Suitability Analysis of Machine Learning Algorithms: Processing 3D Spatial Data for Automated Robot Control

Benjamin Peric and Michael Engler

October 17, 2021

Suitability analysis of machine learning algorithms: Processing three-dimensional spatial data for automated robot control

Benjamin Peric, Michael Engler

Hochschule Furtwangen University, Faculty of Business Administration and Engineering, Robert-Gerwig-Platz 1, D-78120 Furtwangen, Germany

benjamin.peric@hs-furtwangen.de

Abstract. Global competition, rapidly rearranging market requirements and shorter product life cycles are expressed in constantly changing environmental conditions, which further complicate the demands on the production process. Given smaller batch sizes in small to medium-sized companies, the importance of flexibly varying handling tasks, which must be implemented through a robot gripping system, increases. Standardized workflows are difficult to establish in undefined environments since the products to be handled vary strongly in orientation and position.

The work aims to determine whether artificial intelligence can be developed through the combination of a color camera including an infrared depth measurement, which enables industrial robots to interact with the environment. The following two research questions arise: 1. to what extent can the potentials of artificial intelligence and its success of the recent period be adapted for the application of a robot gripping process and 2. how this symbiosis effects the use of industrial applications. The combination of intelligently controlled robotics using artificial intelligence and the processing of data without server-driven computing power at the end device form the basis of the investigation. The behavior of neural networks in scenarios with a small amount of data is the focus of the question. The realization of artificial intelligence is carried out in an iterative approach and the development process is available in written form.

The overall context of the approach is questioned via a suitability analysis to gain an understanding of possible applications and to name the limits of the system in the given scenario. With this approach, it can be examined which factors support the use of neural networks in the outlined context and whether they can be used successfully, despite of additional aggravating environmental influences.

Keywords: Artificial Intelligence; Neural Networks; Small Data; Robotics; 3D-Data

1 Introduction

Self-learning computer programs are conquering economic structures as a sustainable branch of industry. The efficiency increase in all areas of a company indicates the potential of digital data processing. Machine learning algorithms can be identified as an essential driver for monitoring, regulation and control of industrial processes [1]. The degree of automatization is to be recognized as a fundamental prerequisite for the long-term safeguarding of competitiveness in manufacturing industries along with the key technology of artificial intelligence. The increase in flexibility, humanity, quality and productivity lies at the core of digitalization and Industry 4.0

in this branch of industry [2]. Autonomously controlled robot processes, digital quality assurance requirements, operational resource planning and preventive process analyses are mostly based on intelligent sensor technology [3].

Robotic applications are becoming increasingly challenging in the field of automation due to growing demands for flexibility. These requirements increase exponentially because of smaller batch sizes which creates an even more difficult starting situation in small to medium sizes companies. Common computer vision approaches could enable applications to deal with a dynamic scenery characterized by the opposing trend currents of automation and flexibility. Automated object manipulation, motion planning and even navigation through a dynamic scenery are the focus of these problems. Application areas within logistics, assembly and production include these problem issues [4].

In addition, the initial situation inhibits access to data collection and data processing. Large information structures often enable machine learning and are widely accepted as a prerequisite. Therefore, the impact of small data is the core of the following investigation.

Low computing power at the end device is another supplemental condition that is closely rooted in the initial scenario. Digital infrastructures cannot be established as a requirement in the manufacturing industry of small to medium-sized enterprises. Server-driven calculations are therefore not available as a solution option for answering the research questions posed [5].

2 Trend Research

Trend research is a methodological tool that attempts to identify the development process based on observations of technical and social changes. The newly acquired knowledge serves as an aid for the user to name individual trends, avoid surprises, assess interactions and enable interpretations through the collection of observations [6]. The scope is limited here to the analysis of already recognized trends and their origins. This approach aims to narrow down the relevant content of the suitability analysis by identifying the potentials and accurately define the core fields for the use case of the work. The high complexity inevitably requires the reduction of manageable factors.

2.1 Drivers of Artificial Intelligence

The applications of artificial intelligence have gained momentum in the last decade. The rapid increase in development can be attributed to the interaction of several factors, which can be seen in the following figure.

