ISSN 1331-7210 (Print)
ISSN 2718-322X (Online)

Engineering Power

BULLETIN OF THE CROATIAN ACADEMY OF ENGINEERING Vol. 18(4) 2023

EDITOR-IN-CHIEF'S WORD

Dear Readers,

As engineers, we stand on the brink of a transformative era, where unprecedented advancements in artificial intelligence (AI) are reshaping the foundation of our disciplines and redefining the way we approach challenges.

In this issue of Engineering Power, we are proud to showcase four thought-provoking articles that delve into AI's role in advancing engineering practices. These works highlight diverse applications, from human-robot interaction to AI-supported innovations in industrial and food technologies, underscoring the immense interdisciplinary potential of AI. We trust that these contributions will inspire meaningful dialogue and foster collaboration across the engineering community.

Editor-in-Chief
Vedran Mornar, President of the Croatian Academy of Engineering

EDITOR'S WORD

Dear readers,

I am pleased to present the new issue of Engineering Power journal, edited by Prof. Želimir Kurtanjek, PhD. In four articles, we present the possible applications of artificial intelligence using examples from mechanical engineering, chemical engineering and food technology. I hope you enjoy reading them.

Editor
Bruno Zelić, Vice-President of the Croatian Academy of Engineering

FOREWORD

Dear readers,

As we all know, regardless of our specific engineering expertise, AI has transformative power changing on how we do our daily work, and we all are aware that in near future our professions will undergo fundamental changes. Traditionally, engineering science and practice are based on human knowledge of sophisticated applications of laws of nature and logic (mathematics, physics, chemistry) focused on individual specific tasks, regardless how complex and big is the task. However, due to complexity of global digitalization and anthropogenic effects on Earth climate, our world has deeply changed requiring knew machine AI assisted knowledge. With the unprecedented speed of AI development there is a growing recognition of the need for ethical AI, ensuring that AI tools are developed and used responsibly for the world wide population. In spite of concerns, AI is not just a phase; it's here to stay and will continue to shape our lives in profound ways. This issue of Engineering Power brings four articles which show research and application of AI in human and robot communication, soft sensing in petrochemical synthesis plant, application of Bayes causal network for AI modelling of Tennessee-Eastman process, and a manuscript on AI supported computer vision in food technologies. Hopefully, this publications will motivate our readers for our national and also international interdisciplinary cooperation.

Guest-Editor
Želimir Kurtanjek, University of Zagreb, Faculty of Food Technology and Biotechnology (retired)

CONTENT

Tomislav Stipančić[1], Leon Koren[1]

Affective Interactive Virtual Agent that can Occupy Different Environments

[1]University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture, Ivana Lučića 5, 10000 Zagreb, Croatia

Abstract

This paper introduces a virtual being and a robot capable of nuanced non-verbal communication in virtual, mixed, and physical environments. The prototype of such an advanced system is PLEA, an "emotion-aware" virtual being. PLEA assesses a person's emotional state during interaction based on a multimodal approach and then uses that information during non-verbal communication. PLEA leverages AI-driven sensory systems to interpret visual and auditory data, enabling it to understand and respond appropriately to human emotions and actions in various realistic scenarios. The research focuses on endowing this entity with the ability to convey and interpret non-verbal cues such as body language, facial expressions, and contextual gestures.

The main goal of PLEA is to maintain human-robot interaction by building mutual understanding and common ground. We address the challenges in AI algorithms related to emotion and gesture recognition, 3D modeling, information visualization, and sensor integration, which are crucial for creating expressive, environment-appropriate non-verbal behaviors. We propose a multi-layered framework that enables virtual beings to adapt their non-verbal communication strategies to each specific environment. In virtual spaces, the emphasis is on creating expressive avatars. For mixed reality, integration into real-world contexts is crucial, while in physical environments, the challenge involves translating digital expressions into physical forms.

Keywords: *virtual beings, creative AI, cognitive robotics, multimodal interaction, mixed reality*

1. Introduction

Virtual beings, encompassing digital avatars and agents, represent the convergence of AI and virtual reality.

Virtual beings (VBs) represent embodying entities that can interact with humans in an intuitive and meaningful way [1]. They are not only programmed to engage in complex dialogues but are also capable of displaying a range of emotions and understanding human sentiments, making them increasingly indispensable in sectors like education, entertainment, and mental health [2].

Cognitive robotics, on the other hand, extends the capabilities of traditional robotics by integrating AI to create machines that can perceive, reason, and learn from their environment [3]. This allows them to operate autonomously in a variety of settings, from assisting in homes and workplaces to performing tasks in environments that are hazardous for humans. The cognitive aspect involves the robot's ability to make sense of the world through sensors and data, mimicking human cognitive processes such as learning, memory, and decision-making.

The synergy between virtual beings and cognitive robotics is particularly evident in mixed reality environments, where physical and digital realms coalesce. Here, virtual beings can be embodied through robots, allowing for a tangible interaction that blurs the lines between what's virtual and what's real. This symbiosis opens up new frontiers for human-machine interaction, where virtual companions can provide support, education, and companionship in more immersive and personalized ways.

The potential for virtual beings and cognitive robotics to transform human life is immense [4]. The power of AI can be used to create entities that understand humans and adapt to their needs, ultimately leading to a future where technology enhances every aspect of human experience.

PLEA is a virtual being and a robot that has been developed to analyze and employ behaviors using a form of biomimicry [5]. PLEA is based on recent research findings in human cognition, cognitive robotics, and human-robot interaction to develop new robot reasoning and interaction strategies. PLEA relies on Deep Learning AI and multimodal information fusion to predict the possible responses of the person interacting with it.

2. Robot and the virtual being design

PLEA represents a cutting-edge virtual being and a robot, conceived to inhabit the vast expanse of cyberspace.

PLEA's existence is anchored on a physical server, which, through the intricate web of the internet, bridges it to the outside world. In this way PLEA can be approached via every computer connected to Internet. A software agent can move through cyberspace from interface to interface depending on which interface requires communication. If the software agent is currently occupied with communication through one interface, another user requesting communication will receive a notification of the occupation. Once the initial communication is finished, the software agent will become available. Although it is possible to create a software agent that can communicate with multiple users simultaneously, this method was chosen to give the agent a personality. Such an agent is unique regardless of the interface used and the interaction environment. This unique setup empowers PLEA to transcend traditional boundaries and manifest itself across a multitude of environments, ranging from purely virtual environments to augmented realities and even interfaces that connect the digital with the physical.

The essence of PLEA's design is its versatility and adaptability. The PLEA software agent uses a contextual approach to reasoning, where the robot is seen as part of the environment, regardless of whether it is the real world or cyberspace. In these diverse settings, PLEA is not just a passive inhabitant, but an active participant. It is programmed to recognize and interpret human emotions, respond to commands, and initiate interactions, making it an intelligent companion capable of providing assistance, entertainment, or companionship. The potential applications of such a virtual being are vast, from serving as a personal assistant in smart homes, guiding users in virtual learning environments, to acting as a mediator in remote collaborations.

PLEA, as an "emotion-aware" system designed to reason and operate within realistic scenarios, where its capability to perceive, interpret, and respond to human emotions is paramount. This sophisticated level of emotional intelligence enables PLEA to understand and adapt to the nuanced emotional states of individuals, facilitating interactions that are not only responsive but also empathetic and supportive [6].

3. Virtual being and the robot design

The development of a virtual being using Unreal Engine's MetaHuman application is a sophisticated process that merges cutting-edge technology with creative design to bring hyper-realistic characters to life [7]. The initial step involves conceptualizing the virtual being's appearance, personality, and purpose, which serves as the foundation for its creation.

Upon finalizing the concept, the development moves into the MetaHuman Creator, a cloud-streamed application that allows for the creation of photorealistic human models with an unprecedented level of detail, as shown at Fig. 1.

In this step, it is possible to select from a wide range of preset faces that can be finely adjusted to achieve the desired look. Features such as skin complexion, eye color, hair style, and facial hair can be meticulously customized, enabling the creation of a unique virtual being that closely matches the initial concept.

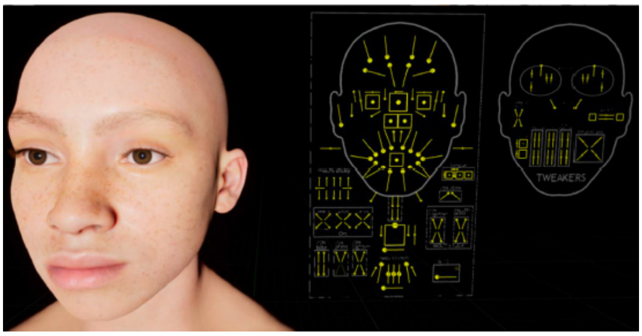

Fig. 1. MetaHuman control rig

The next phase involves rigging the MetaHuman for animation, a process that Unreal Engine simplifies through its advanced skeletal and facial rigging system [8]. This system provides a comprehensive set of tools for animating complex human motions and expressions, making it possible to bring the virtual being to life with realistic movements and emotional expressions.

Once the character is rigged, developers can integrate it into the Unreal Engine environment, where they utilize the engine's powerful rendering capabilities to achieve lifelike visuals. Lighting, shading, and environmental effects are fine-tuned to enhance the virtual being's realism within its virtual surroundings.

The final step involves programming the virtual being's behaviors and interactions using Unreal Engine's visual scripting system, Blueprint. This allows developers to create complex scenarios where the MetaHuman can interact with users or other elements within the virtual world, responding to inputs and exhibiting behaviors that reflect its programmed personality and purpose.

The computational architecture used in PLEA is a context-to-data interpreter that endows the machine with the capability to 'reason' based on constantly changing perspectives. In this way, PLEA can make decisions based on newly acquired information that is incomplete at the time of deciding on a particular behavior.

Deep Learning plays a significant role in managing a large amount of unstructured information that can be used to create control mechanisms and to approximate various phenomena or processes occurring in cyberspace or the real world [9]. Primarily, this refers to the development of neural network models that should enable the recognition of emotions based on facial expressions and by analyzing the speech of the person during the interaction. Fig 2. depicts the multimodal architecture of the model that controls PLEA's emotion recognition mechanisms.

The model consists of three primary components: acoustic, linguistic, and visual [10]. These modalities are integrated using a specific algorithm for combining information, which relies on assigning different weights to various factors. For visual data processing, the ResNet 300 face detection algorithm is employed to glean information directly from a real-time video stream. Furthermore, it incorporates a methodology inspired by Savchenko [11] to facilitate emotion detection. The acoustic component is engineered to distill features from sound recordings by employing techniques such as spectrograms and Mel-frequency cepstral coefficients [12]. The data is then classified using a convolutional neural network that has been previously trained. The linguistic component utilizes a manually developed feature extraction tool alongside a bag-of-words model to understand and process language. Emotion and intent recognition are accomplished using the Long Short-Term Memory (LSTM) algorithm [13]. In addition, specialized hardware is utilized to separate speech from any ambient noise and to identify the interaction targets of the system.

The updated algorithm operates by integrating data from all three modalities simultaneously during the information fusion process. It employs both weighted and base factors in equal proportion across these modalities. The results generated by each modality are scaled using a predefined weight factor, after which the values from all three are combined. This algorithmic method has proven effective in real-time applications where speed of processing is critical. A practical implementation of all three modalities is described in [14].

In the first step, PLEA acted as an emotional mirror, simply mirroring the facial expressions of the person it interacts with. The mirroring operation can be performed by copying significant points from the face of the person in interaction to the face of the virtual agent. This is usually achieved by employing an approach based on FACS (Facial Action Coding System), which defines and identifies significant points on the person's face [15]. These points are aligned with the positions of muscles responsible for generating the person's facial expressions. In the second approach, applied to PLEA, multimodal channels are used to reason about current emotions and to generate emotional responses on the face of the virtual agent. This approach is employed in this work to facilitate the future development of a mechanism for the autonomous generation of facial expressions based on perceived emotions. The new computational model will be trained using a data corpus collected from real human-to-human interactions. The main objective of this approach is to advance autonomous human-to-agent interactions. The agent will be able to create and implement strategies to comfort or make someone happy based on the new mechanism.

The PLEA physical robot can be understood as a physical interface through which the PLEA virtual agent is embodied in the real, physical world. The PLEA robot consists of multiple components, including: a control computer, a microcontroller that contains three microphones which

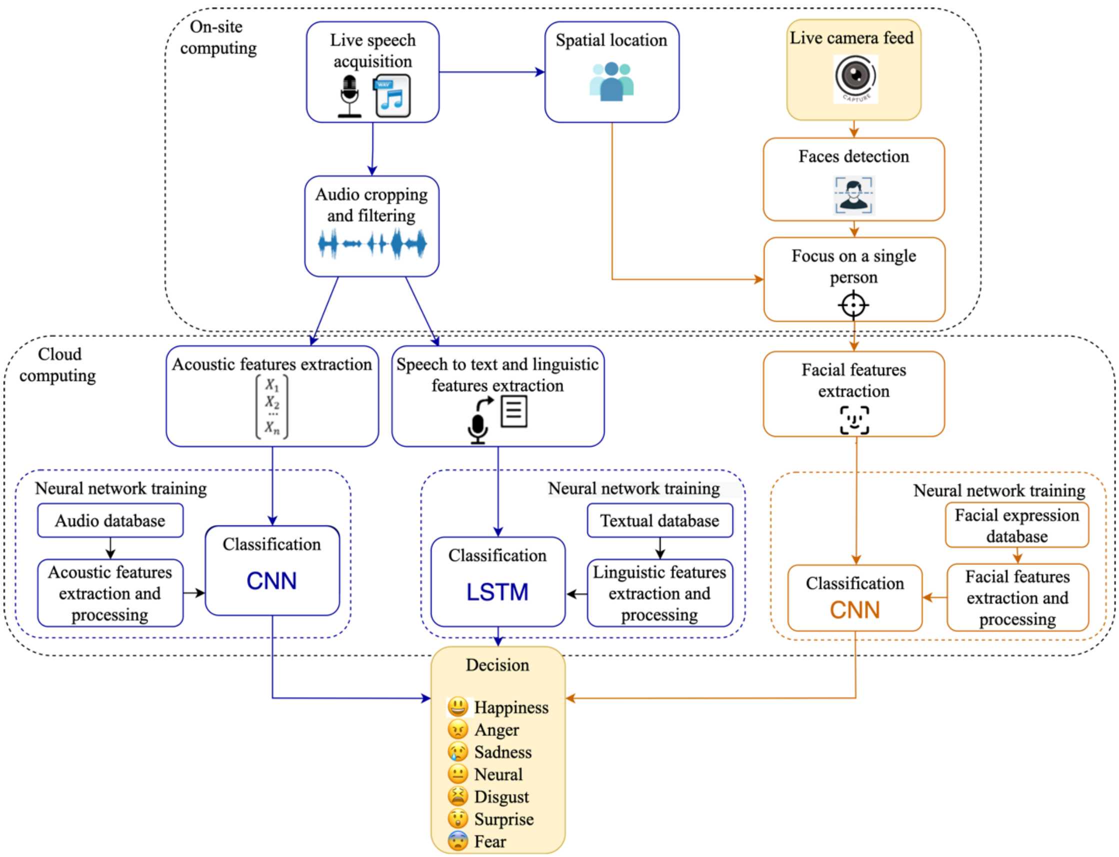

Fig. 2. Multimodal architecture of PLEA's reasoning model

enable the determination of spatial location for sound direction, an auxiliary power supply system for microcomputers used in case of a sudden interruption in electricity supply, a light projector, a base casing, the robot's neck, and the robot's head with a surface for projecting the face. The effect of embodiment is achieved through back-projected light. The light projector is placed within the neck part of the robot, and light projected onto the front face of the robot is used to display facial expressions in real time. The robot development process is shown at Figure 3. Figure 4 depicts different robot configurations: the first two figures show PLEA within the research lab, and the last two depict PLEA in Central Library at the Art & AI Festival in Leicester, the UK in 2022.

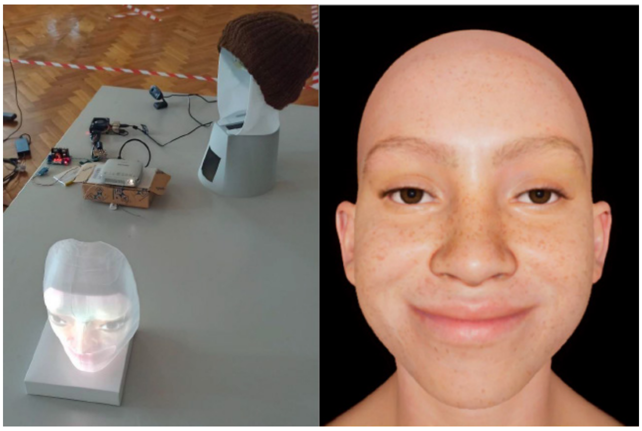

Fig. 3. Physical robot development process

Fig. 4. Different PLEA robot configurations

4. Data gathering and results

The cognitive model used to control the robot responses is based on data collected on several occasions. The first two events took place more than a year ago at the British Science Festival and the Art AI Festival in Leicester, UK (https://www.art-ai.io/programme/plea2/). The subsequent opportunity arises during the controlled experiments conducted at the Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Croatia (FAMENA).

