

Rozkład wielkości kryształów nadzorowany przez agenta uczenia ze wzmocnieniem z wykorzystaniem czujnika tomografii hybrydowej w procesie krystalizacji

Streszczenie. W pracy opisano wykorzystanie uczenia ze wzmocnieniem w modelu Simulink do symulacji sterowania procesem krystalizacji. Przedstawiono również stanowisko akwizycji danych tomograficznych służące do sterowania rzeczywistym procesem krystalizacji zarządzanym agentem uczenia ze wzmocnieniem. Rekonstrukcja obrazu z systemu tomograficznego umożliwia agentowi uczenia ze wzmocnieniem uzyskanie dodatkowych danych w czasie rzeczywistym o stanie środowiska, co z kolei pozwoli kontrolerowi na prowadzenie procesu.

Abstract. This paper describes the use of reinforcement learning in the Simulink model to simulate the control of the crystallization process. A tomographic data acquisition workstation for controlling the actual crystallization process managed by a reinforcement learning agent is also presented. Reconstructing the image from the tomographic system allows the reinforcement learning agent to obtain additional real-time data about the state of the environment, which in turn will allow the controller to guide the process. (*Crystal size distribution supervised by a reinforcement learning agent using a tomography sensor during the crystallization process*)

Słowa kluczowe: Uczenie ze wzmocnieniem, krystalizacja, Simulink, tomografia.

Keywords: Reinforcement learning, crystallization, Simulink, tomography.

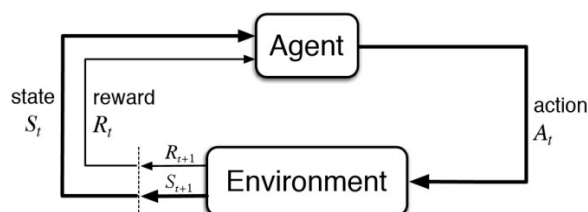
Wstęp

Krystalizacja jest procesem oczyszczania i separacji stosowanym w wielu operacjach w produkcji żywności, w przemyśle chemicznym, w przemyśle farmaceutycznym, w przemyśle nawozowym itp. Większość produktów stałych zawiera krystalizację w swoim procesie produkcyjnym, a zatem optymalizacja procesu krystalizacji może mieć znaczący wpływ na właściwości dużej ilości produktów końcowych. Istnieje wiele metod przesycań wywołującego krystalizację, takich jak chłodzenie, odparowanie, dodawanie antyrozpuszczalnika, przeprowadzanie reakcji chemicznych (np. zmiana pH) itp. W każdym z tych przypadków celem jest przesylenie roztworu do poziomu wywołującego zarodkowanie i rozrost kryształów. Kontrola procesu krystalizacji jest kłopotliwa, nieliniowe warunki mieszania, niekontrolowana szybkość dodawania rozpuszczalnika, opóźnienia w kontroli i sterowania temperaturą, nieoczekiwane przemiany, szybkość krystalizacji w czasie, zmiany rozmiaru kryształów i trudności w skalowaniu procesu z wielkości laboratoryjnych do wielkości przemysłowych powodują niekorzystne efekty. Najczęściej parametrami nadzorowanymi podczas procesu krystalizacji są koncentracja cząstek i rozkład wielkości kryształów, określane na podstawie zarodkowania i wzrostu kryształów. Celem precyzyjnej kontroli i sterowania procesem jest uzyskanie oczekiwanych parametrów produktu końcowego [1-7]. W niniejszej pracy zastosowano symulację krystalizacji wsadowej z kontrolą chłodzenia w celu osiągnięcia przesylenia. Sterownik chłodzenia został zdefiniowany przy użyciu agenta uczenia ze wzmocnieniem [8] w oprogramowaniu Matlab i Simulink [9]. W części drugiej przedstawiono stanowisko pomiarowe wykorzystujące jednocześnie tomografię impedancyjną, pojemnościową oraz ultradźwiękową do pozyskania danych rzeczywistego procesu krystalizacji metodą przesylenia wywołanego dodawaniem rozpuszczalnika, gdzie sterowaniu podlega szybkość podawanego rozpuszczalnika oraz szybkość mieszania roztworu mieszadłem mechanicznym. Rekonstrukcje tomograficzne w tym przypadku zapewniają dostarczanie danych z procesu tym samym są wykorzystane w sprzężeniu zwrotnym zamkniętej pętli kontrolera. Prace związane ze sterowaniem procesem krystalizacji za pomocą przetwarzania obrazu

były już prowadzone [10-12], jednak nie wykorzystywały one rekonstrukcji tomograficznych a obrazy mikroskopowe.