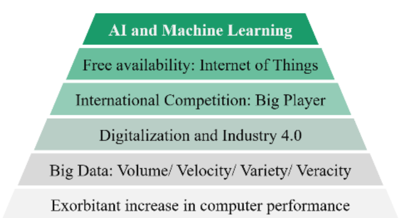

Fig. 1. Factors influencing the development of artificial intelligence

One of the strongest influencing factors is the massive increase in the computing power of intelligent circuits. Already in 1958, a regularity was derived by G. Moore through observations, in which the doubling of the performance can be continuously determined within a period of two years. The so-called Moore's law comprises more than 32 doubling cycles in the present period [7].

The applications of machine learning algorithms benefit especially when processing large amounts of data due to the exorbitant increase in computing power. The handling of large data structures can be described with the term Big Data. At the same time, it implies the possibility of being able to gain target-oriented insights by skillfully processing and analyzing mass data [8]. Symbiosis can be derived in which both the processing task and the medium to be processed will benefit from the constant growth [9].

In a business context, providing useful information, at the right time and in a usable form is critical to success. In addition to real physical data sets, so-called digital twins also serve to calculate and assess real scenarios. These furthermore support the creation process of well-founded data structures [10].

This core value creation process can be derived from the investments within the technology, which have increased fiftyfold to over 15 billion in the last ten years [11]. The leadership position is being fought by big players such as Google, Amazon, and Facebook on an international level. However, access to data sets is not only reserved for the technology giants. The German Federal Statistical Office estimates a tripling of the sensor market by 2025 [12]. The digitalization of industrial production and the high degree of connectivity favor the evaluation of the collected information. In the literature, terms such as "Internet of Things" or "Internet of Everything" are therefore already chosen to describe the development [13].

2.2 Alignment: Research and real-life applications

The following chapter offers a more specific insight into concrete fields of robot applications, which are optimized and implemented with the help of neural networks.

The subject area of robotics mostly covers technical apparatuses that usually interact with the physical world employing mechanical movements. Actuators are operated via kinematic variables by control systems, whereby a specific task can be performed automatically due to its structure [14].

One first milestone was published in October 2020 by Knapp AG in an interview. Through cooperation with Covariant Embodied Intelligence Inc., the transfer of research results to a real-life scenario has taken place. As an operational packaging robot, the system supports an electronics distributor for wholesale within logistics. Smaller goods are automatically transported to the robot arm via a conveyor belt. The gripper system recognizes the type, the lay, and the position of the objects. Transported via a pneumatic suction cup, each of these arrives at the desired packaging location. According to the company's information, this categorizes and recognizes almost 78,000 different small goods, whereby the products are initially unordered in a carton. A short-term maximum of approximately 600 objects per hour is stated. The system detects the products with an accuracy of more than 99% during operation. The gripping process appears very precise considering the error rate of less than one percent. Nevertheless, the manufacturers state a daily working time of just 14 hours [15].

The selected example impressively shows the different demands between research work and real-life solutions. An error rate of less than one percent is to be recognized as a groundbreaking and remarkable performance in the context of new research results. Within one of the most well-known international competitions, such as the ImageNet Challenge, a similar network could replace the previous titleholders [16]. However, these results represent the minimum level of acceptance in real-world use cases. The following short thought experiment illustrates the reason for this fact through an overall system effectiveness calculation.

With an average operating speed of approximately 400 objects per hour, with the assumed error rate of one percent, 96 objects are incorrectly detected during the day and therefore not processed. If the cause of this problem can be solved within five minutes, a machine downtime of approximately eight hours per day can be expected, assuming shift operation as the work design in this example. The downtime under consideration is only caused by a purely technical system error. Other recovery times, such as maintenance and servicing, occupancy, or manning times are disregarded in this consideration. Therefore, the specified working time of the "Pick-it-Easy" robot seems realistically evaluated.

2.3 Challenges of Artificial Intelligence in the manufacturing industry

In the following chapter, some decisive influencing factors are named which limit the use of a machine learning solution in a corporate context. Based on this very compact presentation, the limits of this technology will be critically examined and considered.

Failure of drivers as amplifiers. One of the obvious challenges in implementing meaningful AI-based robotics solutions is the absence of the drivers in business applications.

Hardware requirements. High computing capacities and memory requirements are a condition of practical applications. The necessity has an additional effect on the acquisition and operating costs of the system, which are particularly significant when a battery is used.

Performance standards. Whereas momentary successes of classification in worldwide competitions reach unprecedented accuracies, these results represent only the minimum of acceptance for real requirements.