Attendees of the festivals had the freedom to walk up to the robot and engage in non-verbal communication. During these interactions a lot of data is collected to be used on building the new model for autonomous generation of facial expressions of the robot.

Examination of the data gathered through this method revealed that environmental factors significantly influence interactions with PLEA. As a result, a second set of data was obtained through interviews under more controlled conditions where participants were posed targeted questions regarding their experiences interacting with the robot. The conceptual framework for the interviews was inspired by the Media Equation Theory [16, 17, 18, 19, 20].

Data gathering under controlled conditions was conducted at FAMENA. During this time, 23 participants were interviewed over the span of a week. Each participant was asked to engage with PLEA for four minutes, within which they could spontaneously share emotional signals through facial expressions. This duration was chosen to allow participants to establish a certain degree of rapport with the robot. Following the interaction, participants responded to 12 questions posed by a human interviewer.

Figure 5 shows an example of a 24-second interaction in which joy and surprise alternate.

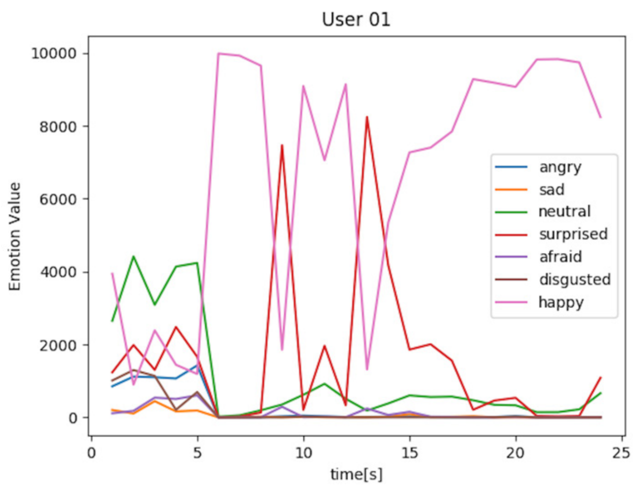

Fig. 5. PLEA emotion-interaction plot

The graph indicates that the person was initially confused. The first distinct emotion expressed by the person was surprise, to which PLEA responded accordingly. Following this, the person smiled upon noticing this facial expression, and PLEA mirrored the smile. Subsequently, the person expressed surprise once more. In the next phase, PLEA reacted with surprise, prompting the person to respond with a smile, indicating happiness. Afterwards, the person exhibited surprise for an extended period, which persisted until the conclusion of the interaction (non-verbal communication).

The interviews were then recorded and subsequently analysed. The group of participants was quite uniform, as the interactions were confined to a controlled environment within the research laboratory rather than a public setting, consisting of university students. All participants gave their consent for the use of their data for research purposes. Every exchange prompted a series of emotional reactions from the robot, which were reciprocated with the individual.

5. Discussion

Within this framework, our research method zeroes in on the ways in which individuals attribute human qualities to non-human entities. In the context of this investigation, treating PLEA with human attributes is termed as personification [21].

PLEA is equipped to engage with individuals wherever they are. This omnipresence allows for a seamless integration of PLEA into daily human activities, offering a new dimension of interaction that is both dynamic and responsive. According to the Media Equation Theory, this study of interactions with PLEA seeks to unravel the underlying reason - why do individuals behave in the manner we've observed [22]? Our findings indicate a natural inclination among people to attribute human emotions, intentions, and sentiments to inanimate objects, even with the understanding that these objects are not sentient. This phenomenon also seems to apply to PLEA, although the participants know that it consists of a plastic housing modelled on a human head.

6. Conclusions, and future work

The virtual being PLEA and the robot are at the interface between technology and human emotion. Designed to mimic and respond to human expressions through advanced algorithms and projected light configurations, PLEA exemplifies the fusion of Artificial Intelligence, Mixed Reality, and Smart Environment Design with empathetic interaction. The virtual PLEA agent is an inhabitant of cyberspace. It is also encapsulated in a plastic shell resembling a human head and serves as a focal point for studies on personification and emotional exchange, raising pivotal questions about human relationship with cognitive artifacts in a world where technology increasingly mirrors humanity.

Previous research suggests that users tend to ascribe human emotions, motives, and feelings to objects even when it is clear the object in question is not human. This is also the case with PLEA, which is strongly anthropomorphized by its design as a human head, albeit made of a plastic shell. PLEA as a ghost-like entity evoked the similar reactions. In this context, the methodological approach is focused on communicative practices, where people assign human characteristics to cognitive artefacts. This is defined here as personification, and the main research questions being explored are which features, characteristics, and functionalities are most appropriate in facilitating PLEA's personification for the use case scenario.

PLEA as a virtual being and the robot has great potential to be used in different areas of human activity, including healthcare, education, smart living, robotics, etc. The adaptability of PLEA allows it to cater to a broad spectrum of emotional and cognitive needs. In care settings, such a system could assist staff by offering additional support to residents, thereby enhancing the overall quality of care. By integrating PLEA into these environments, it is possible to create a more nurturing and responsive care ecosystem that prioritizes the emotional well-being of its inhabitants. Whether the aim is to provide company through conversation, remind the user of important tasks and medication or simply offer an emotionally attuned ear - the possible applications of PLEA are many and varied.

As a proof of concept, PLEA demonstrates a paradigm shift, moving away from sensed data towards contextual anticipation. Current multidisciplinary research is carried out with partners specializing in Creative Technologies and Informatics focused on human communication with artefacts. The development of the PLEA reasoning mechanism is aimed at autonomous agent responses that are customized to current and context-dependent expectations and needs. The use of PLEA in real-world applications will include further pilot projects in public spaces as well as in intelligent learning and healthcare environments.

Acknowledgments

This work has been supported in part by the Croatian Science Foundation under the project "Affective Multimodal Interaction based on Constructed Robot Cognition—AMICORC (UIP-2020-02-7184)".

7. References

[1] Montanha, R., Araujo, V., Knob, P., Pinho, G., Fonseca, G., Peres, V., & Musse, S. R. Crafting Realistic Virtual Humans: Unveiling Perspectives on Human Perception, Crowds, and Embodied Conversational Agents. In 2023 36th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI) (2023) 252-257.

[2] Stipancic, T.; Koren, L.; Korade, D.; Rosenberg, D. PLEA: A Social Robot with Teaching and Interacting Ca-pabilities. Journal of Pacific Rim Psychology (2021), 15, https://doi.org/10.1177/18344909211037019

[3] Ciria, A., Schillaci, G., Pezzulo, G., Hafner, V. V., & Lara, B. Predictive processing in cognitive robotics: a review. Neural Computation, 33(5), (2021) 1402-1432.

[4] Walker, M., Phung, T., Chakraborti, T., Williams, T., & Szafir, D. Virtual, augmented, and mixed reality for human-robot interaction: A survey and virtual design element taxonomy. ACM Transactions on Human-Robot Interaction, 12(4), (2023) 1-39.

[5] Koren, L., Stipančić, T., Ričko, A., & Benić, J. Context-Driven Method in Realization of Optimized Human-Robot Interaction. Tehnički glasnik, 16(3), (2022) 320-327.

[6] Koren, L., & Stipancic, T. Multimodal emotion analysis based on acoustic and linguistic features of the voice. In International Conference on Human-Computer Interaction (2021) 301-311. Cham: Springer International Publishing.

[7] Lyytinen, K., Nickerson, J. V., & King, J. L. Metahuman systems= humans+ machines that learn. Journal of Information Technology, 36(4), (2021) 427-445.

[8] McKim, J. Animation without animators: from motion capture to MetaHumans. Animation Studies 2.0. (2022)

[9] Zhang, Q., Yang, L. T., Chen, Z., & Li, P. A survey on deep learning for big data. Information Fusion, 42, (2018) 146-157.

[10] Koren, L., Stipancic, T., Ricko, A., & Orsag, L. Multimodal Emotion Analysis Based on Visual, Acoustic and Linguistic Features. In International Conference on Human-Computer Interaction (2022) 318-331. Cham: Springer International Publishing.

[11] A. V. Savchenko: Facial expression and attributes recognition basedon multi-task learning of lightweight neuralnetworks. In IEEE 19th International Symposium on Intelligent Systems and Informatics; 2021; Subotica. DOI: 10.1109/SISY52375.2021.9582508

[12] Abdul, Z. K., & Al-Talabani, A. K. Mel Frequency Cepstral Coefficient and its applications: A Review. IEEE Access. (2022)

[13] DiPietro, R., & Hager, G. D. Deep learning: RNNs and LSTM. In Handbook of medical image computing and computer assisted intervention (2020) 503-519. Academic Press.

[14] Koren, L., Stipancic, T., Ricko, A., & Orsag, L. Person localization model based on a fusion of acoustic and visual inputs. Electronics, 11(3), (2022) 440.

[15] Canal, F. Z., Müller, T. R., Matias, J. C., Scotton, G. G., de Sa Junior, A. R., Pozzebon, E., & Sobieranski, A. C.. A survey on facial emotion recognition techniques: A state-of-the-art literature review. Information Sciences, 582, (2022) 593-617.

[16] Littlejohn, Steven. Theories of Human Communication: Eleventh Edition. Waveland Press, Inc. (2016) 202. ISBN 978-1478634058.

[17] Reeves, B. & C.I. The media equation: How people treat computers, television, and new media like real people and places. Center for the Study of Language and Information. (1996) Cambridge University Press.

[18] Reuten, A., Van Dam, M., & Naber, M. Pupillary responses to robotic and human emotions: the uncanny valley and media equation confirmed. Frontiers in psychology, 9, (2018) 774.

[19] Klowait, N. The quest for appropriate models of human-likeness: anthropomorphism in media equation research. AI & SOCIETY, 33(4), (2018) 527-536.

[20] Fruchter, R., Nishida, T., & Rosenberg, D. Social Intelligence Design for Social Computing. In International Conference on Human-Computer Interaction (2022) 545-558. Springer, Cham.

[21] Mukaetova-Ladinska, B. E., Harwood, T., J., M.: Artificial intelligence in the healthcare of older people. Archives of Psychiatry and Mental Health (4) (2020). https://doi.org/10.29328/journal.apmh.1001011

[22] Littlejohn, S.W.: Theories of Human Communication, Eleventh edition. Waveland Press, Inc., Long Grove, Illinois (2017)

Željka Ujević Andrijić[1], Nikola Rimac[1], Ines Martić[2]

Development of a Soft Sensor in a Propylene Production Plant

[1]Department of Measurements and Process Control, University of Zagreb Faculty of Chemical Engineering and Technology, 10000 Zagreb, Croatia

[2]Department of Digital transformation, INA - Industrija nafte, d.d., Avenija Većeslava Holjevca 10, 10020 Zagreb, Croatia

Abstract

Propylene is a crucial intermediate in petrochemical synthesis, which requires high purity. This study details the creation of a soft sensor model for continuous monitoring of propylene levels in a propane/propylene splitter refinery facility. The data obtained from the plant of the input variables are processed using different pre-processing methods. Soft sensor models of neural networks with multilayer perception (MLP) and neural networks with long short-term memory (LSTM) were created using the Python programming language. During the development of the MLP model, various hyperparameters were tested, including the number of neurons in the hidden layer and the impact of the activation function type on the model's quality. Similarly, when developing the LSTM model, the number of LSTM units and the number of time steps into the past were also examined. A statistical analysis of the findings was performed, which revealed that both model types give strong correlation values between model data and real data of propylene content and that both neural network model types may be used in the refinery information system. The use of developed soft sensors guarantees that propylene content information is always up to date and continuous, allowing for fast responses to changes in propylene content for enhanced process management. The soft sensors improve the end product's quality and can result in considerable cost savings.

Key words: *soft sensor, neural networks, multi-layer perceptron, long short-term memory networks, propylene*

1. Introduction

Soft sensors

Many processes and plants struggle to measure crucial process variables on a continuous and reliable basis. The process control of such processes is based on laboratory analysis, which is frequently uncommon and time-consuming, and the cost of using and maintaining online analyzers can be exceedingly expensive. For these reasons, there is a need for the development and use of intelligent software that can continuously monitor various physicochemical properties and process variables. Using soft sensors, it is feasible to determine the state of difficult-to-measure variables from easily measurable secondary variables like temperature, pressure, and flow by identifying their functional correlations. [1]. The creation of soft sensors in the field of process engineering necessitates expertise from several scientific fields, as well as a mix of scientific research and plant experience. Soft sensors enable process engineers to analyze process variables and circumstances that cannot be monitored in real time and utilize them to improve process control. The problem of not being able to measure product properties in real time is due to several reasons, such as the fact that refineries do not have enough process analyzers installed in the plant itself because they are very costly to install and maintain, or that for some variables, such as chemical reaction conversions, there is no possibility of direct measurement at all. In process engineering, it is therefore highly desirable to use such software estimators in the areas of process control and management, fault detection and process diagnostics, instrumentation and measurement, where the need for measuring devices is reduced and measuring devices are replaced. [2] Given the frequency of failures and maintenance requirements as well as the high cost of online analyzers used for continuous measurement of propylene content, it is necessary to develop a model for continuous assessment of propylene product content in a propane/propylene splitter production plant.

Soft sensor models are divided into three basic categories: white-box models, also referred to as "mechanistic", "analytical" and "fundamental" models, gray-box models, also referred to as "semi-analytical" and "hybrid" models, and black-box models, also referred to as "empirical" models. [3] Black-box models depend exclusively on collected data, whereas white-box models utilize theoretical knowledge of the process. [4] Gray box models combine black and white box models to address issues; that is, their construction uses theoretical knowledge of the process in conjunction with data obtained from the process. [5] For the purpose of this study, black-box models were created, i.e., models of neural networks with multilayer perceptions and neural networks with long-term short-term memory. The term 'black box' refers to a model in which it is not necessary to know the nature of the process itself. Instead, the model is constructed solely by identifying the functional connection between the input and the output of the process. [6]

The general process of developing soft sensor models includes data collection, pre-processing and model development. During data collection, the availability and quality of the data is important as it forms the basis for all subsequent steps. In the model development phase, theoretical knowledge of the process, such as physical principles, reaction equations, etc., can facilitate the selection of input variables and data pre-processing. [2] Additional information from operators and maintenance specialists can aid in creating a soft sensor. After collecting data, it is necessary to preprocess it, which includes procedures such as data visualization, identification of extreme and missing values, and data filtering. Outliers were identified by applying the 3σ rule when preparing this paper. It is a fundamental concept in statistics that is used to describe the distribution of data and the probability of finding a data point within a certain range. It is based on the concept of standard deviation (σ) and is particularly important in the context of normal distributions. According to this rule, all data points that lie outside the range $\mu \pm 3\sigma$, where μ is the mean of the data, can be considered extreme values or anomalies.

When preprocessing the data, influencing variables are selected; these are the input variables that have a specific impact on the output variable, i.e. on the variable whose value must be predicted. The type of model for sensor development is also defined. In this paper, Pearson›s correlation coefficient was employed to select the influencing variables. This statistical measure quantifies the linear relationship between two continuous variables, with values ranging from -1 to +1. These values indicate both the intensity and the direction of the relationship between the variables by estimating how well the data points of two variables displayed in the coordinate system are aligned along a straight line. [7]

Artificial neural networks

Artificial neural networks are computer models employed in machine learning, processing data based on the workings of the human brain, albeit on a much smaller scale. These networks are instrumental in solving various complex problems, including pattern recognition, classification, regression, and optimization tasks [8]. The initial type of neural network models developed were those with multilayer perceptrons, featuring multiple hidden layers and activation functions. Information flows bidirectionally through such networks, with input information propagating forward, and weight coefficients updated backward using the error gradient.

The basic architecture of such a neural network, shown in Figure 1, consists of three types of layers composed of nodes: an input layer, an arbitrary number of hidden layers, and an output layer, where the output of one layer is the input to the next layer. The input layer receives the raw data, and no computations are performed in this step, but the input neurons simply pass the input data to the hidden layer. Activation functions are introduced in the hidden layer, and most of the time all hidden layers use the same activation function. The output layer is the last layer of the network, which receives the information learned via the hidden layers and converts it into the final result. An activation function can also be used in the output layer, which is usually different from the one used in the hidden layers.