Uczenie ze wzmocnieniem

Agent uczenia ze wzmocnieniem do sterowania procesem wykorzystuje doświadczenie zdobyte podczas uczenia z symulacji lub bezpośrednio z procesu krystalizacji. W przeciwieństwie do tradycyjnych metod sterowania przebiegiem procesu za pomocą równania bilansu populacji (population balance equation) [13] kontroler oparty na uczeniu ze wzmocnieniem nie musi rozwiązywać w czasie rzeczywistym problemu optymalizacyjnego polegającego na doborze parametrów sterowania prowadzącym do uzyskania oczekiwanego rozkładu wielkości kryształów i ich koncentracji, gdyż sieć neuronowa taką optymalizację przeszła na etapie uczenia się reakcji na stan środowiska.



Rys. 1. Schemat procesu uczenia się ze wzmocnieniem

Podstawowy schemat uczenia ze wzmocnieniem przedstawiony został na rysunku 1. Schemat zawiera dwa bloki: Agent i Środowisko. Agent wchodzi w interakcję ze środowiskiem za pomocą predefiniowanych akcji A. W pierwszym kroku Agent obserwuje stan S_t i wartość nagrody R_t , a następnie agent wykonuje akcję A_t z zestawu akcji A, która zmienia stan środowiska z S_t na S_{t+1} zgodnie z poniższym rozkładem prawdopodobieństwa.

$$p(s', r|s, a) = \Pr \{S_t = s', R_t = r | S_{t-1} = s, A_{t-1} = a\}$$

gdzie

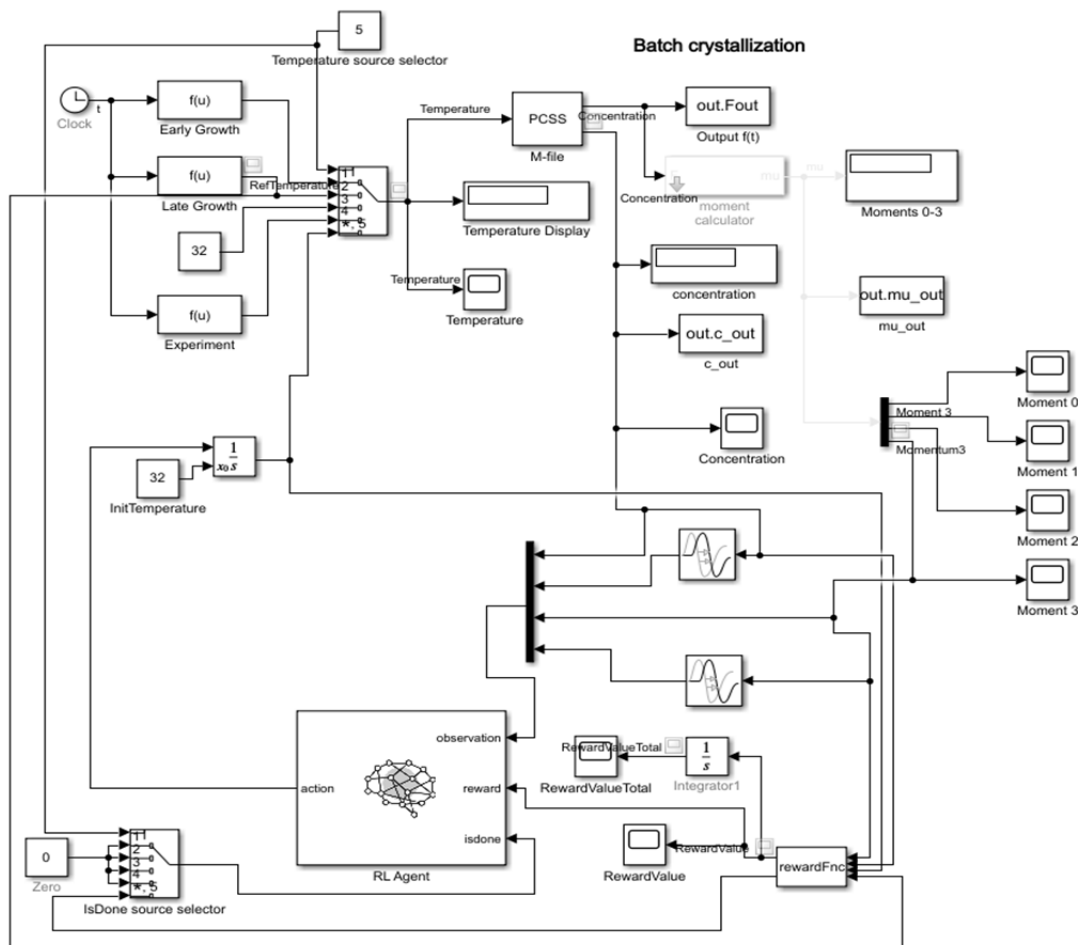
$$\sum_{s' \in S} \sum_{r \in R} p(s', r|s, a) = 1$$