Specialist personnel and interface disciplines. The pure basic knowledge of machine learning processes can only be implemented successfully in combination with specialist know-how. The generated results can only be validated based on a critical examination. The application of AI-based systems is to be defined as a highly interdisciplinary work process, which requires the combination of trained professional competencies [17].

Flexible in use - rigid in applications. Artificial intelligence can be embedded in almost any core activity of a company across industries. The integrity of the technology is highly dependent on each individual dataset used. The quality and quantity of training data ultimately determine the validity of the entire system. Changing environmental conditions limit the use of neural networks enormously. As soon as these changes are not reflected in the dataset used, this harms the capabilities of the whole system [18].

2.4 Experimental approach

The simulation of a simple gripping and joining process serves as the essence of the investigation. The analysis is based on a wooden game with small geometric wooden figures. The handling of unknown figures in a partly undefined environment is to be tested. The following investigation addresses common computer vision tasks, which allow object recognition. Besides the classification and detection, the localization and positioning of objects is the focus of this work. Six degrees of freedom are available to the target objects, which must be defined before grasping. In addition to three translational movements, the geometric bodies can rotate around all three spatial axes [19].

Applications within the sub-discipline of "Transfer-Learning" represent a variant to deal with very small amounts of data. This application area uses pre-trained networks, which are connected beforehand of the algorithms. The stored knowledge of the exorbitantly large networks transfers to the use case by interconnection and the probability of successful generalization increases in the ideal case [20].

The number of parameters in structures such as AlexNet, VGGNet, or ResNet reaches into the high millions. Due to the predefined application criteria, this discipline is left out in the following analysis. Consequently, approaches are chosen which hardly require any major computing power in operation. In addition to the redefinition of learning processes via a so-called Siamese network, compression via knowledge distillation, synthetic data generation, and data augmentation is the focus of the experiments.

A lightweight and very compact time of flight design with stereovision is offered by the infrared depth camera "Intel Realsense D435". The manufacturer of the product offers a comprehensive development platform compatible with free programming libraries and includes the most widely used programming languages [21]. Therefore, this depth sensor technology is the basis of the following work to create three-dimensional datasets and perception of the environment.

3 One-Shot-Learning: Rotation determination of unknown objects

Grasping objects are often located in a two-dimensional disordered initial situation. The localization of the target objects is only one part of the necessary scene determination. The rotation of an object cannot be determined with the help of the image segmentation illustrated above. If the components lie on a surface, it is necessary to extract the position and rotation of the objects, based on which the gripping and joining process can be derived.

It is assumed that the position of the target objects is already clearly determined. Consequently, the grasping objects are perceived in the zenith from the top view. Now the rotation of the object figures on the image plane must be determined. The rotation symmetry properties of the objects specify the maximum rotation on the image plane.

Synthetic data preparation. A single image is captured of each target object. Using image processing techniques, the subsequently visualized test figures are aligned and rotated with a rotation step size of half a degree. Consequently, each class receives a single original image to train. As a test set, five new images are taken of each object, which is augmented using the same methodology. The trainingset of the objects can be seen in the figure below.

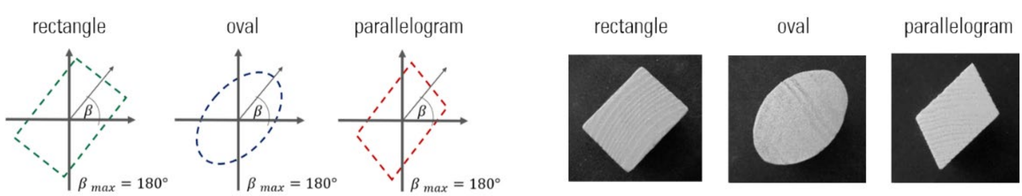

Fig. 2. Rotation determination of three geometric figures

Problem definition. The rotation is to be determined via a regression. Trigonometric relations represent periodic processes as mathematical elementary functions. Therefore, the use of a direction vector allows a regression with the support of oscillation functions. All rotationally symmetrical properties can be mapped via compression and extension of the oscillation functions. Here, the determination of the direction vectors is the focus of the neural regression. The neural network outputs the real vector elements in this approach.