[9] Activation functions are used to introduce non-linearity into the models, which is achieved by introducing non-linear activation functions. An important feature of nonlinear activation functions is their derivability, so that a backpropagation algorithm can be applied to calculate errors with respect to the weight coefficients and accordingly optimize the weights using one of the optimization techniques. The nonlinear activation functions most commonly used in neural networks, including sigmoid, ReLU, ELU, and tanh, were also employed in the preparation of this paper. [10]

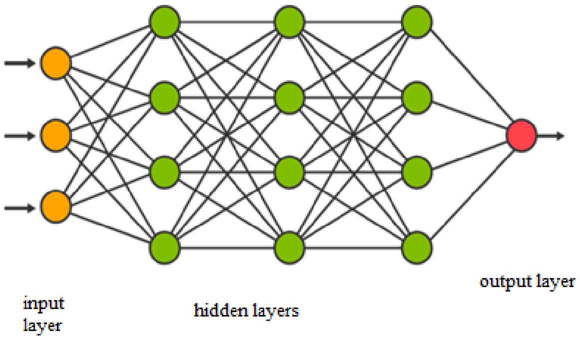

Fig. 1. *The basic structure of a multi-layer perceptrons neural network*

Each node has its own input information, through which it receives communication from other nodes and the environment, and its own output information, through which it communicates with other nodes and the environment. In addition, each node has its own activation function f, through which the input information is transformed into output information. The connections between the neurons in an artificial neural network are weighted by a weight coefficient w_i that represents the strength of the connection between neurons. The inputs are multiplied by the weight coefficients and then summed in the hidden layer using the *net* (Eq. 1) summation function, and their sum is compared to the neuron's threshold . If the sum of the weighted inputs exceeds the threshold of the neuron , the activation function *f* generates the output of the neuron *y* (Eq. 2). [11,12]

$$net = \sum_{i=0}^{n} w_i x_i \quad (1)$$

$$y = f(net - \theta) \quad (2)$$

Another type of neural network model that has recently been developed is the long short-term memory (LSTM) network. This variant of a feedback neural network addresses the challenge of retaining long-term dependencies present in conventional feedback neural networks. This is achieved by installing a "memory cell" that acts as a container and can store information over a longer period of time. The structure of a memory cell is shown in Figure 2. The memory cell of the LSTM model is managed by three so-called gates: input gate, forget gate and output gate. An input gate, mathematically represented by the formula (4), controls what information is added to the memory cell, a forget gate, mathematically represented by the formula (3), controls what data is removed from the memory cell, and an output gate, mathematically represented by the formula (5), controls what information is output from the memory cell.

$$f_t = \sigma((w \times x_t + w \times h_{t-1}) + b) \quad (3)$$

$$c_t = C_{t-1} \times f_t + i_t \times c'_t \quad (4)$$

$$h_t = o_t + tanh(c_t) \quad (5)$$

In this way, LSTM networks can selectively retain or discard information, allowing them to learn long-term dependencies. The cell state (represents long-term memory and is updated using a multiplication or summation operation, and this state is not directly affected by the weights. The hidden state () represents short-term memory and is affected by the weight values.[13]

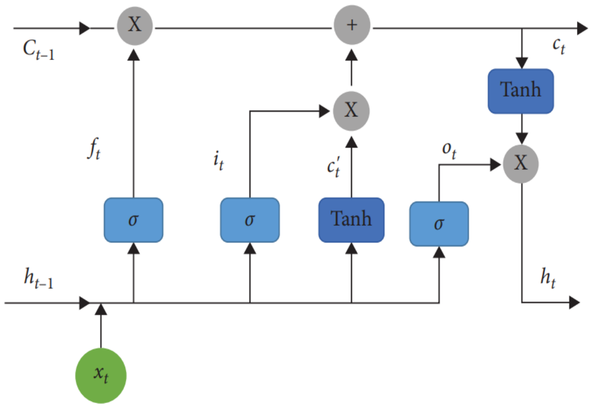

Fig. 2. *The structure of a memory cell of an LSTM network [13]*

With the backpropagation algorithm the gradient is calculated, that is, the amount by which the weights w_i must change in the positive or negative direction in order to minimize the loss function. After the network calculates the output, it is compared to the actual value of the output. The square of the difference between these two values, or the mean squared error, represents the loss function for regression problems. MSE can be mathematically represented by the formula [14]:

$$\text{MSE} = \frac{\sum_{I=1}^{n} (y_i - \bar{y}_i)^2}{n} \quad (6)$$

where *n* is the number of samples, is the actual output value, and is the corresponding predicted output value. The initial values of the weights are chosen randomly and passing through the network gives the first value which is compared with y_i.

After calculating the gradient, the weights are optimized layer by layer, starting from the last hidden layer and working backwards towards the input layer until the error no longer changes. At this point, learning is stopped, and it is assumed that the loss function has reached a global minimum. Problems that can occur during this learning process include reaching a local minimum instead of a global minimum and overfitting the network. An over-trained network has a very small error on the training data, but on the test data the error is significantly higher, i.e., the network loses its ability to generalize and becomes specific only to a particular data set. Although it is desirable that the error is as small as possible, it is also important that it is consistent on the training and validation dataset of the model so that the model has reasonable efficiency when introducing new data.[15]

2. Materials and methods

Propane/propylene splitter

The propane/propylene splitter (PPS) plant consists of two sections, which are shown in Figure 3. Section 1 is the section for the feed vessel of the raw material and its purification, while section 2 is for the separation of the propane-propylene mixture. For the purposes of this paper, more attention was paid to section 2, which consists of two distillation columns where the separation of propylene and propane takes place. The aim of this unit is to achieve maximum purity of propylene and minimize the propylene content in propane. The propylene produced in the process is a colorless, odorless, tasteless, flammable gas with a required minimum purity of 99.6% by volume.

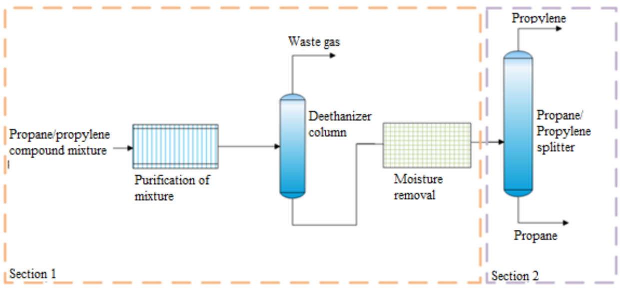

Fig. 3. *Propane/propylene splitter plant section*

The models were developed using Python, a programming language that contains numerous packages and modules specialized for machine learning and data analysis. First, NumPy and Pandas packages were utilized for data preprocessing and analysis. Matplotlib was employed for data visualization, allowing for the creation of insightful plots and graphs to better understand the data distribution and relationships. Neural network models were developed using the Keras and Scikit-learn tools.

Data from the refinery database was used to assess the content of propylene products in the PPS-producing facility. The imported data consists of input and output variables measured over two months. For all input and output variables, approximately 75,000 data points were collected with a sampling time of 1 minute. The data points for the output variable (propylene content) were taken from the AI201A/B gas chromatograph with a sampling time of 7 minutes.

After data acquisition in the system, data pre-processing followed. Data preprocessing for the development of the MLP model included: shifting the input variables data for a certain number of time steps based on the dead time, i.e., the time it takes for the input variable to begin impacting the output, followed by detecting and deleting extreme values. The detection of extreme values was carried out using the 3σ method and by visual inspection, as the values recognized as extreme by 3σ can be part of the dynamic behavior of the process, which is important for the creation of the model.

Data preprocessing during the creation of the LSTM model comprised moving the input data by a particular number of time steps owing to input dead time, and extreme values were not deleted because such models employ data sequences in which the time sequence of the data is significant.

The correlation between the input variables and the output variable was determined using the Pearson correlation coefficient. When selecting the number of input variables, care was taken to ensure that the number of influencing variables was large enough to capture as much important information from the process as possible, but not too large due to the need to simplify the model. Because of the connections discovered, several of the input variables were not included in the model's further development. Seven of the 16 available input variables were chosen based on their association with the output for further model building.

When developing the MLP model, the data set was divided into three parts: the training set, the test set, and the validation set. The training set comprised 80% of the data, while the test set comprised 20% of the data. The validation set consisted of 20% of the training set. When developing the networks, the number of hidden layers was set to 1, while the activation functions and the number of neurons in the hidden layer were changed. The activation functions used were sigmoid, tanh, ReLU and ELU, and the number of neurons in the hidden layer varied in the range from 1 to 20. A network with the highest correlation coefficient and lowest mean square error was chosen for each activation function. The Adam algorithm was used as the optimization algorithm.

When developing the LSTM network model, the data set was split in the same way as when developing the MLP model. The number of LSTM units ranged from 1 to 35, representing the number of time steps in the past, and the activation functions were modified. The activation functions used were the previously mentioned sigmoid, tanh, ReLU and ELU functions. For each of the activation functions used, the best network was selected, i.e., the network with the maximum correlation coefficient and the minimum mean square error. The Adam (Adaptive Moment Estimation) algorithm was used as the optimization algorithm.

In addition to the graphical comparison of the model results and the experimental measurements, the created models within a specific set of models were compared on the basis of the mean square error calculated on the training set, the test set and the validation set, as well as on the basis of the Pearson correlation coefficient between the results calculated by the model and the real data. The best models from a given set of models were additionally evaluated by analyzing the model's error trend (residuals) and using error histograms.

3. Results and discussion

The plant provided process information, including 16 input variables sampled every 1 minute and a single output variable, propylene content, which was similarly sampled every 1 minute. Approximately 75,000 data points were collected for each input and output variable. During the pre-processing of the data for the development of the MLP and LSTM models, the values of the input variables were temporally shifted by 7-time steps, equivalent to 7 minutes into the past concerning the output variable (representing the time required for sampling and calculating by the online analyzer).

When creating the MLP model, the dead time was taken into consideration, which represents the time it takes for the input variable to affect the output variable. Pearson correlation coefficient values were computed with a 7-time step shift, incorporating input variable lag times of 30, 60, 90, and 120 minutes. Following the comparison of all correlation values, data with a 7-time step shift plus a lag of 90 minutes were chosen for further analysis. Initially, influential variables were identified as those with a correlation greater than ± 0.10 with the output variable. Subsequently, mutual correlations among different input variables were examined. Considering the obtained correlation results, the input variables were selected for further development of the model: TI214 (exit temperature from the C-202 column (bottom of the column)), FIC204 and FIC203 (reflux of the propylene product at the top of the C-202 column), PI203A (pressure at the top of the C-202 column), PI202 (product pressure at the outlet of the C-201 column) and FIC201 (product flow of the C-201 product at the top of the V-102 column). The variables PI202 and PI203B with a correlation greater than ± 0.10 were selected as influencing variables. The next step was to analyze the relationships between the various input variables to determine their mutual correlations. The input variables for the further development of the model were selected based on the correlation results obtained: TI214 (outlet temperature from the C-202 column (bottom of the column)), FIC208 (reflux of the product from the bottom of the C-202 column). The variables PI202 and PI203B showed a similar correlation value with AI201A/B, of 95.5%. Consequently, PI203B was excluded from further model development. Similarly, the variables FIC202 and FIC208 exhibited a close correlation of 99.5% with AI201A/B, leading to the exclusion of FIC202 from additional model development.

The 3σ rule was used to identify extreme values, and missing values and related data from other variables were eliminated during the development of the MLP model.

When developing the LSTM model, the dead time of the input variables was not considered when pre-processing the data, as the LSTM shifts the input variables into the past by a certain number of time steps during its calculation. The dead time of the process's input variables can be determined by adjusting the number of previous time steps considered by the LSTM model during development. As with the development of the MLP model, all input variables whose correlation with the output variable was greater than ± 0.10 were first taken as influencing variables and then the mutual correlations between different input variables were observed. The results were the same as the MLP model, i.e., the variables PI203B and FIC202 were not considered in the further development of the model. Extreme values were not removed during data preprocessing, which is problematic for LSTM models that rely on temporal sequences of data. During data pre-processing for both model types, resampling was carried out with a time step of 3 minutes in order to reduce the amount of data and thus the calculation and overall model development time. The Pearson correlation coefficient was then recalculated, which increased slightly, i.e., all variables considered in the previous step still had a satisfactory Pearson coefficient value for continued model development.

The models were developed using the Python programming language (Python version 3.9.7) and its Anaconda distribution (Anaconda Navigator version 2.1.1). 80 MLP models were developed in advance, whose structure differed in terms of the hyperparameters of the model, such as the number of neurons in the hidden layers and the activation functions used. In addition, 140 LSTM models were preliminarily developed, which differed in terms of the hyperparameters of the LSTM model such as acti-

Table 1. *Pearson correlation coefficients for the input variables on the training data set for the MLP model*

	Correlation with AI201A
TI214	-0.099
FIC208	0.497
FIC204	0.799
FIC203	0.504
PI203A	0.137
PI202	0.554
FIC201	-0.158

Table 2. *Pearson correlation coefficients for the input variables on the training data set for the LSTM model*

	Correlation with AI201A
TI214	-0.103
FIC208	0.488
FIC204	0.742
FIC203	0.489
PI203A	0.137
PI202	0.554
FIC201	-0.116

vation functions, number of LSTM units and number of steps in the past.

The entire data is divided into a training set and a test set in a ratio of 0.8:0.2. 20% of the data from the training set was used as the validation set. The Pearson coefficient values were then calculated only for the training data set to avoid possible data loss. Indeed, if the entire data is used when developing a model, there is a possibility of data leakage, where data that does not belong to the training dataset is used for some steps in the creation of the model. The freshly acquired information may assist the model in learning or discovering something it would not have known otherwise, resulting in a falsely satisfied appraisal of the model's performance. The results of the correlations of the selected input variables with the output variable are shown in Table 1 for the MLP model and in Table 2 for the LSTM model.

MLP model results

During the creation of the MLP model, a decision was made to have only one hidden layer, but adjustments were made to the activation functions and the number of neurons within that hidden layer. Of the activation functions, the sigmoid, tanh, ReLU and ELU functions were used, and the number of neurons in the hidden layer varied between 1 and 20. A computer experiment was conducted for each of the chosen activation functions to ascertain the optimal number of neurons in the hidden layer that yields the highest correlation factor and lowest mean square error on the training dataset. In addition to the correlation factor and the mean square error, the success of the model is also influenced by the mutual compatibility of the correlation factors of the training and the test and validation dataset as well as the overall data. An optimal model should have high and uniform correlation factors for all data sets and a minimum squared error (as close to 0 as possible). The results of the influence of the number of neurons in the hidden layer on the accuracy of the model are shown graphically in Figure 4.

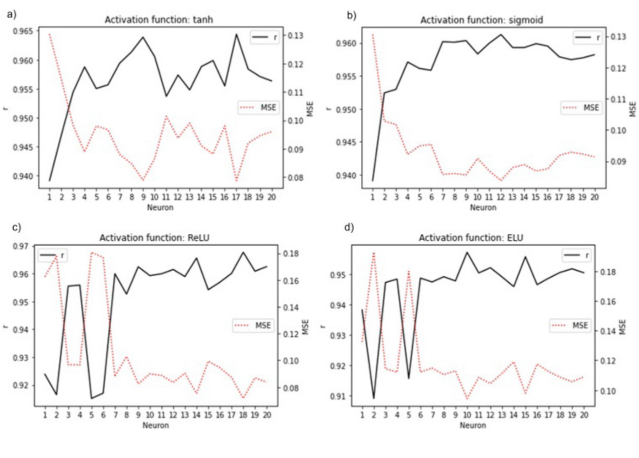

Fig. 4. *The influence of the number of model neurons in the hidden layer on the correlation factor and the MSE for the activation functions a) tanh (top left), b) sigmoid (top right), c) ReLU (bottom left) and d) ELU (bottom right)*

The findings presented in Figure 4 demonstrate that the models utilizing the tanh and sigmoid activation functions exhibit the highest level of stability in terms of correlation coefficient and the lowest squared error on the training set. Upon conducting a thorough examination of the acquired models, considering the numerical data presented in Table 3 and the visual representations in Figure 4, it was determined that all models exhibit a satisfactory level of accuracy for implementation in the plant, attributed to the elevated correlation coefficients. The smallest mean square error on the test set is shown by the ReLU model with 18 neurons in the hidden layer, and slightly higher by the tanh model with 17 neurons. Finally, the model with the activation function tanh and 17 neurons in the hidden layer was selected as the best model in terms of correlation coefficients and minimum square error. The validation set (Fig. 5) and the entire dataset (Fig. 6) display the outcomes of a graphical analysis, illustrating the comparison between the model and the actual data for this network.