Dla każdego $s \in S, a \in A$.

Stan S_t opisuje aktualną sytuację w środowisku (obserwowane właściwości środowiska), a funkcja nagrody zwraca wartość R_t , czyli metrykę obliczaną dla każdego kroku skorelowaną z aktualnym stanem środowiska. Algorytm działa w pętli, generując kolejną akcję A_{t+1} i odczytując wyniki: stan S_{t+1} i nagrodę R_{t+1} . Celem uczenia przez wzmocnienie jest nauczenie agenta interakcji ze środowiskiem w taki sposób, aby w wyniku działania procesu doprowadzić środowisko do możliwie najlepszego stanu [14, 15]. W omawianym przypadku środowiskiem jest proces krystalizacji, zbiór akcji A zawiera listę możliwych zmian temperatury dla każdego kroku czasowego procesu, stan S_{t+1} zawiera wartość koncentracji kryształów oraz trzeci moment rozkładu wielkości kryształów dla dwóch ostatnich kroków t i $t+1$, a nagroda R_{t+1} zwraca wartość obliczoną przez funkcję nagrody, łączącą różnicę aktualnej i

oczekiwanej koncentracji oraz różnicę trzeciego momentu rozkładu wielkości cząstek dla aktualnej i oczekiwanej wartości. Zestaw kroków wykorzystywanych przez agenta do sterowania procesem nazywany jest strategią. Strategia budowana jest na podstawie zestawu rekordów interakcji i nagród dla każdego ze stanów. Taki zestaw nazywany jest funkcją wartości. Funkcje wartości są zazwyczaj tablicami z danymi lub też funkcjami liniowymi. W związku z ograniczeniami takich rozwiązań oraz coraz powszechniejszym wykorzystaniem sztucznej inteligencji, jako funkcje wartości zaczęto używać sieci neuronowe, są one w stanie uczyć się znaczenie większej ilości różnorodnych problemów, niż odpowiedniki funkcji liniowych, przy użyciu algorytmów gradientowych, minimalizując rozbieżności pomiędzy oczekiwanymi wynikami a wynikami przewidywanymi.

Rys.2. Model SIMULINK krystalizacji wsadowej z nadzorem agenta Reinforcement Learning



Algorytmy głębokiego uczenia ze wzmocnieniem odnoszą sukcesy przewyższając jakość wyników sterowania w odniesieniu do sterowania przez człowieka. Przykładem może być gra AlphaGo [16] czy gry Atari [17].

