., 2020, Digitization Capability and the Digitalization of Business Models in Business-to-Business Firms: Past, Present, and Future. *Ind. Mark. Manag.*, 86, 180–190, DOI:10.1016/j.indmarman.2019.11.019.
- Tao F., Qi Q., Liu A., Kusiak A., 2018, Data-Driven Smart Manufacturing., *J. Manuf. Syst.*, 48, 157–16.
- Teng S.Y., How B.S., Leong W.D., Teoh J.H., Siang Cheah A.C., Motavasel Z., Lam H.L., 2019a, Principal Component Analysis-Aided Statistical Process Optimisation (PASPO) for Process Improvement in Industrial Refineries. *J. Clean. Prod.*, 225, 359–375.
- Teng S.Y., Loy A.C.M., Leong W.D., How B.S., Chin B.L.F., Máša V., 2019b, Catalytic Thermal Degradation of *Chlorella Vulgaris*: Evolving Deep Neural Networks for Optimization. *Bioresour. Technol.*, 292, 121971.
- Teng S.Y., Máša V., Stehlík P., Lam H.L., 2019c, Deep Learning Approach for Industrial Process Improvement. *Chem. Eng. Trans.*, 76, 487–492.
- Yang Y., Saleemi I., Shah M., 2013, Discovering Motion Primitives for Unsupervised Grouping and One-Shot Learning of Human Actions, Gestures, and Expressions. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35, 1635–1648.
- Yin X., Niu Z., He Z., Li Z., Lee D., 2020, An Integrated Computational Intelligence Technique based Operating Parameters Optimization Scheme for Quality Improvement Oriented Process-Manufacturing System. *Comput. Ind. Eng.*, 140, 106284, DOI:10.1016/j.cie.2020.106284.
- Yip K., Sussman G.J., 1997, Sparse Representations for Fast, One-Shot Learning, In: Proceedings of the National Conference on Artificial Intelligence. Providence, Rhode Island, 521–527.
- Yong J.Y., Nemet A., Varbanov P.S., Kravanja Z., Klemeš J.J., 2016, Data Reconciliation for Total Site Integration. *Chem. Eng. Trans.*, 52, 1045–1050.
- Zhang J.Z., Chen J.C., 2008, Tool Condition Monitoring in an End-Milling Operation based on the Vibration Signal Collected through a Microcontroller-Based Data Acquisition System. *Int. J. Adv. Manuf. Technol.*, 39, 118–128.