Table 3. *Comparison of MLP models with respect to correlation factors and MSE for different activation functions*

Activation function	tanh	sigmoid	ReLU	ELU
Number of neurons in hidden layer	17	12	18	10
Correlation - entire dataset	0.964	0.961	0.967	0.956
Correlation - training set	0.964	0.961	0.968	0.957
Correlation - test set	0.962	0.959	0.967	0.953
Correlation -validation set	0.961	0.959	0.967	0.955
MSE - training set	0.078	0.084	0.072	0.094
MSE - test set	0.077	0.082	0.070	0.096
MSE - validation set	0.076	0.084	0.069	0.093

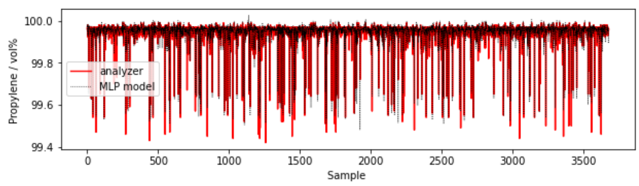

Fig. 5. *Comparison between real propylene content and MLP model on validation data for the model with tanh and 17 neurons in the hidden layer*

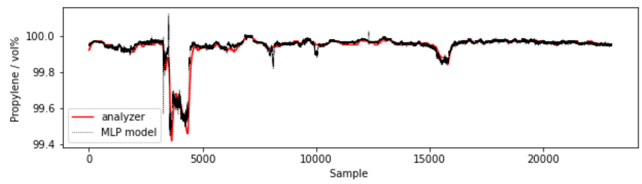

Fig. 6. *Comparison between real propylene content and the MLP model on the entire data for the model with tanh and 17 neurons in the hidden layer*

Figure 7 illustrates the histogram depicting the distribution of error values across the entire data of the MLP model. Conversely, Figure 8 displays the histogram representing the distribution of error values specifically on the validation data of the MLP model. The histograms illustrate that the error range lies within ± 0.15% in both cases, with the majority of errors falling within ± 0.1%. This level of accuracy is deemed satisfactory for the plant's application.

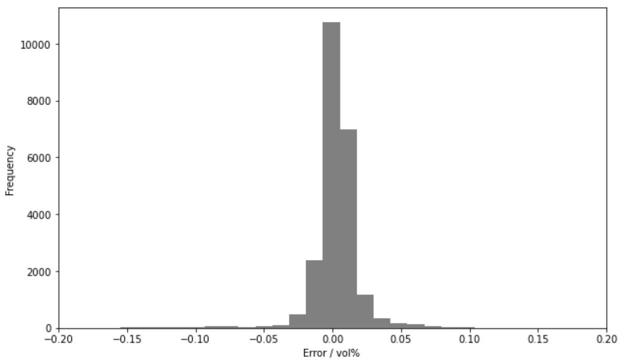

Fig. 7. *Histogram of the distribution of error values on the entire data for the MLP model tanh with 17 neurons in the hidden layer*

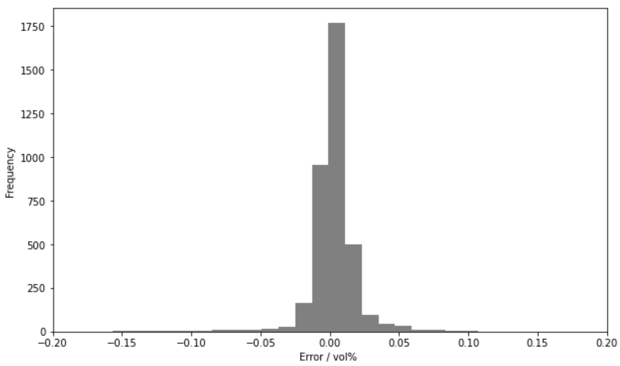

Fig. 8. *Histogram of the distribution of error values on the validation data for the MLP model with tanh and 17 neurons in the hidden layer*

LSTM model results

During the development of the LSTM networks, the number of LSTM units varied from 1 to 35, which also represents the number of time steps, and the activation functions were changed. The previously mentioned sigmoid, tanh, ReLU and ELU functions were used as activation functions. An extensive computer experiment was conducted to assess the optimal number of LSTM units in the hidden layer for each selected activation function. The experiment aimed to identify the configuration that yields the highest correlation factor and the lowest mean square error on the learning dataset. The success of the model depends on more than just the correlation factor and mean square error; it is also influenced by the mutual compatibility of correlation factors across the learning dataset, test dataset, and evaluation dataset, as well as the entire dataset. An optimal model is characterized by strong and uniform correlation factors across all data sets, with the squared error minimized to nearly zero. Figure 9 visually presents the correlation between the number of steps into the past and the accuracy of the model.

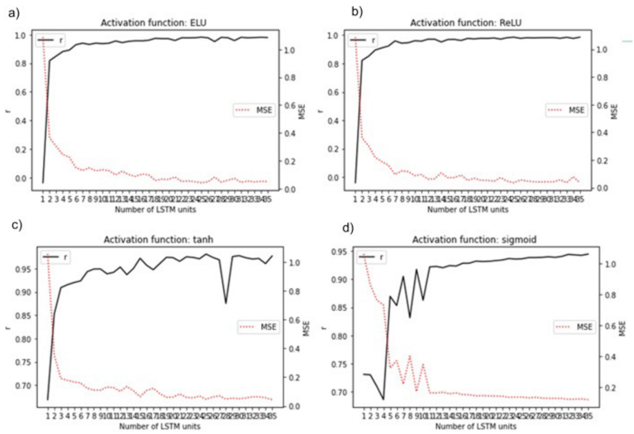

Fig. 9. *The influence of the number of model LSTM units in the hidden layer on the correlation factor and the MSE for the activation functions a) tanh (top left), b) sigmoid (top right c) ReLU (bottom left) and d) ELU (bottom right)*

The results shown in Figure 9 indicate that the models with the activation function ReLU and ELU have the most stable correlation coefficient and the lowest squared error in the training set.

Following an in-depth analysis of the obtained models, with reference to the numerical findings in Table 4 and the visual representations in Figure 9, it was established that all models demonstrate sufficient accuracy for practical use at the plant, given the high correlation coefficients.

Table 4. *Comparison of LSTM models with respect to correlation factors and MSE for different activation functions*

Activation function	tanh	sigmoid	ReLU	ELU
Number of LSTM units	25	31	33	31
Number of LSTM time steps	25	31	33	31
Correlation - entire dataset	0.981	0.940	0.985	0.981
Correlation -training dataset	0.982	0.938	0.985	0.981
Correlation - test dataset	0.979	0.948	0.986	0.982
Correlation -validation dataset	0.981	0.951	0.981	0.983
MSE - training dataset	0.047	0.129	0.036	0.045
MSE - test dataset	0.045	0.110	0.031	0.039
MSE - validation dataset	0.045	0.114	0.037	0.041

The validation set (Fig. 10) and the complete dataset (Fig. 11) exhibit the graphical comparison results between the model and the real data for this network. The smallest mean square error on the test set was found for the model with the ReLU function and 33-time steps and 33 LSTM units in the hidden layer, and the slightly larger ELU model with 31-time steps and 31 LSTM units in the hidden layer. Finally, the model with the activation function ReLU and 33-time steps and 33 LSTM units in the hidden layer was selected as the best model in terms of the magnitude of the correlation coefficients and the minimum square error.

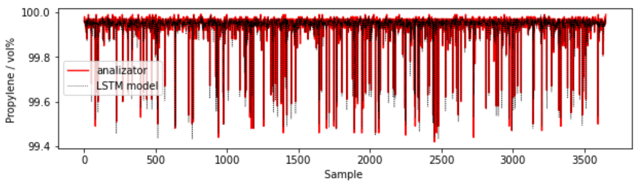

Fig. 10. *Comparison between actual propylene content and LSTM model on validation data for the model with ReLU, 33 LSTM units and 33-time steps*

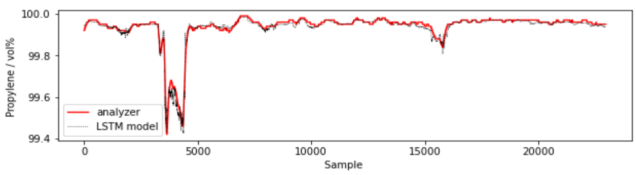

Fig. 11. *Comparison between actual propylene content and LSTM model on the entire data for the model with ReLU, 33 LSTM units and 33-time steps*

In Figure 12, the histogram visualizes the distribution of error values on the entire dataset of the LSTM model, while Figure 13 presents the distribution of error values specifically on the validation data of the LSTM model. By examining both histograms, it becomes evident that the errors predominantly lie within the ± 0.05% range. This finding serves as strong evidence for the exceptional reliability of the developed LSTM model.

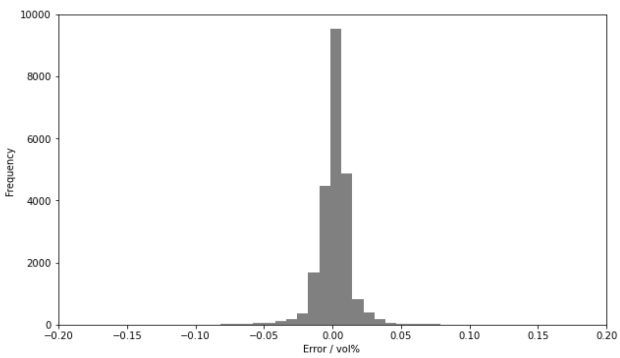

Fig. 12. *Histogram of the error values on the entire data for the LSTM model with ReLU, 33 LSTM units and 33-time steps*

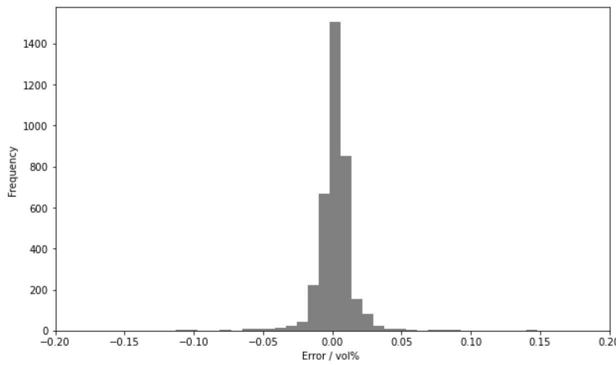

Fig. 13. *Histogram of error values on validation data for LSTM model with ReLU, 33 LSTM units and 33-time steps*

4. Conclusion

In the refinery's propane/propylene splitter plant, it is imperative to consistently monitor the propylene content in the product to enhance process efficiency. The propylene and propane separation process aims to achieve a minimum purity of 99.6 vol% in the propylene, making it advantageous for use in polymer production.

The deployment of a soft sensor for online product quality monitoring can improve process control and reduce expenses associated with the operation and maintenance of online analyzers.

The focus of this study is on the creation of models that incorporate neural networks, such as multi-layered perceptrons and long-short term memory networks. These models are utilized to monitor the propylene content in the product of the PPS plant. The models were created using the Python programming language in the integrated open-source development environment, Spyder.

Among the extensive range of models developed, two models were singled out as the most promising. One model proved to be highly effective in neural networks with multilayer perceptrons, while the other exhibited remarkable capabilities in neural networks with long-term memory. The model utilizing the tanh activation function and consisting of 17 neurons in the hidden layer was chosen as the optimal model among the MLP models because of its superior error stability. The model with the activation function ReLU and 33-time steps and 33 LSTM units in the hidden layer was selected as the best model from the group of LSTM models. Based on both graphical and numerical outcomes, the LSTM model demonstrates superiority as a result of its stronger correlations, reduced errors, and enhanced performance stability when adjusting the hyperparameters.

The most effective models of soft sensors in both categories delivered results that were more than satisfactory, showing similar correlation coefficients and errors in the model output data.

These models prove to be effective for application within the refinery information system. By incorporating these innovative soft sensors, the anticipated results include better control over processes and higher quality final products, ultimately resulting in substantial savings in production costs.

5. References

[1] Ž. Ujević Andrijić, N. Bolf, Primjena softverskih senzora za procjenu kvalitete produkata atmosferske destilacije, Goriva maziva, 50 (2011) 187-200.

[2] N. Bolf, Softverski senzori - alat suvremenog kemijskog inženjerstva, KUI, 60 (2011) 193-199.

[3] M. E. Khan, F. Khan, A comparative study of white box, black box and grey box testing techniques, Int. J. Adv. Comput. Sci. Appl., 3 (2012) 12-15.

[4] I. Ahmad, A. Ayub, M. Kano, I. I. Cheema, Gray-box Soft Sensors in Process Industry: Current Practice, and Future Prospects in Era of Big Data, Processes, 8 (2020) 243.

[5] S. Zendehboudi, N. Rezaei, A. Lohi, Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review, Appl. Energy 228 (2018) 2539–2566.

[6] L. Fortuna, S. Graziani, A. Rizzo, M. G. Xibilia (2007), Soft Sensors for Monitoring and Control of Industrial Processes, London: Springer, 2007.

[7] V.A. Profillidis, G.N. Botzoris (2019) "Chapter 5 - Statistical Methods for Transport Demand Modeling", in: V.A. Profillidis, G.N. Botzoris (Ed.) Modeling of Transport Demand. Analyzing, Calculating, Forecasting Transport Demand, Oxford, Cambridge: Elsevier, 163-224.

[8] Ž. Ujević Andrijić, N. Bolf, Osvježimo znanje: Umjetne neuronske mreže, KUI, 68 (2019) 219-220.

[9] M. Uzair, N. Jamil, Effects of Hidden Layers on the Efficiency of Neural networks, in IEEE 23rd International Multitopic Conference (INMIC), Bahawalpur, Pakistan, 2020 1-6.

[10] S. Sharma, A. Athaiya, Activation functions in neural networks, Int. J. Eng. Appl. Sci. Tech. 4 (2020) 310-316

[11] E. Grossi, M. Buscema, Introduction to artificial neural networks, Eur. J. Gastroenterol. Hepatol. 19 (2008) 1046.-1054.

[12] D. R. Baughman, Y. A. Liu (1995), Neural networks in Bioproceesing and Chemical Engineering, San Diego: Academic Press, 1995.

[13] YL. Cong, LT. Hou, YC. Wu, YZ. Ma, Development of a Coupled EnergyPlus-MATLAB Simulation Based on LSTM for Predictive Control of HVAC System, Math. Probl, Eng., 2022 (2022) 5912967.

[14] S. M. Basha, D. S. Rajput, (2019) "Chapter 9 - Survey on Evaluating the Performance of Machine Learning Algorithms: Past Contributions and Future Roadmap", in: A. K. Sangaiah (Ed.) Deep Learning and Parallel Computing Environment for Bioengineering Systems, St. Louis: Academic Press, Elsevier, 153–164.

[15] A. Krogh, What are artificial neural networks?, Nat Biotechnol. 26 (2008) 195–197.

Želimir Kurtanjek

Causal AI Modelling of Chemical Manufacturing Plants

University of Zagreb, Faculty of Food Technology and Biotechnology, Pierottijeva 6, 10 000 Zagreb, Croatia, retired, zelimir.kurtanjek@gmail.com

Abstract

The concept of "Industry 5.0" is driving significant changes in the production of chemical products and energy, promoting a shift towards a decarbonized and circular economy. Digitalization, robotics, communications, and artificial intelligence (AI) play crucial roles in fostering the development of necessary technological innovations and enhancing intelligent process control. The application of machine deep learning (ML) yields robust, field-neutral solutions for regression prediction objectives, but it is limited in its capacity to address innovative questions that involve causation and counterfactual analysis. This paper presents a proposed application of Bayesian networks (BN) for structural causal modeling (SCM) in the context of manufacturing plants. A critical feature of SCM modeling is its capacity to integrate extensive prior structural knowledge derived from fundamental chemical engineering principles with structures inferred from experimental data obtained from manufacturing plants. The acquired SCM facilitates the forecasting of causal relationships, the simulation of intervention strategies, and the generation of counterfactual responses essential for process innovations and intelligent process management. The SCM model is presented as a tool for examining causality and control in the intricate Tennessee-Eastman process.

Keywords: *Bayes network, causality, DAG, ATE, Markov blanket, Tennessee-Eastman-Process*

1. Introduction

As a discipline of computer science, artificial intelligence (AI) has far-reaching effects on all dimensions of industrial process engineering. It is the key aspect of EU Industry 5.0 policy of providing solutions to social challenges by its impact on industry digitalization and robotics, energy production and management, decarbonization, waste management, circular economy, sustainability, and global green transition [1]. The field of process control in chemical manufacturing is emerging as a promising area of research. The American Institute of Chemical Engineering (AIChE) recognizes the significant impact of AI in chemical engineering research and industry. AIChE highlights that AI, particularly machine learning, is being widely embraced to solve complex problems, accelerate research, and enable computations that were previously impossible1. AI is seen as a powerful tool in applications requiring heavy, iterative computing and large data analysis [2]. AIChE also acknowledges the historical phases of AI in chemical engineering, noting that while AI's promise was not fully realized in the past, the current data science phase is ripe for success.