Model procesu

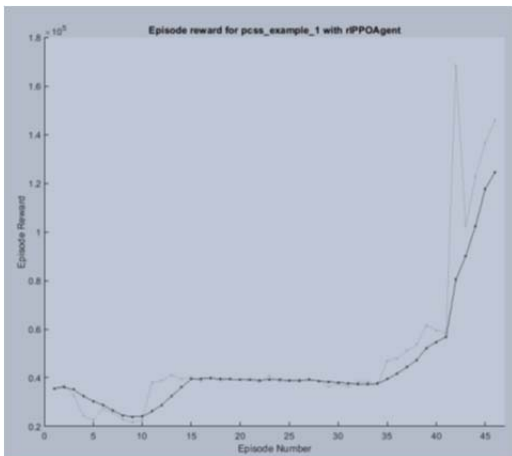
Model sterowania krystalizacją wsadową oparty na uczeniu ze wzmocnieniem skonfigurowano zgodnie z rysunkiem 2 w systemie symulacji procesu w czasie Simulink. Jest to rozszerzona wersja schematu blokowego Simulinka dla procesu krystalizacji KNO₃ (azotanu potasu) wykorzystującego blok PCSS [18]. Głównym elementem regulatora jest agent uczenia ze wzmocnieniem (ang. Reinforcement Learning Agent). Agent zasilany jest danymi z symulacji procesu generowanymi przez moduł PCSS na

podstawie temperatury czynnika, wartością nagrody oraz flagą zakończenia procesu.

Table 1. Struktura sieci neuronowej aktora agenta PPO

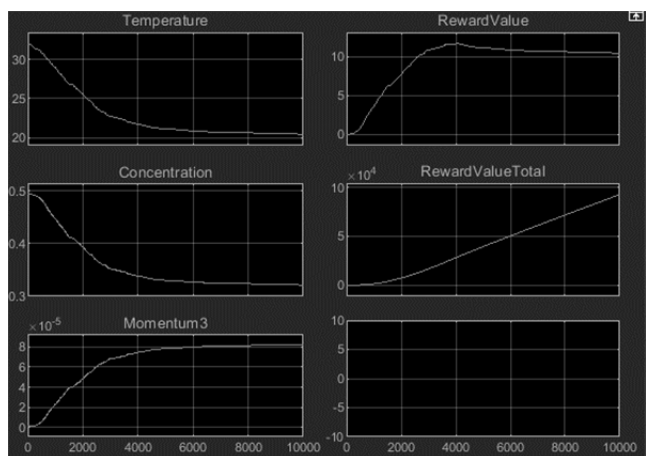
Name	Type	Activations	Learnable Prope...
1 input_1 4 features	Feature input	4(C) × 1(B)	-
2 fc_1 256 fully connected layer	Fully Connected	256(C) × 1(B)	Weights 256 × 4 Bias 256 × 1
3 relu_body ReLU	ReLU	256(C) × 1(B)	-
4 fc_body 256 fully connected layer	Fully Connected	256(C) × 1(B)	Weights 256 × 256 Bias 256 × 1
5 body_output ReLU	ReLU	256(C) × 1(B)	-
6 fc_action 2 fully connected layer	Fully Connected	2(C) × 1(B)	Weights 2 × 256 Bias 2 × 1
7 output softmax	Softmax	2(C) × 1(B)	-

Wektor danych przekazywanych z PCSS do agenta zawiera wartości stężenia i trzeciego momentu rozkładu wielkości kryształów zarówno dla bieżącego, jak i poprzedniego kroku. Pozostałe parametry są zwracane przez funkcję nagrody. Jako kalkulator nagrody wykorzystano funkcję Matlab S-Function, w której wartość nagrody jest obliczana na podstawie koncentracji kryształów, wartości trzeciego momentu rozkładu wielkości kryształów i temperatury roztworu. W systemie wykorzystano agenta PPO (ang. Proximal Policy Optimization) [19] który naprzemiennie pobiera dane ze środowiska i optymalizuje funkcję celu zastępczego wykorzystując algorytmy gradientowe. Cechą charakterystyczną algorytmu PPO jest fakt że każda optymalizacja realizowana jest na podstawie wielu epok.