CNN for rotation determination. Two convolutional layers with a kernel of size 3 x 3 scan the characteristics of the figures in 6 and 12 feature maps using relu activation functions. Via a dense layer, the data reach the output, which contains two neurons. The tangent hyperbolic function has the same range of values as the sine and cosine, so the network can determine the real vector elements using the mean squared error.

Table 1. Comparison: Vector and unit vector regression for rotation determination

Approach	**Δ Mean**	**Δ Standard deviation**	**Δ Maximum**	**Intolerance < 2.5°**
Vector output	4.28°	2.49°	21.46°	54.91%
Unit vector output	1.25°	0.95°	5.63°	90.65%

The validation shows that the regressive determination of the output over a simple vector fails. Only 55% of the test data fall within the set tolerance. The regression is therefore extended by the determination of the unit vector. The unit vector amount must always have the length of one. As soon as this condition is included in the output layer, it improves the rotation determination of the three target objects enormously, as shown in the table above.

4 Handling small data sets: Classification of unknown objects

The small geometric figures are to be classified and recognized based on their shapes. This enables the robot system to deal with unknown scenery.

Data acquisition. Four small target figures are measured as a test by the depth camera and stored in a data set. In total, this comprises 10,000 images per class. The captured images have a size of 100 x 100 pixels. The subsequent neural networks each receive a greatly reduced data number of 500 images per class to approximate the problem. The remaining data points are available to the test as validation. The data set represents the point clouds of the target objects from all viewing directions and varies strongly in the distance to the target objects.

The problem of small data. The phenomenon of overfitting occurs especially with small data sets, where the networks over-specify on the existing data points. Overfitting massively counteracts the primary goal of generalization, which is why neural networks can respond poorly to new input in this case [22].

Augmentation. Synthetic augmentation provides an effective method for extracting meaningful information structures despite having few representative data points [23].

The operations to augment the set of data points can be implemented using an integrated development tool called "Keras-Experimental" within the TensorFlow programming library and prepended to the architecture of the network.

The following methods are randomly selected and executed for each input.

- Zooming in and out of the original image up to a maximum of 20% of the existing image dimension
- Mirroring around the horizontal and vertical axis
- Average intensity contrast change by a maximum of 10% of the original image
- Shifting the image by a maximum of 10% in the horizontal or vertical direction

Structure of Siamese Neural Network. The Siamese Neural Network (SNN) is an architecture that enables the handling of very small amounts of data. It constantly receives temporally staggered impulses, which are to be linked together. The two input images are processed in parallel by the same convolutional neural network (CNN) architecture [24]. The SNN receives input pairs of data images with a size of 100 x 100 pixels. Two convolutional layers including max-pooling, alongside 24 and 48 feature maps respectively, process the data each with a kernel of 3 x 3 weights. Using a dropout of 50%, the data is transformed into the dimension of a vector with 50 elements as a result of a dense layer. The parallel processing of the input images enables the SNN to compare the vectors over the Euclidean distance of the intermediate output $\vec{v}_1$ and $\vec{v}_2$, which represent the similarity of the input pair in an abstract form. This scalar value is calculated using a sigmoid activation function.

$$\phi(\vec{v}_1, \vec{v}_2) = \phi(|\vec{v}_1 - \vec{v}_2|) = 1/(1 + e^{-|\vec{v}_1 - \vec{v}_2|}) \quad (1)$$

The more similar the input images are, the smaller is the Euclidean distance of the vectors and the output approaches zero. Hereby the similarity of the three-dimensional shapes of the objects can be measured. A simple comparison image set of the figures to be compared is sufficient for classification since the closest match of the unknown input to the comparison set can be used as a discriminator.

To validate this architecture, almost identical CNNs are used, which map the number of classes to be distinguished in the output layer. With the help of a softmax activation function, the class probability distributions of the target object can be directly specified. The SNN uses the binary cross-entropy due to the binary similarity output structure, whereas the simple CNNs use the pure cross-entropy loss.

Table 2. Comparison: CNN with augmentation vs. Siamese Neural Network

Network	Epochs	Train: 500 images	Test: 9500 images
CNN without augmentation	100	99.17 %	66.24 %
CNN with augmentation	1000	91.17 %	81.87 %
SNN with augmentation	500	99.01 %	90.21 %

The very small datasets cannot be successfully distinguished without augmentation. The classification task of only four targets cannot be performed by simple CNNs for this small dataset. The augmentation methods allow the approximation of the problem, which can counteract the phenomenon of overfitting. The SNN outperforms the results by more than 10%, despite a nearly equal architecture of the model. However, a tenfold calculation time of an epoch must be accepted by doubling the input data. The validation of the SNNs is repeatedly done by pairing the test data. The achieved performance of more than 90% is to be considered as very good in this context since the use of several comparison images per class can additionally raise the result.