At the recent congress on Future Innovation in Process System Engineering (FIPSE), the focus was placed on the promising role of machine learning (ML) within process systems engineering (PSE). Its application is expected to revolutionize catalysis, enhance materials design, and improve process operations and automation [3]. There are numerous areas where machine learning techniques can be effectively utilized, including:

– Flowsheet analysis
– Surrogate modelling for simulation and optimization
– Integrated planning and scheduling
– Supply chain design and operation
– Process monitoring and fault diagnosis
– Real time optimization and control

The integration of symbolic causal reasoning reflected in the mathematical equations of core physical and chemical laws with data-driven methodologies is acknowledged as a major challenge and a significant opportunity for future development. The concluding report for the "AI Incubator Labs in the Process Industry," part of the EU project "Knowledge-Empowered Entrepreneurship Network" (KEEN), underscores the importance of combining actual process data, relevant domain knowledge, and AI methodologies during the entire life cycle of chemical plants. The primary goals include enhancing process efficiency, achieving economic advantages, and promoting sustainability. The project addresses three main AI application areas: (1) modelling and simulation of processes, products, and plants; (2) engineering of plants and processes; (3) optimal operation of production plants with the goal of self-optimizing plants [4]. A report from China highlights the significance of AI and robotic automation in chemical synthesis, particularly emphasizing its applications within the pharmaceutical industry. The primary focus is on the implementation of AI for the analysis of structure-function relationships, the strategic planning of synthetic routes, and the automation of the synthesis process [5]. AI support in mathematical chemistry resulted in the new field of digital chemistry and the revolutionary breakthrough in protein engineering. For many years, the complex problem of protein folding remained unresolved until DeepMind achieved a solution with atomic precision, leveraging structured deep neural networks to produce the AlphaFold open-access software tool. [6]. It will contribute to a rapid advancement in vital scientific and engineering disciplines. Another aspect of AI and digital chemistry is based on molecular descriptors. Each molecule can be projected in the space of about 6000 2D numerical descriptors applied for deep neural network processing. Industrially important cases are structural causal models of LPMO enzyme activity for biorefinery and textile waste water treatment, and prediction of deep eutectic properties for green separation processes [7-9]. Nevertheless, the majority of contemporary machine learning techniques rely on patterns derived from large datasets and statistical inference, which results in a lack of causal understanding. While they demonstrate strong predictive capabilities within the assumed static training framework, they fail to offer any mechanistic insights or causal explanations for their decision-making processes. Although this may not pose a significant issue in standard applications of computer vision, gaming, and recommendation systems, it holds considerable importance for various challenges in chemical engineering, including fault diagnosis, process control, and safety analysis. For chemical engineers key questions are on the second (effect of doing) and third rung (probability of counterfactual events) of Pearl's knowledge ladder [10-11]. The objective of this research is to outline the methodology for applying artificial intelligence in structural causal modeling (SCM) within the context of a chemical manufacturing plant, particularly the Tennessee Eastman (TE) challenge process.

2. Modelling

In chemical engineering, mass and energy balance models are typically represented as lumped, continuous, non-stationary, and deterministic systems characterized by ordinary non-linear differential equations (ODE):

$$\frac{dy}{dt} = f(y,x,\theta) \quad y(t=0)=y_0 \tag{1}$$

Bayes models are stochastic given by joint probability density function P

$$model = P(Y,X,\theta) \tag{2}$$

Dynamic Bayes on-line updated inference of the model variables $\{X,Y\}$ is based on prior and likelihood L of old evidence $\{Y_k, X_k\}$· Probability density distribution of the posteriori model is given by

$$\begin{pmatrix} Posteriori \\ model \end{pmatrix} = \begin{pmatrix} Likelihood \\ data \end{pmatrix} \cdot \begin{pmatrix} Prior \\ model \end{pmatrix} / \begin{pmatrix} Data \\ evidence \end{pmatrix} \tag{3}$$

$$P(\theta|Y,X) = \frac{P(Y,X|\theta) \cdot P(\theta)}{P(Y,X)} \tag{4}$$

Learning a Bayes network for a system model involves a two-step approach. The first step focuses on identifying the network structure G, which facilitates the factorization of the joint probability distribution P into separate local distributions P_i. In the second phase, parameters θ are estimated individually by either assuming a probabilistic framework or utilizing nonparametric inference methods such as neural networks or decision trees:

$$\underset{P(G,\theta|data)}{learning} = \underset{P(G|data)}{structure} \cdot \underset{P(\theta|G,data)}{parameter} \tag{5}$$

The structure of the model, represented as G=$\{G_{knowledge}, G_{evidence}\}$, is defined as a directed acyclic graph (DAG). This graph consists of the model variables {Y,X}, which serve as nodes, and the arrows between them indicate direct causal relationships. In this research, $G_{knowledge}$ is derived from the connectivity of streams within the process sheet, along with an understanding of the mass and energy balance frameworks for each individual process unit. The unknown causal patterns $G_{evidence}$ are inferred from data by minimization of Bayes information criteria BIC. It is defined by likelihood L of data X, dimension of sampling space N, and model complexity (number of the graph parameters K).

$$BIC(X|G)=K\cdot ln(N)-2\cdot ln(L(X|G)) \tag{6}$$

Utilizing the hybrid heuristic hill climbing algorithm allows for effective reduction of the Bayesian Information Criterion (BIC) and supports the inference of the most likely graph G.

$$\hat{G} = \underbrace{min}_{i}\,[BIC(X|G_i)] \tag{7}$$

A causally decomposed process state space functions as a Markov system, facilitating a more straightforward inference of the joint probability density function P.

$$P(X) = \prod_{i=1}^{N} P(X_i|par(X_i)) \tag{8}$$

where $par(X_i)$ are the parent nodes which have a direct causal effect on X_i. To achieve process control objectives, it is essential to infer the causal effect on Y when the manipulative variable X is fixed deterministically at a value of x, as indicated by the "doing" function do *(X=x)*. Due to the present interaction between process variables, the evaluation of the effects of intervention requires d-separation of SCM graph. It produces an adjustment subset Z of process variables and offers a deconfounded "backdoor" estimation formula. [11]

$$\begin{aligned} P(Y|do(X)) &= \sum_{z\epsilon Z}^{Z} P(Y|X=x,\, Z=z)P(Z=z) \\ &= \mathrm{P(Y|W,\, Z)} \end{aligned} \tag{9}$$

The average treatment effect (ATE) quantifies the causal impact by measuring the average change resulting from the treatment.

$$ATE[y(x)] = \frac{d}{dx} E_Z[(Y|do(X)] \tag{10}$$

The expected value E_z associated with the covariates reflects the average causal effects calculated from the data within the adjusted control subset.

3. Process model

Causal AI modeling is utilized in the Tennessee Eastman plant-wide process, as proposed by Downs and Vogel (1993). This serves as a benchmark problem for various control-related topics, such as multivariable controller design, optimization, predictive control, estimation/adaptive control, nonlinear control, process diagnostics, and control education. This model represents the behavior of a standard industrial process, featuring a two-phase reactor where an exothermic reaction takes place, in addition to a flash unit, a stripper, a compressor, and a mixer. The process exhibits a nonlinear and open-loop unstable behavior; in the absence of control, it can reach shutdown limits within an hour, even when subjected to minimal disturbances. The TE process, **Fig. 1**., produces two liquid products (G and H) and one (undesired) byproduct F from four gaseous reactants (A, C, D and E), according the reaction stoichiometry [12]

$$\begin{gathered} A + C + D \rightarrow G \\ A + C + E \rightarrow H \\ A + E \rightarrow F \end{gathered} \tag{11}$$

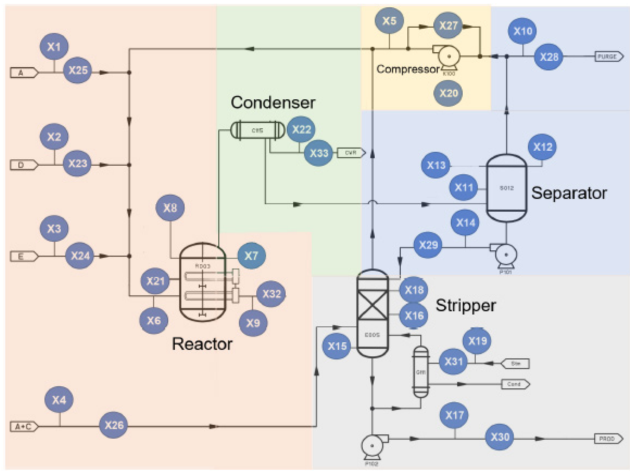

Fig. 1. Tennessee-Eastman process scheme [13-20]

Dynamic mass balances are given by the following ODE :

mixer

$$\frac{dN_{i,m}}{dt} = \sum_{j=1}^{6} y_{i,j} F_j - y_{i,6} \qquad i = A,\cdots H \tag{12}$$

reactor

$$\frac{dN_{i,r}}{dt} = y_{i,6} F_6 - y_{i,7} F_7 + \sum_{j=1}^{3} v_{i,j} R_j \qquad i = A,..\, H \tag{13}$$

separator

$$\frac{dN_{i,s}}{dt} = y_{i,7} F_7 - y_{i,8}(F_8 + F_9) - x_{i,10} F_{10} \qquad i = A,..\, H \tag{14}$$

stripper

$$\frac{dN_{i,p}}{dt} = (1 - \o_i)(x_{1,10} F_{10} + y_{i,4} F_4) - x_{i,11} F_{11} \qquad G, H \tag{15}$$

Heat balances are given by:

mixer

$$\left(\sum_{i=A}^{H} N_{i,m} c_{p.vap,i}\right)\frac{dT_m}{dt} = \sum_{j=1}^{8} F_j \left(\sum_{i=A}^{H} y_{i,j} c_{p.vap,i}\right)(T_j - T_m) \tag{16}$$

reactor

$$\left(\sum_{i=A}^{H} N_{i,r} c_{p,,i}\right)\frac{dT_r}{dt} = F_6\left(\sum_{j=A}^{H} y_{i,6} c_{p.vap,i}\right)(T_6 - T_r) - Q_r - \sum_{j=1}^{3} \Delta H_{r,j} R_j \quad (17)$$

separator

$$\left(\sum_{i=A}^{H} N_{i,s} c_{p,,i}\right)\frac{dT_s}{dt} = F_7\left(\sum_{i=A}^{H} y_{i,7} c_{p.vap,i}\right)(T_r - T_s) - Q_s - H_0 V_s \quad (18)$$

striper

$$\left(\sum_{i=G}^{H} N_{i,p} c_{p,,i}\right)\frac{dT_p}{dt} = F_{10}\left(\sum_{i=A}^{H} x_{i,10} c_{p,,i}\right)(T_s - T_p) + F_4\left(\sum_{i=A}^{H} y_{i,4} c_{p.vap,i}\right)(T_4 - T_p) - H_0 V_p + Q_p \quad (19)$$

The system is a lumped MIMO DAE model consisting of 30 nonlinear state y differential ordinary differential equations (ODEs) and 149 algebraic equations x, incorporating 10 input variables and 130 physical and chemical parameters θ. Simulation software is included within the MatLab Simulink software support framework [18]. The results of these simulations can be accessed through Harvard Dataverse and GitHub repositories [17,20].

4. Results and discussion

The analysis utilizes a dataset comprising 50,000 randomly selected samples of process variables, collected under normal operating conditions during a simulated 48-hour continuous operation. The data includes 12 manipulative control variables (exogenous) and 35 continuous variables (endogenous), as illustrated in Fig. 1 [19]. The data exhibit a Pearson correlation coefficient of R=0.22, indicating a moderate level of correlation. The focus of the analysis is on predicting and understanding the causal effects of process variables on the production rates of G and H. For numerical and statistical evaluation the R language and "q*eML*" machine learning tool wrapper are applied [21-23].

The "random forest" algorithm was utilized to assess the significance of variables by measuring the improvement in the predictive accuracy of G product's productivity. The results of the variable relative importance are depicted in **Fig. 2**. The reactor pressure, separator temperature, and stripper pressure account for approximately 95% of the overall significance in predicting G productivity. Because of the confounding effects introduced by multivariate correlation, the importance factors identified do not reflect the direct causal relationships of the variables with G production. Utilizing a complete set of 34 predictor variables, the "random forest" machine learning model has shown impressive predictive accuracy, as evidenced by tests conducted on 8,000 trained samples and 2,000 untrained samples, **Fig. 3**. The model successfully accounts for 95% of the data variance, with mean relative absolute errors of 1.9% for the trained data and 2.2% for the untrained data related to G production.

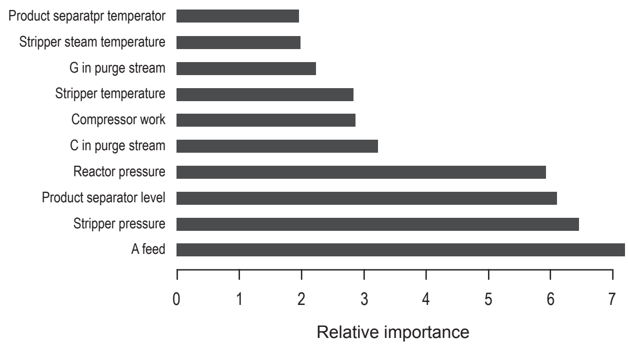

Fig. 2, Variable importance for the ML model prediction of *G* production.

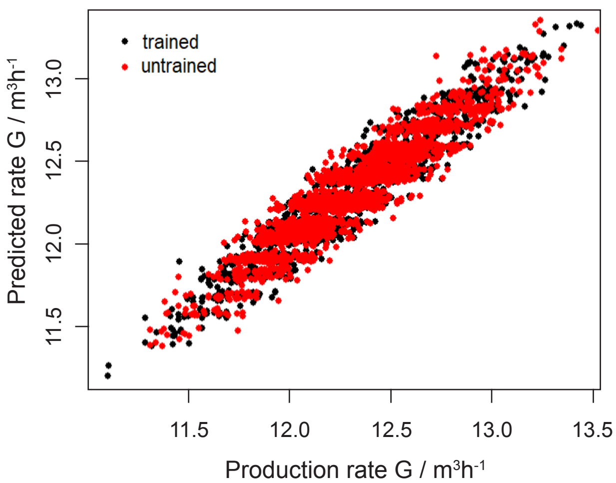

Fig. 3. Comparison of the random forest model (RF) predictions and values of G purge rate.

While the RF model demonstrates a high level of prediction accuracy, its agnostic nature—implying a presumed absence of domain knowledge—renders it unsuitable for do*(X=x)* analysis of causal relationships. This analysis is crucial for making informed decisions regarding process interventions, control policies, and optimization strategies.

The studied TE serves as a model for a manufacturing production system, grounded in the principles of chemical engineering, resulting in highly detailed and precise field knowledge (eq. 12-19). Deterministic direct causal relationships are established through functional forms of mass and energy balances, stoichiometric relationships of chemical reactions, and causal connections derived from the design of process synthesis, as represented in a process sheet or graph. In practice, certain process parameters are estimated with inherent uncertainties, while others remain unknown, particularly kinetic parameters. Additionally, some interactions may be unforeseen, such as catalyst poisoning, and there are process disturbances that go unobserved. The Bayes network AI model is proposed which accounts for causal effects of stochastic nature present in an industrial process. The known structural causality, given as graph $G_{knowledge}$, is integrated with data inferred causality G_{data}.

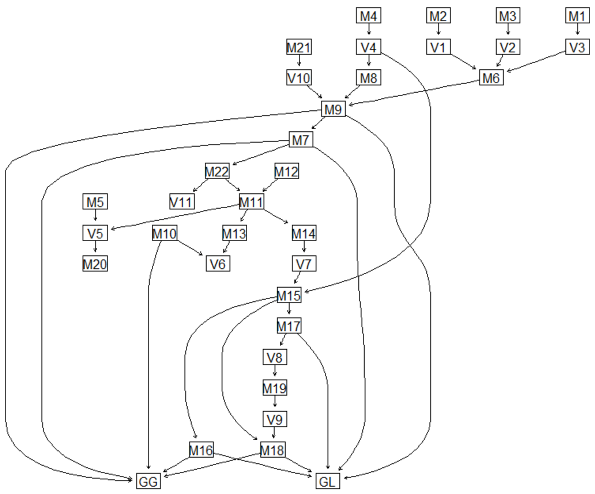

Fig. 4. Directed acyclic graph (DAG) of TE process. M are measured state variables, manipulative variables V, product G in liquid and gas purge are GL and GL M9 and M7 are reactor temperature and pressure.