Rys.3. Nagroda w kolejnych epokach uczenia sztucznej sieci neuronowej

Strukturę sztucznej sieci neuronowej dla zastosowanego algorytmu PPO przedstawiono w tabeli 1. Wyniki uczenia agenta wykazują, że agent potrafi sterować procesem w sposób prowadzący do uzyskania zarówno wysokich nagród w kolejnych epokach (rysunek 3) jak również oczekiwanych wartości koncentracji i rozkładu wielkości kryształów, a dodatkowo ma tendencję do uzyskiwania oczekiwanych rezultatów w możliwie krótkim czasie.



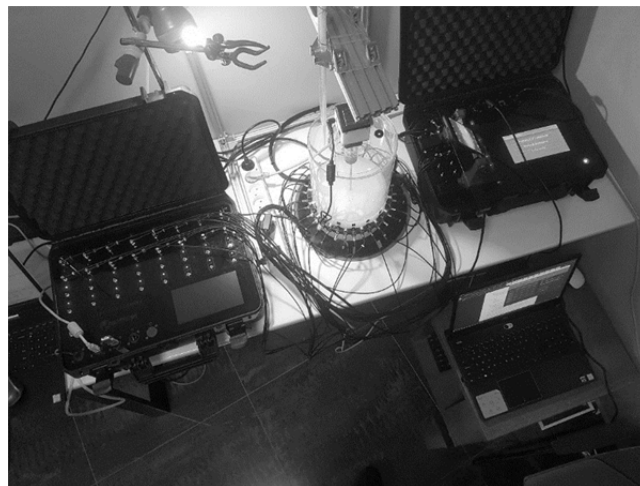
Rys.4. Parametry procesu sterowanego wytrenowanym agentem uczenia ze wzmocnieniem

Prezentowane wykresy (rysunek 4) przedstawiają przebieg temperatury zaproponowany przez agenta uczenia ze wzmocnieniem oraz stężenie cząstek i trzeci moment rozkładu wielkości kryształów, a także wartość nagrody dla

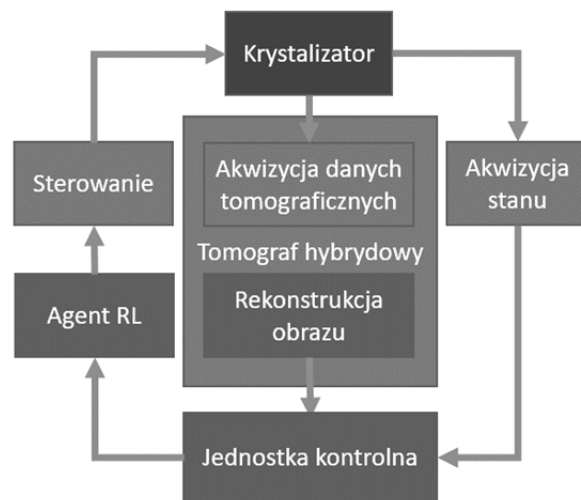
każdego kroku czasowego i wartość nagrody w całym procesie (10000 kroków).

Hybrydowy system tomograficzny

Prowadzone są nowatorskie badania z wykorzystaniem systemów tomografii elektrycznej (EIT, ECT) oraz tomografii ultradźwiękowej (UST) do monitorowania procesów produkcyjnych, takich jak fermentacja, krystalizacja i podobne [20-29]. W niniejszej pracy przeprowadzono pomiary procesu krystalizacji z wykorzystaniem hybrydowego systemu tomograficznego (rysunek 5). Diagram przepływu danych w procesie sterowania i akwizycji przedstawiono na rysunku 6.



Rys.5. Hybrydowy tomograf EIT i ECT, Hybrydowy Sensor Tomograficzny, Tomograf UST



Rys.6. Diagram przepływu danych, w procesie akwizycji i sterowania obiektem

Blok granatowy symbolizuje obiekt kontrolowany, w tym przypadku kryształizator. Bloki zielone reprezentują część sprzętową systemu, odpowiednio systemy wykonawcze jak pompy, mieszacze, zawory, systemy akwizycji danych tj. sensory i sondy tomograficzne oraz czujniki wielkości fizycznych np. temperatury. Bloki pomarańczowe symbolizują część programową systemu, w tym moduły rekonstruujące obraz z systemów tomograficznych, jednostkę kontrolną której zadaniem jest przygotowanie i zarządzanie przepływem danych do agenta uczenia ze wzmocnieniem, i ostatecznie agent RL, którego zadaniem jest optymalne sterowanie procesem produkcji kryształów.