5 U-Net Compression with Knowledge Distillation: Image Segmentation

The three-dimensional scenery only acquires a complete meaning through the understanding within the pixel level. In addition to the classification, the object recognition and extrapolation of the object surfaces is therefore the focus of the next task. The recognition of a reference surface enables the robot system to form a normal, which can be used to determine gripping points. Object recognition is one of the central issues of a gripping process, in addition to grip evaluation, behavior coordination, and the determination of minimum holding forces [25].

Data acquisition. The question reflects a binary image segmentation of the point clouds. The generation of the target masks is typically done by hand, which is why the data preparation itself takes a very time-consuming process. However, the Intel-Realsense sensor technology offers some synergy effects. The recorded point clouds can be overlaid with a color image, whereby both data streams reproduce the same scenery. With the help of a simple color filter, the color channels can be used to distinguish the body surfaces. The top of the searched target figure is cut out from the rest of the environment in the color stream, allowing the corresponding mask to be formed to the depth data. The captured scene contains eight different geometric figures. The objects are set in motion on a tabletop, creating a dynamic environment. Different viewing directions and distances are again included in the dataset. The point clouds and the mask images assume a dimension of 200 x 200 pixels and contain 7,500 data points each. The test-training split assumes the ratio of 30 - 70%.

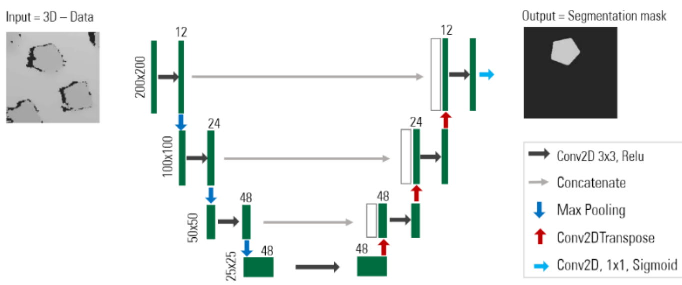

Fig. 3. U-NET Architecture for binary image segmentation

U-Net architecture. The original U-Net structure is used for binary image segmentation and reduced in size for the present application as shown in the figure above [26]. Via convolutional layers, the input data is compressed and brought back to the original size. During deconvolution, an additional dropout of 10% is integrated after each max-pooling layer, which is not visible in the figure above.

Knowledge Distillation. Considering a binary pixel classification, a sigmoid activation function is present in the output layer. The pure U-Net architecture outputs the pixel-probability mask q, which is compared to the target masks y via the binary cross-entropy.

$$D_{BCE}(q, y) = -\sum_{(i,j)} [\, y_{ij} \log(q_{ij}) + (1 - q_{ij}) * \log(1 - q_{ij}) \,] \tag{2}$$

The architecture is trained using the existing data sets and their target masks. This Teacher Network is then used to develop a compressed architecture using Knowledge Distillation. The compressed U-Net structure receives as knowledge distillation loss the pure binary cross-entropy in combination with the Kullback-Leibler divergence [27].

$$D_{KD}(q_s, q_t, y) = \alpha D_{BCE}(y_t, q) + (1 - \alpha) * D_{KL}(q_t^* \| q_s^*) \quad (3)$$

The products of the teacher and student architecture are aligned via the Kullback-Leibler divergence. However, even softer probability distributions q^* of the student and teacher outputs are used. The divergence considers the temperature τ during the final computation with the activation function $q^* = \phi(x/\tau)$. The binary cross-entropy falls with the value $\alpha = 0.1$ to 10% in the combined loss of the student. The temperature τ receives the value 3. Hereby, the student not only receives the hard mask targets but also the corresponding soft probability distributions of the larger teacher network to approximate the problem.