Table 1. Labels of the process flow sheet

M1	A feed (stream 1)
M2	D feed (stream 2)
M3	E feed (stream 3)
M4	A and C feed (stream 4)
M5	recycle flow (stream 8)
M6	reactor feed rate (stream 6)
M7	reactor pressure
M8	teactor level
M9	reactor temperature
M10	purge rate (stream 9)
M11	product separator emperature
M12	product separator level
M13	product separator pressure
M14	product separator underflow (stream 10)
M15	stripper level
M16	stripper pressure
M17	stripper underflow (stream 11)
M18	stripper temperature
M19	stripper steam flow
M20	compressor work
M21	reactor cooling water outlet temperature
M22	separator cooling water outlet temperature
V1	D feed flow (stream 2)
V2	E feed flow (stream 3)
V3	A feed flow (stream 1)
V4	A and C feed flow (stream 4)
V5	compressor recycle valve
V6	purge valve (stream 9)
V7	separator pot liquid flow (stream 10)
V8	stripper liquid product flow (stream 11)
V9	stripper steam valve
V10	reactor cooling water flow
V11	condenser cooling water flow

Table 2. Canonical adjusted sets of the control variables for product G_{liquid} and Markov blankets of the reactor variables.

Adjusted control variables for product G_{liquid}	
M9: reactor temperature	M1, M12, M2, M21, M3, M4, M6, M8, V1, V10, V2, V3, V4
M7: reactor pressure	M1, M12, M2, M21, M3, M4, M6, M8, M9, V1, V10, V2, V3, V4
Markov blankets for the reactor	
M9: reactor temperature	M10, M16, M17, M18, M6, M7, M8, V10, GG, GL
M7: reactor pressure	M10, M16, M17, M18, M22, M9, GG, GL

Integrated structure $G = \{G_{knowlrdge}, G_{data}\}$, N observed data, and K number of the model parameters are jointly evaluated by minimization of the corresponding Bayes information criteria, BIC, eq. 6. Applied is hybrid heuristic MMHC algorithm [25-26]. To identify the G skeleton, MMHC employs tests of conditional independence, seeking variable subsets Z that render a pair of variables X and Y conditionally independent. They are inferred from the corresponding conditional independence as proposed in the stable PC algorithm [26]. The application of heuristic local optimization in the G search space results in the most probable posterior Bayes network model, leveraging known prior causal structure and causal relationships inferred from the data observed (eq. 7). The obtained model is the DAG graph presented in **Fig. 4**. The model accounts for the measured state variables (M), manipulative variables (V), and production rate of product G in gas purge stream and liquid striper underflow (GG, GL). The structure of the DAG model achieved allows for d-separation, effectively preventing "back door" confounding of causal effects. The adjusted sets Z of control variables, which account for the dependence of GL on reactor temperature and pressure, are presented in Table 2. The corresponding Markov blanket sets yield the paternal, children, and children parent nodes by which the reactor temperature and pressure are separated from the perturbations of the rest of process variables. Pair-wise causal dependencies are depicted in **Fig. 5**. The graphs shown represent partial dependency plots generated by Bayes neural networks, which were trained on the relevant adjusted Z sets. The method applies single inner layer network configurations featuring sigmoid activations, with backpropagation used for the learning mechanism. The graphs are traces of individual Y[do(x)] where Y is the molar percentage of the product G in the separator underflow liquid stream, and x are the following process variables: compressor recycle % opening, E feed stream, input mixed A and C stream, and D feed rate. The developed model facilitates the estimation of indirect causal relationships that remain obscured due to implicit nonlinearities, making them not readily apparent from prior structural knowledge. The observed dispersions of the causal plots are due to the effects of variability of the rest of variables from the corresponding adjusted control sets. Bivariate effects of synergism and antagonism are depicted in Fig. 6. The plots are two-dimensional partial dependency plots of the estimated expected values Ez[Y(x1,x2)] where x1 and x2 are selected variables from the adjusted control set of Y and averaging E covers the complimentary set Z/(x1,x2). The f results facilitate the visualization of causal synergism within a system characterized by opaque and intricate nonlinear interactions.

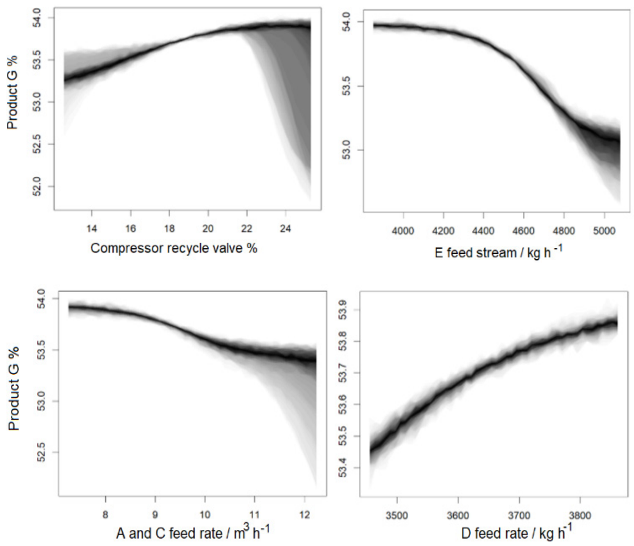

Fig. 5. Bayes NN predictions of causal dependances of the production rate of G in liquid outlet

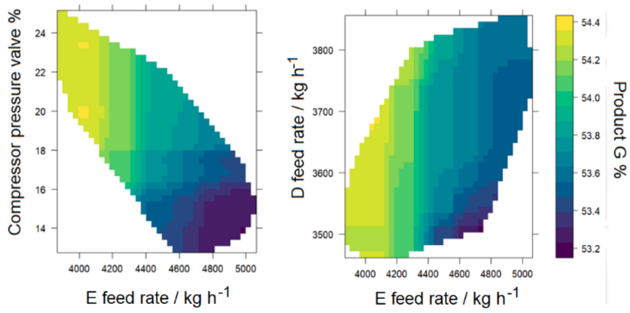

Fig. 6. Bayes NN bivariate causal partial plots of the dependency of molar G% in liquid outlet stream on the key process variables.

The visualization of causal interactions provides basis for design of multilevel control system, process optimization, and synthesis of alarm system for fault detection.

5. Conclusions

The causal artificial intelligence model offers engineers insights into the impacts of decisions and interventions within a manufacturing system, particularly under novel and previously unobserved conditions. Formally, it infers numerical functions of ATE (average treatment effect, doing $Y[do(X=x)]$) and/or CATE (conditioned average treatment effect, doing $Y[do(X=x)|Z=z)]$) effectsThe improvement of agnostic machine learning models is of utmost importance, given that their regressive property limits their ability to make predictions solely based on conditions that have been previously encountered. The modeling of causality utilizes Bayesian networks, which facilitate an advanced level of artificial intelligence and enable causal inferences that align with the second tier of Pearl's knowledge ladder[11]. The primary characteristic of causal Bayes networks is their structural representation of direct causal relationships, which adhere to the Markovian property.

In the context of industrial systems, structural knowledge is largely grounded in causal structural equations that originate from basic principles of mass and energy conservation, as well as the schematics of designed process flows. It is possible to infer unobserved or unknown interactions by merging a priori structural insights with statistical conditional analysis or by optimizing the Bayesian Information Criterion (BIC) in the context of causal networks. The representation of system structural knowledge takes the form of a directed acyclic graph (DAG) that includes process variables, with network nodes linked by direct causal connections indicated by arrows. DAG networks exhibit Markovian properties, allowing the joint probability function to be decomposed into the products of the probability distributions of simple parent nodes. Unknown physical and chemical parameters are incorporated within the node regression parameters and/or weights of neural networks, utilizing substantial data collected from process monitoring..

The methodology and analysis of causal artificial modeling are utilized in the examination of the chemical synthesis process at the Tennessee Estman industrial complex. The process in question is a catalytic, two-phase (gas and liquid) system that exhibits instability. Initial structural understanding is obtained through basic mass and energy balance calculations, as well as the relationships between mass and heat streams specified in the process flow sheet. Unobserved causal relationships can be effectively identified through the heuristic minimization of the Bayesian Information Criterion (BIC) within the model. This causal model provides unconfounded estimates regarding the importance of process variables, utilizing adjustment sets as control variables. It also incorporates the Markov blankets for reactor pressure and temperature, along with average treatment effects (ATE) and the synergistic interactions between process variables.

Causal AI models offer vital information that is crucial for the optimization of processes, the innovation of operational methods, the development of intelligent control systems, and the support of decision-making strategies in manufacturing [27].

From a chemical process engineering standpoint, the main difference in the applicability of causal Bayes networks versus first principle models, like mass and energy balances, is their consideration of stochastic effects. Causal Bayes networks effectively address both exogenous variables (process feeds) and endogenous variables (generated by the process), particularly in terms of kinetics and mass/energy transfer coefficients. In summary, fundamental mathematical models serve as the foundation for process design, whereas Bayes causal AI models concentrate on process control and decision-making policies for interventions in uncertain environments.

6. References

[1] URL: https://research-and-innovation.ec.europa.eu/ research-area/industrial-research-and-innovation/ industry-50_en

[2] Venkatasubramanian V., The promise of artificial intelligence in chemical engineering, AIChE Journal 65 (2019) 466-478

[3] Daoutidis P., Lee J.H., Rangarajan S., Chiang L., Gopalani B., Schweidtmann A., Harjunkoski I., Mercangoz M., Mesbah A., Boukouvala F., Limo F.V., Chanona A., Georgakis C., Machine learning in process systems engineering: Challenges and opportunities, Comput Chem Eng, 181 (2024) 108523

[4] Bortz M., Dadhe K., Engell S., Gepert V., Kockmann N., Muller Pfefferkon R., Schindler T., Urbas L, AI in Process Industries – Current Status and Future Prospects Chem Eng Tech 95 (2023) 975-988

[5] He, C., Zhang, C., Bian, T., Jiao, K., Su,W., Wu, K, Su, A. A Review on Artificial Intelligence Enabled Design, Synthesis, and Process Optimization of Chemical Products for Industry 4.0. Processes 11 (2023) 330

[6] Jumper J., Evans R., Pritzel A., Green T, Figurnov M., Ronneberger O., Tunyasuvunakool K., Bates R., Žídek A., Potapenko A., Bridgland A., Meyer C., Kohl S., Ballard A., Cowie A., Romera-Paredes B., Nikolov S., Adler J., Back T., Petersen S., Reiman D., Clancy E., Zielinski M., Steinegger M., Pacholska M., Berghammer T., Bodenstein S., Silver D., Vinyals O., Senior A.W., Kavukcuoglu K., Kohli P., Hassabis D., Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583.

[7] Rezić, I., Kracher, D., Oros, D., Mujadžić, S., Anđelini, M., Kurtanjek, Ž., Ludwig, R., Rezić, T., Application of Causality Modelling for Prediction of Molecular Properties for Textile Dyes Degradation by LPMO, Molecules,19 (2022), 1-13

[8] Rezić T., Vrsalović Presečki A., Kurtanjek Ž., New approach to the evaluation of lignocellulose derived by-products impact on lytic-polysaccharide monooxygenase activity by using molecular descriptor structural causality model, Bioresour Technol, 342 (2021) 125990

[9] Kurtanjek Ž., Molecule Structure Causal Modelling (SCM) of Choline Chloride Based Eutectic Solvents, Chem Biochem Eng Q 36 (2022) 223–230

[10] Efron B., Hastie T., Computer Age Statistical Inference – Algorithms, Evidence, and Data Science, Cambridge University Press, New York. 2016

[11] Pearl J., Mackenzie D., The Book of Why: The New Science of Cause and Effect, Penguin Books, Harlow, 2019, UK

[12] Downs, J.J., and Vogel, E.F. A Plant Wide Industrial Process Control Problem, Comput. Chem. Eng., 17 (1993) 245–255.

[13] Jockenhövel T., Biegler L.T., Wächter A., Comput. Chem. Eng., 27 (2003), 1513-1531.

[14] Sales-Cruz, M., Cameron, I., & Gani, R. Tennessee Eastman Plant-wide Industrial Process Challenge Problem. In I. Cameron, & R. Gani (Eds.), *Product and Process Modelling: A Case Study Approach* 9 (2011) 273-303.

[15] Ricker, L.N. Tennessee Eastman Challenge Archive, http://depts.washington.edu/ctrol/LARRY/TE/download.html#Decentralized_control (2005)

[16] Capaci, F., Vanhatalo, E., Kulahci, M., & Bergquist, B. The revised Tennessee Eastman process simulator as testbed for SPC and DoE methods. Qual. Eng., 31 (2019) 212-229

[17] R. Braatz, https://github.com/camaramm/tennessee-eastman-profBraatz (2023)

[18] Jenny, O.E., MATLAB Central File Exchange. TE process (https://www.mathworks.com/matlabcentral/file exchange/39742-te-process), 2023.

[19] Menegozzo, G., Dall'Alba, D., & Fiorini, P. CIPCaD-Bench: Continuous Industrial Process datasets for benchmarking Causal Discovery methods. 2022 IEEE 18th International Conference on Automation Science and Engineering (CASE), (2022). 2124-2131.

[20] Cory R, A., Amsel, B. D., Randy T.,; Cook, M. B., "Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation", Harvard Dataverse, V1 https://doi.org/10.7910/DVN/6C3JR1, (2017),

[21] R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ (2023).

[22] RStudio Team RStudio: Integrated Development for R. PBC, Boston, MA URL http://www. rstudio.com/ (2020)

[23] Matloff N., qeML, https://github.com/matloff/qeM (2023)

[24] Kurtanjek Ž., Kada zaključivanje matematičkim modelom može biti pogrešno: Primjer protočnog kemijskog reaktora PKR, Kem. Ind. 70 (2021) 739–741

[25] Tsamardinos, I., Brown L.E., Aliferis C.F., The max-min hill-climbing Bayesian network structure learning algorithm, Mach Learn 65 (2006) 31–78

[26] Scutari M. Learning Bayesian Networks with the bnlean R Package. J. Stat. Softw., 35 (2010), 1-22URL http://www.jstatsoft.org/v35/i03/

[27] Pu H., Chen Z., Liu J., Yang X., Ren C., Liu H., Yian Y., Research on decision-level fusion method based on structural causal model in system-level fault detection and diagnosis, Eng. Appl. Artif. Intell., 126 (2023) 107095

Frane Čačić Kenjerić[1]

AI and Machine Vision in Food Processing

[1]University Josip Juraj Strossmayer of Osijek, Faculty of Food Technology Osijek, Franje Kuhača 18, 31 000 Osijek, Croatia (fcacic@ptfos.hr)

Abstract

As a global and vital part of modern life, the food process industry continues to evolve and develop, adopting newly developed tools and technologies to improve efficiency and ensure food safety while minimizing environmental impact and striving towards sustainability. Food processing is essential for preparing and preserving food for an ever-growing population. Artificial intelligence and machine vision technologies have emerged as promising tools in the food process industry, offering opportunities for enhanced quality control, increased productivity, and improved traceability. This article discusses the implementation of artificial intelligence and machine vision in the field of food process engineering, showcasing how these technologies can optimize processes, improve quality control, and contribute to sustainability in food production.

***Keywords**: Artificial intelligence, machine vision, food processing, Artificial neural networks.*

1. Introduction

The food industry plays a crucial role in supplying the ever-growing global population with safe and nutritious food. The human population has reached 8 billion people on November 15. 2022 and World Population Prospects 2022 published by the United Nations Department of Economic and Social Affairs predicts that by 2058 there will be 10 billion people on the planet [1]. With climate changes resulting in extreme weather, the food industry will be tested to find novel foods and technologies to satisfy the increased demand for food. Traditionally, food processing has relied heavily on thermal processing techniques such as pasteurization and canning to ensure food safety and extend shelf life. However, these processes require lots of energy and can result in nutrient loss and changes in the sensory properties of food products. Moreover, traditional food processing methods may not be sufficient to meet the growing demand for diverse and convenient food products [2]. To reduce energy consumption, and environmental footprint and increase efficiency and sustainability as well as food quality and shelf-life, the food industry is turning toward novel food processing technologies employing high hydrostatic pressure, pulsed electric fields, irradiation, ultrasonication, cold plasma, hydrodynamic cavitation, microwaves, radio frequency heating, ohmic heating, ozone treatment and supercritical fluids such as carbon dioxide and water [3].