Podsumowanie

Rekonstrukcja obrazu z systemu tomograficznego umożliwi agentowi uczenia ze wzmocnieniem uzyskanie dodatkowych danych w czasie rzeczywistym o stanie środowiska, co z kolei pozwoli kontrolerowi na prowadzenie procesu w sposób bardziej optymalny tym samym uzyskując lepszą jakość produktu. Przyszłe prace będą koncentrować się na nauczaniu agenta uczenia ze wzmocnieniem sterowania procesem krystalizacji z wykorzystaniem pomiarów tomograficznych i rekonstruowanych obrazów.

Autorzy: mgr inż. Konrad Niderla, Centrum Badawczo Rozwojowe Netrix S.A. ul. Związkowa 26, 20-148 Lublin, Wyższa Szkoła Ekonomii i Innowacji, ul. Projektowa 4, 20-209 Lublin, E-mail: konrad.niderla@netrix.com.pl;
dr inż. Grzegorz Kłosowski, Politechnika Lubelska, ul. Nadbystrzycka 38, 20-618 Lublin, E-mail: g.klosowski@pollub.pl

LITERATURA

- [1] A. S.Myerson, *Handbook of Industrial Crystallization*, Elsevier Science & Technology Books, 2001.
- [2] X.Y.Woo, Z.K.Nagy, R.B.H.Tan, R.D.Braatz, *Adaptive Concentration Control of Cooling and Antisolvent Crystallization with Laser Backscattering Measurement*, *Crystal growth & design* 9 (2009), No.1 182-191.
- [3] M.Fujiwara, Z.K.Nagy, J.W.Chew, R.D.Braatz, *First-principles and direct design approaches for the control of pharmaceutical crystallization*, *Journal of process control*, 15 (2005), 493-504.
- [4] Z.K.Nagy, M.Fujiwara, R.D.Braatz, *Recent advances in the modelling and control of cooling and antisolvent crystallization of pharmaceuticals*, *IFAC Symposium on Dynamics and Control* 40 (2007), 29–38.
- [5] Z.K.Nagy, R.D.Braatz, *Advances and New Directions in Crystallization Control*, *Annual reviews Further* (2012).
- [6] Z.K.Nagy, E.Aamir, *Systematic design of supersaturation controlled crystallization processes for shaping the crystalsize distribution using an analytical estimator*, *Chemical Engineering Science*, 84 (2012), 656–670.
- [7] R.D Braatz, *Advanced control of crystallization processes*, *Annual reviews in control*, (2002) 87-99.
- [8] R.S.Sutton, A.G.Barto *Reinforcement Learning. An Introduction*, MIT Press Ltd (2018).
- [9] MATLAB, MathWorks, Inc., Natick, Massachusetts, United and States. [Online]. Available: <https://www.mathworks.com>.
- [10] V.Maneea, R.Barattib, J.A.Romagnolia *Learning to navigate a crystallization model with Deep Reinforcement Learning*, *Chemical Engineering Research and Design*, 178 (2022), 111-123
- [11] Manee, W.Zhu, J.A.Romagnoli *A Deep Learning Image-Based Sensor for Real-Time Crystal Size Distribution Characterization*, *Industrial & Engineering Chemistry Research* 58 (2019), No. 51, 23175-23186
- [12] Z.Gao, Y.Wu,Y.Bao, J.Gong, J.Wang, S.Rohani *Image Analysis for In-line Measurement of Multidimensional Size, Shape, and Polymorphic Transformation of L-Glutamic Acid Using Deep Learning-Based Image Segmentation and Classification*, *Crystal Growth & Design*, 18 (2018), No. 8, 4275-4281.
- [13] S.Mostafa Nowee, A.Abbas†, J.A.Romagnoli *Model-Based Optimal Strategies for Controlling Particle Size in Antisolvent Crystallization Operations* *Cryst. Growth Des.*, 8 (2008), 8, 2698–2706.
- [14] M. Sewak, *Deep Reinforcement Learning*, Springer, 2019.