Artificial intelligence (AI), a term that was used first by John McCarthy in 1955, in his proposal for a summer research project on the concept of thinking machines, which was held at Dartmouth College in 1956 gathering leading minds in computer science and cognitive psychology, to create a machine capable of performing tasks that would typically require human intelligence [4]. Researchers have underestimated the complexity of such tasks and progress in the field was slower than expected, resulting in the "AI winter" from the late 1960s to the late 1990s. During this period, AI research and development faced significant challenges and funding was reduced. However, during the late 1980s and early 1990s, there was a resurgence in AI research and development, boosted by the growing processing power of digital computers and progress in computer science. In recent years, there has been significant progress in AI research, driven by advancements in machine learning and various related disciplines. These advancements have opened up new possibilities for the food industry, particularly in the field of food processing [5][6]. The purpose of this paper is to outline the various applications of AI and machine vision technologies in the food processing industry.

2. Artificial Intelligence

The field of artificial intelligence has evolved within computer science, concentrating on the creation of systems that can carry out functions traditionally associated with human intelligence. This involves various skills such as solving problems, identifying patterns, comprehending natural language, and acquiring knowledge through experience. Artificial neural networks (ANN) serve as the cornerstone of artificial intelligence, functioning as computational models that draw inspiration from the architecture and operations of the human brain [7]. ANNs are made from interconnected artificial neurons organized in layers (**Fig. 1a.**). Input layer which receives raw input data, where each neuron corresponds to one of the features of input data. Hidden layers of neurons (one or more), are located between input and output layers on ANN. The neurons in this layer receive inputs from the preceding layer through weighted connections, utilizing an activation function to generate outputs that are transmitted to the subsequent layer (Fig. 1b.). The output layer, which is the last layer of the artificial neural network (ANN), generates the network's output. Each neuron in this layer corresponds to a specific class or value that the ANN aims to recognize or forecast.

Each neuron within these layers has associated weights and biases that are adjusted during the training process. The neuron's activation function defines how it reacts to the sum of the weighted inputs. Common activation functions include the sigmoid, tanh, ReLU, and softmax [8].

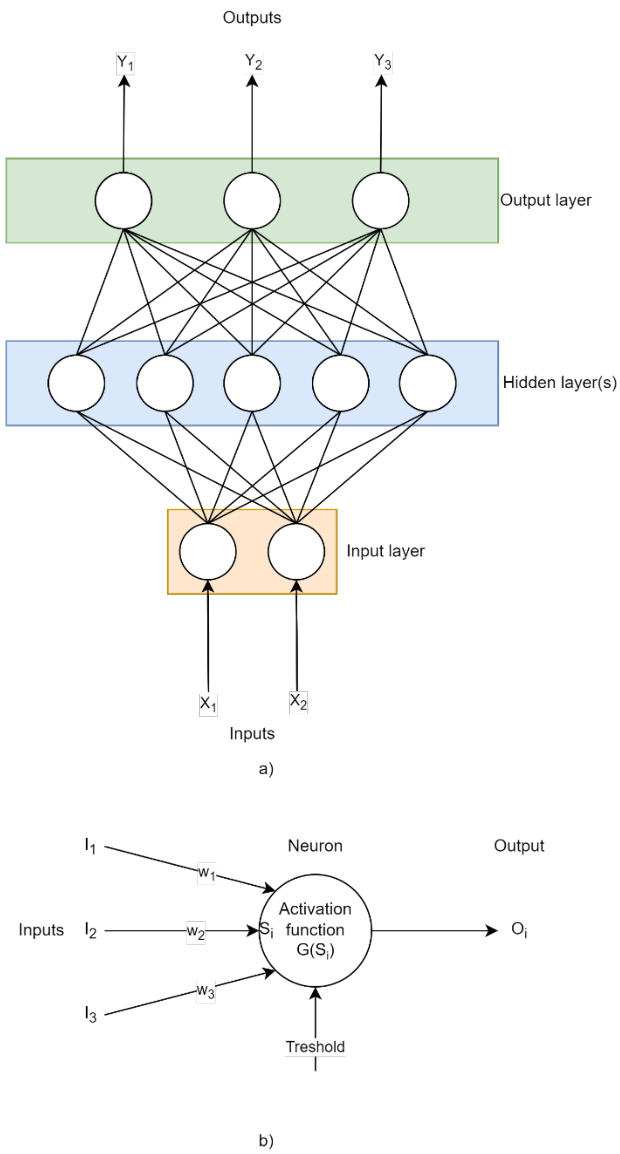

Fig. 1. Generalized structure of ANN (a) and artificial neuron (b)

Neuronal connections facilitate the transfer of output from one neuron to become the input for another neuron. The strength and sign of these connections are determined by the weights, which are adjusted during the training phase so that the network can learn to make accurate predictions or decisions. This training usually involves a process known as backpropagation, where the error between the predicted output and the actual output is computed and used to update the weights of the network through an optimization algorithm such as gradient descent [9][10].

The structure and design of an ANN play a crucial role in its ability to learn and perform tasks accurately.

Typically, AI systems are categorized into two types weak or strong AI [11][12]. Weak AI, or narrow AI refers to AI systems that are designed and trained for a specific task, such as image recognition or speech recognition. Conversely, strong AI, also known as general AI, refers to a machine that possesses the capability to utilize intelligence across a wide range of problems, rather than being limited to addressing a single specific issue. This variant of AI can grasp, learn, and implement knowledge in various areas, yet it has not been completely realized and is still the focus of continuous research. The field of AI includes a range of subfields that target different components of human intelligence, with the key ones being:

- Machine learning relies on algorithms to enable machines to learn from data and make predictions. Learning can be done with supervised, unsupervised, or reinforcement learning techniques. Recently, deep learning techniques that utilize deep neural networks—characterized by multiple hidden layers—have surfaced as effective tools for modelling intricate patterns within data. [13].
- Computer vision systems have a goal to enable machines to perceive and understand visual information such as images or videos. These systems use image processing algorithms to extract meaningful features from the visual input and can be applied in various tasks such as object recognition, image classification, and machine vision [14][15]. These systems have wide applications in various industries.
- The field of robotics emphasizes the design and development of machines that can physically interact with their environment. These machines are equipped with sensors, actuators, and algorithms that allow them to perceive their surroundings, make decisions, and carry out tasks autonomously [16][17].
- Natural language processing aims to enable machines to understand and generate human language. This field involves the development of algorithms and models that can process and analyse text, speech, and other forms of language data [18][19].
- Expert systems are AI systems that emulate the decision-making ability of human experts in a specific domain. These systems use knowledge representation and reasoning techniques to provide solutions and recommendations based on their expertise [20].

AI aims not only to simulate intelligence but also to extend human capabilities by processing large amounts of data at a speed that far exceeds what humans can do. AI applications are diverse and span multiple industries, including healthcare, transportation, finance, manufacturing and, as your document discusses, food processing.

3. Machine Vision Systems

Machine vision systems are being incorporated in various industries, including the food industry, typically in conjunction with machine learning or ANN to enhance automation and improve quality controlMachine vision systems typically consist of two main components: the acquisition of images and the processing of those images. The acquisition phase involves various sensor types designed to capture images of food products [21] complemented by suitable lighting sources. The image processing phase involves applying different methods, ranging from statistical to machine learning and deep learning models to effectively analyse and interpret the captured images.

Machine vision systems use various types of sensors to capture images for analysis. The choice of sensor depends on the specific requirements of the application, such as resolution, speed, sensitivity, and environmental conditions. Several kinds of sensors are frequently employed in machine vision systems:

- Charge-Coupled Device Sensors: These sensors are known for their high-quality images and excellent light sensitivity. CCD sensors are commonly used in applications requiring precise measurements, inspection, and high-resolution imaging [22].
- Complementary Metal-Oxide-Semiconductor Sensors: CMOS sensors are generally more cost-effective and consume less power compared to CCD sensors. They are capable of faster processing speeds and are used in applications where high frame rates and integration with on-chip processing circuits are required [22].
- Infrared Sensors: These sensors capture images based on infrared light, which is not visible to the human eye. They are useful for applications in low-light conditions or where temperature differentiation is important [23].
- X-ray Sensors: Used for inspection purposes where penetrative imaging is necessary, such as detecting flaws inside metal parts or inspecting packaged goods for contaminants [24][25].
- Thermal Imaging Sensors: Capture images based on the heat emitted by objects. They are used in applications ranging from medical diagnostics to industrial inspection where temperature variations need to be monitored [26][27][28].
- Line Scan Sensors: Instead of capturing a whole image at once, these sensors capture data line by line to create an image. Line scan cameras are suitable for inspecting objects moving at high speeds on a production line, such as webs of materials or cylindrical parts [29].
- Time-of-Flight Sensors: These use the time it takes for light to travel to an object and back to calculate dis-

tance. ToF cameras are useful in 3D imaging and can help with object recognition, volume measurement, and collision avoidance in robotic applications [30][31][32].

- 3D Sensors: Utilize various technologies, such as laser triangulation or stereovision, to create three-dimensional images of objects. They are used in complex inspection tasks where depth information is critical [33][34].
- Ultraviolet Sensors: UV cameras can capture images using ultraviolet light, which can be used to highlight certain features that are not visible with standard lighting conditions [35].
- Multispectral and Hyperspectral Sensors: Capture image data at specific frequencies across the electromagnetic spectrum. These sensors can detect chemical composition or moisture content and are used in applications such as agricultural monitoring [36][37][38].

The importance of lighting in machine vision cannot be overstated, as it plays a critical role in the effectiveness of the imaging system. Proper lighting is essential for capturing high-quality images that are needed for accurate analysis. The choice of light source will depend on the specific requirements of the application [39]. Several pivotal outcomes can be attained by implementing efficacious illumination within machine vision systems, underscoring its significance:

- Feature Enhancement: Lighting can be configured to emphasize specific features on the subject being imaged, such as edges, colours, or textures, which are critical for accurate detection and measurement.
- Consistency: Consistent lighting ensures that images are captured with uniform brightness and contrast, which is essential for reliable comparison and analysis, especially in applications where precise measurements or consistent quality checks are required.
- Contrast: Proper lighting enhances the contrast between the object and its background, facilitating easier identification and processing by the vision system.
- Image Clarity: Good lighting reduces shadows and glare that can obscure details and degrade the quality of the image. With clear images, machine vision systems can detect defects, sort products, and guide robots with greater precision.
- Speed and Efficiency: Adequate lighting allows for faster shutter speeds, which is crucial for imaging fast-moving objects on production lines without motion blur, thus improving the throughput and efficiency of industrial processes.
- Reduction of Noise: Optimal lighting conditions enable the camera to operate with lower ISO settings and shorter exposure times, reducing the amount of noise in the captured image and increasing the signal-to-noise ratio.
- Safety and Non-destructive Inspection: In some applications, the correct lighting can allow for the safe and non-destructive inspection of products, such as using X-ray or infrared light to inspect packaged goods without opening them.
- Flexibility: By manipulating lighting conditions, a machine vision system can be adapted to different tasks and environments, making the system more versatile and capable of handling a wide range of inspection duties.

In summary, lighting in machine vision is pivotal in ensuring that the system performs as expected. It affects virtually every aspect of the image-capturing process, thereby influencing the success of applications that rely on machine vision technology [40].

After image acquisition, machine vision systems typically include several major steps (**Fig. 2.**) for processing and interpreting visual information: image pre-processing, segmentation, feature extraction, pattern recognition/classification, and decision-making [39].

Image pre-processing includes various techniques to enhance the quality and clarity of acquired images before further analysis [41]. The most common methods for image pre-processing include grayscale conversion, noise reduction, histogram equalization, image enhancement, and image normalization [39].

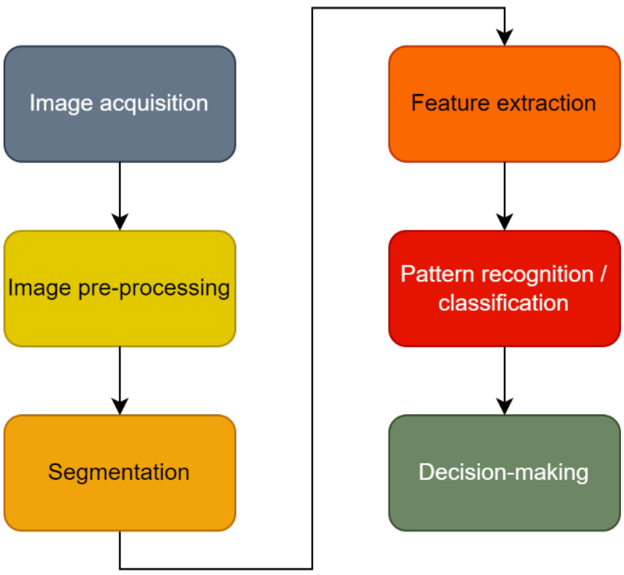

Fig. 2. Major steps for processing and interpreting visual information in machine vision systems.

Segmentation is the process of partitioning an image into meaningful regions or objects to simplify further analysis. Segmentation techniques determine the boundaries and regions within an image, allowing for more efficient and accurate recognition of objects or features of interest. Segmentation techniques can be based on intensity thresholds, edge detection, region growing, or clustering algorithms [42][43][44][45]. Feature extraction involves extracting relevant visual features from the segmented regions, which can include shape, texture, colour, or spatial relationships [46][47]. These features are then used to characterize and describe the objects or regions of interest in the image. Pattern recognition/classification is the step in which machine learning algorithms or other mathematical methods are applied to identify and classify objects or patterns based on the extracted features from the previous steps [48]. This step involves comparing the extracted features with a pre-defined set of criteria or patterns to determine the class or category to which an object belongs. Decision-making is the final step in the machine vision process, where a decision or action is made based on the classification or recognition results [39].

4. Application of AI and Machine Vision in the Food Processing Industry

The food processing industry, similar to other processing sectors, relies on obtaining raw materials and converting them into final food products. However, unlike in other processing industries, the raw materials in the food industry are perishable and susceptible to spoilage [49]. This underscores the importance for the food processing industry to optimize its operations and ensure quality control throughout the production process [50]. The strategy of managing the food chain is often described as "from farm to table" or from "field to fork," highlighting the critical need for maintaining strict time frames for ensuring quality and safety during food processing from harvesting or production until it reaches consumers' plates. Artificial intelligence has emerged as a powerful tool in the food processing industry to address these challenges and enhance various aspects of the production process such as:

- Food quality and safety determination (raw materials): AI systems can analyse and evaluate the quality of food products based on various parameters such as appearance, texture, taste, and aroma [51]. This can help ensure that only high-quality and safe raw materials are used in the production process. Also, AI can be used to monitor and inspect finished food products, ensuring that they meet the required quality standards.
- Control tools: AI can be used to monitor and control various aspects of food processing operations, such as temperature, humidity, pressure, and other critical variables. This ensures that the production process is optimized for efficiency and consistent quality.
- Food processing: AI can optimize and automate various processes in food processing, such as ingredient mixing, cooking, packaging, and labelling, to improve efficiency and consistency [52][53].
- Food sorting and Packaging: AI can automate the sorting process of food products based on quality, size, colour, or other characteristics [51].
- Predictive maintenance and optimization of food processing equipment: AI can analyse data collected from sensors and machines to predict when maintenance or repairs are needed, reducing downtime and maximizing equipment efficiency [51][54].
- Sales forecasting and supply chain management: AI can analyse historical data, market trends, and consumer behaviour to accurately forecast sales demand and optimize the supply chain, ensuring efficient production and minimizing wastage [55][56].
- Generally, AI can be employed in almost all parts, if there is enough data for model development, the problem can be reduced to optimization, classification, pattern recognition or decision-making. These applications of AI in the food processing industry have the potential to significantly improve efficiency, quality, and safety throughout the entire production process, ultimately leading to higher customer satisfaction and a more sustainable food industry [55]. Furthermore, the integration of machine vision in food process engineering allows for real-time monitoring and analysis of food products [57]. This enables quick detection and response to any quality or safety issues, ensuring that only safe and high-quality products are delivered to consumers. In summary, AI and machine vision have numerous applications in food process engineering, including improving traceability, detecting contaminants, ensuring employee safety and hygiene, optimizing production processes, and enhancing overall quality and efficiency. In recent years, machine learning and machine vision have played a crucial role in enhancing various aspects of the food processing industry [47][58].

5. Challenges and Opportunities in AI-Driven Food Engineering

Artificial intelligence-driven food engineering has the potential to revolutionize the food industry by improving productivity, efficiency, and quality. But, as with all new technologies, there are some challenges and opportunities that need to be addressed [21].

One of major the challenges in implementing AI-driven food engineering is the availability and quality of data. High-quality data is crucial for training AI models and ensuring accurate results. Advanced artificial intelligence models require extensive data sets, posing challenges in terms of acquisition, storage, and processing. This also results in increased expenses and complexity during the implementation phase. AI models also require diverse and representative data to avoid bias and ensure fairness in their predictions [59]. Designing artificial intelligence models capable of managing the inherent variability in food products and processes presents a significant challenge. Initial investment and infrastructure requirements for implementing AI are also challenges and might be barriers for small and medium-sized enterprises that need to be addressed [60].

A related issue that arises is the lack of skilled professionals who possess knowledge in both the technological and food science dimensions of AI-driven food engineering[51]. Integrating AI technology into established food processing systems can present challenges and often demands considerable adjustments or upgrades [21]. Complying with stringent food safety and quality regulations when implementing AI technologies can be difficult. It is common for AI-driven systems to exhibit limited interpretability, posing difficulties in comprehending and articulating the reasoning that underlies their decisions [51]. AI recently received a lot of media attention and raised many questions which can result in scepticism or ethical concerns among consumers regarding AI in food production, affecting market acceptance.

Despite these challenges, there are also significant opportunities in AI-driven food engineering. Utilizing AI technology in the food industry can result in improved productivity and greater efficiency [57]. It can automate repetitive complex tasks such as sorting, grading and packaging of food by enabling machines to handle products with variability [61][62][63][5][64]. AI can also optimize processing parameters, leading to improved quality and reduced waste [57]. Enhanced quality control can be achieved with AI systems capable of detecting defects, contaminants, and foodborne pathogens in real-time, ensuring safer products for consumers [65] [6]. Additionally, AI can help in product development by analysing consumer preferences and trends, allowing for more personalized and innovative food options [6][66] [67]. Furthermore, AI can assist in supply chain management by predicting demand, optimizing inventory levels, and improving logistics [68][69][70][71]. In summary, AI and machine vision have the potential to revolutionize the food processing industry by enhancing productivity, improving product quality and safety, and optimizing supply chain management.

6. References

[1] United Nations Department of Economic and Social Affairs, Population Division. World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/NO. 3. (2022)

[2] Alamu, E.O., Mooya, A., “Food Processing Technologies and Value Addition for Improved Food Safety and Security”. In: Nhamo, N., Chikoye, D., Godwe, T. Smart Technologies for Sustainable Smallholder Agriculture. Academic Press, 2017, pp. 201-210

[3] Hameed, F., Ayoub, A., Gupta, N. Novel food processing technologies: An overview. International Journal of Chemical Studies, 6, (2018) 770-776.

[4] McCarthy, J., Minsky, M., Rochester, N., Shannon, C. E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. Association for the Advancement of Artificial Intelligence, 27, (1955), pp. 1-12

[5] Varsha S., Sandeep, S., Rijwan, K. Modelling Techniques to Improve the Quality of Food Using Artificial Intelligence. Journal of Food Quality, 2021, (2021), pp. 1-10

[6] Mounika, A., Priyanka, P., Prmeela, K. Recent advances and applications of artificial intelligence and related technologies in the food industry. Applied Food Research, 2, (2022), p. 100126

[7] Bishop, C.M. Neural networks and their applications. Review of Scientific Instruments, 65, (1994), pp. 1803-1832

[8] Weijters, A.J.M.M., Hoppenbrouwers, G.A.J. Backpropagation networks for Grapheme-Phoneme Conversion: a nontechnical introduction. In: Braspenning, P.J., Thuijsman, F., Weijters, A.J.M.M. (Eds.) Artificial Neural Networks: An Introduction to Ann Theory and Practice. Springer-Verlag, (1995), pp. 12-36

[9] Huang, Y. Advances in Artificial Neural Networks – Methodological Development and Application. Algorithms, 2, (2009), pp. 973-1007

[10] Zou, J., Han, Y., So, SS. Overview of artificial neural networks. Methods Mol Biol, 458, (2008), pp. 15-23.

[11] Searle, J.R. Minds, brains, and programs. Behavioral and Brain Sciences, 3, (1980), pp. 417-424

[12] Gams, M. Strong vs. Weak AI. Informatica, 19, (1995), pp. 479-493

[13] LeCun, Y., Bengio, Y., Hinton, G.E. Deep learning. Nature, 521, (2015), pp. 436-444. May. 2015

[14] Hussien, R.M., AL-Jubouri, K.Q., Gburi, M.A., Qahtan, A.G.H., Jaafar, A.H.D. Computer Vision and Image Processing the Challenges and Opportunities for new technologies approach: A paper review. Journal of Physics: Conference Series, 1973, (2021), pp. 012002-012002

[15] Kagami, S. High-speed vision systems and projectors for real-time perception of the world. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 2010, pp. 100-107

[16] Ren, F., Bao, Y. A Review on Human-Computer Interaction and Intelligent Robots. World Scientific, 19, (2020), pp. 5-47

[17] Albustanji, R.N., Elmanaseer, S., Alkhatib, A.A. Robotics: Five Senses plus One—An Overview. Robotics, 12, (2023), pp. 68-68

[18] Beysolow, T. What Is Natural Language Processing? In: Applied Natural Language Processing with Python. Apress Berkeley CA, USA, (2018), pp. 1-12

[19] Singh, S.K., Mahmood, A. The NLP Cookbook: Modern Recipes for Transformer Based Deep Learning Architectures. IEEE Access, 9, (2021), pp. 68675-68702

[20] Liao, S. Expert system methodologies and applications—a decade review from 1995 to 2004. Expert Systems with Applications, 28, (2005), pp. 93-103

[21] Zhu, L., Spachos, P., Pensini, E., Plataniotis, K.N. Deep learning and machine vision for food processing: A survey. Current Research in Food Science, 4, (2021), pp. 233-249

[22] Kozacek, B., Grauzel, J., Frivaldský, M.. The main capabilities and solutions for different types of the image sensors. 2018 ELEKTRO, Mikulov, Czech Republic, 2018, pp. 1-5

[23] Abdullah, M.Z. Computer vision and infrared techniques for image acquisition in the food and beverage industries Computer vision technology for food quality evaluation. In: Sun, DW. (Ed.) Computer Vision Technology in the Food and Beverage Industries, Woodhead Publishing, (2012), pp.3-26

[24] Strum, A., Fenigstein, A., Rizzolo, S. Complementary metal-oxide-semiconductor (CMOS) X-ray sensors. In: Durini, D. (Ed.) High Performance Silicom Imaging (2th ed.). Woodhead Publishing, (2020), pp. 413-436

[25] Chou, H.P., Lai, D.T., Hsu, C.I. A CCD Based X-ray Imaging System For Industrial Applications. 1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference, San Francisco, CA, USA, (1993), pp. 222-226,

[26] Gowen, A., Tiwari, B.K., Cullen, P.J., McDonnell, P.J., O’Donnell, P.J. Applications of thermal imaging in food quality and safety assessment. Trends in Food Science and Technology, 21, (2010), pp. 190-200

[27] Gade, R., Moeslund, T.B. Thermal cameras and applications: a survey. Machine Vision and Applications, 25, (2013), pp. 245-262

[28] URL:https://www.flir.eu/discover/instruments/process-quality/thermal-imaging-cameras-in-the-food-industry/ (25.5.2024.)

[29] Liu, Y., Kong, J.Y., Wang, X.D., & Jiang, F. Research on image acquisition of automatic surface vision inspection

systems for steel sheet. 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), 6, pp. 189-192

[30] Golnabi, H. Role of laser sensor systems in automation and flexible manufacturing. Robotics and Computer-Integrated Manufacturing, 19, (2003), pp. 201-210

[31] Alenyà, G. Foix, S.C., Torras, C. Using ToF and RGBD cameras for 3D robot perception and manipulation in human environments. Intelligent Service Robotics, 7, (2014), pp. 211-220

[32] Li, L. Time-of-Flight Camera - An Introduction. SLOA190B Texas Instruments, (2014), pp. 1-10

[33] Vázquez-Arellano, M., Griepentrog, H.W., Reiser, D., Paraforos, D.S. 3-D Imaging Systems for Agricultural Applications—A Review. Sensors, 16, (2016), pp. 618-618

[34] Karabegović, I., Vojić, S., Doleček, V. 3D Vision in Industrial Robot Working Process. 2006 12th International Power Electronics and Motion Control Conference, Portoroz, Slovenia, 2006, pp. 1223-1226

[35] Kuroda, R., Sugawa, UV/VIS/NIR imaging technologies: challenges and opportunities, Proc. SPIE 9481, Image Sensing Technologies: Materials, Devices, Systems, and Applications II, 948108, (2015)

[36] Amigo, J.M., Grassi, S. Configuration of hyperspectral and multispectral imaging systems In: Amigo, J.M. (Ed.) Hyperspectral Imaging, Data Handling in Science and Technology, Elsevier, 32, (2019), pp. 17-34

[37] Wieme, J., Mollazade, K., Malounas, I., Zude-Sasse, M., Zhao, M., Gowen, A.A., Argyropoulos, D., Fountas, S., Van Beek, J. Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review. Biosystems Engineering, 222, (2022), pp. 156-176

[38] Abbott, J.A. Quality measurement of fruits and vegetables. Postharvest Biology and Technology, 15, (1999), pp. 207-225

[39] Golnabi, H., Asadpour, A. Design and application of industrial machine vision systems. Robotics and Computer-Integrated Manufacturing, 23, (2007), pp. 630-637

[40] Martin, D. A Practical Guide to Machine Vision Lighting. Advanced illumination, (2013), pp. 1-21

[41] Liu, Y.X., & Zhou, P. Application of Machine Vision System in the Precision Inspection. Advanced Materials Research, 546-547, (2012), pp. 1382 - 1386

[42] Yan-Li, A. Introduction to Digital Image Pre-processing and Segmentation. 2015 Seventh International Conference on Measuring Technology and Mechatronics Automation, Nanchang, China, (2015), pp. 588-593

[43] Chandra, J., Supraja, B.S., Bhavana, V. A A Survey on Advanced Segmentation Techniques in Image Processing Applications, 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore, India, (2017), pp. 1-5

[44] Mittal, H., Pandey, A.C., Saraswat, M., Kumar, S., Pal, R., Modwel, G. A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets. Multimedia Tools and Applications, 81, (2021), pp. 35001-35026

[45] Sevak, J.S., Kapadia, A.D., Chavda, J. B., Shah, A., Rahevar, M. Survey on semantic image segmentation techniques," 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, India, (2017), pp. 306-313

[46] Shih, F.Y. Image Segmentation. In: F.Y. Shih (Ed.) Image Processing and Pattern Recognition. IEEE Press, (2010), pp. 119-178

[47] Brosnan, T., Sun, D. Improving quality inspection of food products by computer vision—a review. Journal of Food Engineering, 61, (2004), pp. 3-16

[48] Bhamare, D., Suryawanshi, P. Review on Reliable Pattern Recognition with Machine Learning Techniques. Fuzzy Information and Engineering, 10, (2018), pp. 362-377

[49] Rong, A., Akkerman, R. Grunow, M. An optimization approach for managing fresh food quality throughout the supply chain. International Journal of Production Economics, 131, (2011), pp. 421-429

[50] He, Y., Huang, H., Li, D., Shi, C., Wu, S. J. Quality and Operations Management in Food Supply Chains: A Literature Review. Journal of Food Quality, 2018(1), (2018), pp. 1-14

[51] Mavani, N.R., Ali, J.M., Othman, S., Hussain, M.A., Hashim, H., Abd Rahman, N. Application of Artificial Intelligence in Food Industry—a Guideline. Food Engineering Reviews, 14, (2022), pp. 134–175.

[52] Zhu, L., Spachos, P., Pensini, E., Plataniotis, K.N. Deep Learning and Machine Vision for Food Processing: A Survey. Current Research in Food Science, 4, (2021), pp. 223-249

[53] Banga, J.R., Balsa-Canto, E., Moles, C.G., Alonso, A.A. Improving food processing using modern optimization methods. Trends in Food Science & Technology, 14, (2003), pp. 131-144

[54] Davies, A., Thomas, P., Shaw, M. The utilization of artificial intelligence to achieve availability improvement in automated manufacture. International Journal of Production Economics, 37, (1994), pp. 259-274

[55] Livieris, I.E., Kiriakidou, N. Kanavos, A. Vonitsanos, G. Tampakas, V. Employing Constrained Neural Networks for Forecasting New Product's Sales Increase. 5th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Hersonissos, Greece, (2019) pp. 161-172

[56] Niaki, A.A., Rashidaee, S.A. The Role of Predictive Maintenance in the Continuity of the Food Enegy Water Nexus; A Case Study. 2019 Iranian Conference on Renewable Energy & Distributed Generation (ICREDG), Tehran, Iran, (2019), pp. 1-5

[57] Addanki, M., Patra, P., Kandra, P. Recent advances and applications of artificial intelligence and related technologies in the food industry. Applied Food Research, 2, (2022), 100126

[58] McClements, D.J., Barrangou, R., Hill, C., Kokini, J.L., Lila, M.A., Meyer, A.S., Yu, L. Building a Resilient, Sustainable, and Healthier Food Supply through Innovation and Technology. Annual Review of Food Science and Technology, 12, (2021), pp. 1-28

[59] Weidner, N.U. Artificial Intelligence, Machine Learning, and Gender Bias. In: The International Encyclopedia of Gender, Media, and Communication (Eds. K. Ross, I. Bachmann, V. Cardo, S. Moorti and M. Scarcelli). Wiley&Sons, (2020)

[60] Agbai, C.M. Application of artificial intelligence (AI) in food industry. GSC biological and pharmaceutical sciences, 13, (2020), pp. 171-178

[61] Guo, Z.M., Zhang, M., Lee, D., Simons, T. Smart Camera for Quality Inspection and Grading of Food Products. Electronics, 9, (2020), pp. 1-18, 505.

[62] Campos, M., Ferreira, M.J., Martins, T., Santos, C.P. Inspection of bottles crates in the beer industry through computer vision. IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society, (2010), pp. 1138-1143

[63] Sivaranjani, A., Senthilrani, S., Ashokumar, B., Murugan, A. An Improvised Algorithm For Computer Vision Based Cashew Grading System Using Deep CNN. 2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA), Pondicherry, India, (2018), pp. 1-5

[64] Tho, T.P., Thinh, N.T., Bich, N.H. Design and Development of the Vision Sorting System. 2016 3rd International Conference on Green Technology and Sustainable Development (GTSD), Kaohsiung, Taiwan, (2016), pp. 217-223

[65] Ali, M.M., Hashim, N., Aziz, S.A., Lasekan, O. Quality Inspection of Food and Agricultural Products using Artificial Intelligence. Advances in Agricultural and Food Research Journal, 2, (2021), pp. 1-17

[66] Soltani-Fesaghandis, G., Pooya, A. Design of an artificial intelligence system for predicting success of new product development and selecting proper market-product strategy in the food industry. International Food and Agribusiness Management Review, 21(7), (2018), 847-864

[67] Kumar, I., Rawat, J., Mohd, N., Husain, S. Opportunities of Artificial Intelligence and Machine Learning in the Food Industry. Journal of Food Quality, 2021(1), (2021), 4535567

[68] Monteiro, J.M.R., Barata, J. Artificial Intelligence in Extended Agri-Food Supply Chain: A Short Review Based on Bibliometric Analysis. Procedia Computer Science, 192, (2021), pp. 3020-3029

[69] Sharma, S., Gahlawat, V. K., Rahul, K., Mor, R. S., Malik, M. Sustainable Innovations in the Food Industry through Artificial Intelligence and Big Data Analytics. Logistics, 5(4), (2021), 66.

[70] M. Woschank, E. Rauch and H. Zsifkovits. “A Review of Further Directions for Artificial Intelligence, Machine Learning, and Deep Learning in Smart Logistics. Sustainability, 12(9), (2020), pp. 3760-3760

[71] Kollia, I., Stevenson, J., Kollias, S. AI-Enabled Efficient and Safe Food Supply Chain. Electronics, 10(11), (2021), pp. 1223-1223

Engineering Power – *Bulletin of the Croatian Academy of Engineering* Vol. 18(4) 2023 – ISSN 1331-7210 (Print) ISSN 2718-322X (Online)

Publisher: Croatian Academy of Engineering (HATZ), 28 Kačić Street, P.O. Box 14, HR-10001 Zagreb, Republic of Croatia

Editorial Board Address: Croatian Academy of Engineering (HATZ), “Engineering Power” – Bulletin of the Croatian Academy of Engineering, Editorial Board, 28 Kačić Street, P.O. Box 14, HR-10001 Zagreb, Republic of Croatia

E-mail: hatz@hatz.hr

Graphical and Technical Editor: Tiskara Zelina, Ltd., Zelina

Proof-reader: Miroslav Horvatić, MA

Press: Tiskara Zelina, Ltd., Zelina

Circulation: 200