- [15] T.P.Lillicrap, J.J.Hunt, A.Pritzel, N.Heess, T.Erez, Y.Tassa, D.Silver, D.Wierstra, *Continuos control with deep reinforcement learning*, arXiv, (2019)
- [16] D.Silver, A.Huang, C.J.Maddison, A.Guez, L.Sifre, G.Van Den Driessche, J.Schrittwieser, I.Antonoglou, V.Panneershelvam, M.Lanctot, S.Dieleman, D.Grewe,J.Nham, N.Kalchbrenner, I.Sutskever, T.Lillicrap, M.Leach, K.Kavukcuoglu, T.Graepel, D.Hassabis, *Masteringthe game of go with deep neural networks and tree search*. (2016) *Nature*.
- [17] V.Mnih, K.Kavukcuoglu, D.Silver, A.Graves, I.Antonoglou, D.Wierstra, M.Riedmiller *Playing Atari with Deep Reinforcement Learning*, arXiv (2013)
- [18] J. D. Ward and C.-C. Yu, *Population balance modeling in Simulink: PCSS*, *Computers and Chemical Engineering*, (2008), 2233-2242.
- [19] J.Schulman, F.Wolski, P.Dhariwal, A.Radford, O.Klimov, *Proximal Policy Optimization Algorithms*, (2017) arXiv
- [20] Rymarczyk T., Kłosowski G., Hoła A., Sikora J., Tchórzewski P., Skowron Ł., *Optimising the Use of Machine Learning Algorithms in Electrical Tomography of Building Walls: Pixel Oriented Ensemble Approach*, *Measurement*, 188 (2022), 110581.
- [21] Koulountzios P., Rymarczyk T., Soleimani M., *Ultrasonic Time-of-Flight Computed Tomography for Investigation of Batch Crystallisation Processes*, *Sensors*, 21 (2021), No. 2, 639.
- [22] Kłosowski G., Rymarczyk T., Niderla K., Rzemieniak M., Dmowski A., Maj M., *Comparison of Machine Learning Methods for Image Reconstruction Using the LSTM Classifier in Industrial Electrical Tomography*, *Energies* 2021, 14 (2021), No. 21, 7269.
- [23] Rymarczyk T., Król K. Kozłowski E., Wołowicz T., Cholewa-Wiktor M., Bednarczuk P., *Application of Electrical Tomography Imaging Using Machine Learning Methods for the Monitoring of Flood Embankments Leaks*, *Energies*, 14 (2021), No. 23, 8081.
- [24] Majerek D., Rymarczyk T., Wójcik D., Kozłowski E., Rzemieniak M., Gudowski J., Gauda K., *Machine Learning and Deterministic Approach to the Reflective Ultrasound Tomography*, *Energies*, 14 (2021), No. 22, 7549.
- [25] Kłosowski G., Rymarczyk T., Kania K., Świć A., Cieplak T., *Maintenance of industrial reactors supported by deep learning driven ultrasound tomography*, *Eksploracja i Niezawodność – Maintenance and Reliability*; 22 (2020), No 1, 138–147.
- [26] Gnaś, D., Adamkiewicz, P., *Indoor localization system using UWB*, *Informatyka, Automatyka, Pomiary W Gospodarce I Ochronie Środowiska*, 12 (2022), No. 1, 15-19.
- [27] Styła, M., Adamkiewicz, P., *Optimisation of commercial building management processes using user behaviour analysis systems supported by computational intelligence and RTI*, *Informatyka, Automatyka, Pomiary W Gospodarce I Ochronie Środowiska*, 12 (2022), No 1, 28-35.
- [28] Korzeniewska, E., Krawczyk, A., Mróz, J., Wszyńska, E., Zawiślak, R., *Applications of smart textiles in post-stroke rehabilitation*, *Sensors (Switzerland)*, 20 (2020), No. 8, 2370.
- [29] Sekulska-Nalewajko, J., Goćławski, J., Korzeniewska, E., *A method for the assessment of textile pilling tendency using optical coherence tomography*, *Sensors (Switzerland)*, 20 (2020), No. 13, 1–19, 3687.