Table 3. Comparison: Compression of U-Net Architectures with Knowledge Distillation

U-Net Network	Parameter	Accuracy	Loss	IoU of the target class
Teacher	80,485	99.02 %	0.0257	85.72 %
Student from scratch	7,753	97.63 %	0.0592	67.54 %
Student with KD	7,753	98.74 %	0.0323	81.27 %

Given the number of parameters that can be trained, the student U-Net architecture represents a tenfold reduction of the teacher, while the basic architecture remains the same. Accuracy does not correctly represent the binary image segmentation problem. Therefore, the Intersection over Union (IoU) is used for validation. The direct comparison of the student architectures with and without knowledge distillation shows optimization of the IoU of 14 percentage points. The tenfold compression therefore takes place with a drop of less than 4 percentage points.

6 Conclusion

The development of neural networks on the end device extremely complicates the development of neural networks. The experiments address extreme cases of an application under the chosen conditions. It was shown that some approaches facilitate the use of small data. Considering real application scenarios, however, the gap between the research results and real requirements becomes strongly apparent.

7 References

1. BMWi: Strategie Künstliche Intelligenz der Bundesregierung. pp. 15-20. BMWi Publikation Schlüsseltechnologien, Berlin (2018).
2. Jeske T., Lennings F.: Produktivitätsmanagement 4.0: Praxiserprobte Vorgehensweisen zur Nutzung der Digitalisierung in der Industrie. pp. 15-18. Springer Vieweg, Berlin (2021).
3. Seifert I. et al.: Potentiale der Künstlichen Intelligenz im produzierenden Gewerbe in Deutschland. pp. 12-15. Institut für Innovation und Technik in der VDI/ VDE Innovation + Technik GmbH, Berlin (2018).
4. Wittpahl V.: iit – Themenband: Künstliche Intelligenz Technologie | Anwendung | Gesellschaft. pp. 110-112. Springer Vieweg, Berlin (2019).
5. Arnold N. und Wangermann T.: Digitalisierung und Künstliche Intelligenz: Orientierungspunkte. pp. 7. Konrad-Adenauer-Stiftung, Berlin (2018).

6. Blechschmidt J.: Quick Guide Trendmanagement: Wie Sie Trendwissen in Ihrem Unternehmen wirksam nutzen. pp. 13. Springer Gabler-Verlag, Berlin (2020).
7. Kreutzer R., Sirrenberg M.: Künstliche Intelligenz verstehen: Grundlagen – Use Cases – unternehmenseigene KI-Journey. pp. 74. Springer Fachmedien, Wiesbaden (2019).
8. Wagener A.: Künstliche Intelligenz im Marketing – ein Crashkurs. pp. 38-43. Haufe-Lexware GmbH & Co. KG, Freiburg (2019)
9. Scheier C., Held D.: Künstliche Intelligenz in der Markenführung: Der effiziente Weg den Erfolg von Marken zu steuern. pp. 33. Haufe-Lexware GmbH & Co. KG, Freiburg (2019).
10. Pistorius J.: Industrie 4.0 – Schlüsseltechnologien für die Produktion: Grundlagen Potentiale Anwendungen. pp. 10 and 41-46. Springer Verlag, Berlin (2020).
11. Lünendonk L. et al.: Künstliche Intelligenz: Eine Studie zum Status quo in deutschen Unternehmen und zu zukünftigen Anwendungsfällen. pp. 7. Lünendonk & Hossenfelder GmbH, Mindelheim (2019).
12. Statistisches Bundesamt: Dossier: Internet of Things. [Data set]. pp. 15 and 29. [Status: 12th August 2021]. Available from: https://www.statista.com/topics/2637/.
13. Pistorius J.: Industrie 4.0 – Schlüsseltechnologien für die Produktion: Grundlagen Potentiale Anwendungen. pp. 9 and 12. Springer Verlag, Berlin (2020).
14. Haun M.: Handbuch Robotik: Programmieren und Einsatz intelligenter Roboter. pp. 18-22. Springer Verlag, Berlin. (2007).
15. Covariant: Case Study: Making Obeta´s warehouse more resilient with KNAPP and COVARIANT´s AI-powered robot. Embodied Intelligence Inc. [Status: 11th August 2021]. Available from: https://covariant.ai/case-studies/obeta.
16. CC-BY-SA 4.0.: Image Classification on ImageNet. [Status: 13th Mai 2021]. Available from: https://paperswithcode.com/sota/image-classification-on-imagenet.
17. Stowasser S., Suchy O. et al.: Einführung von KI-Systemen in Unternehmen: Gestaltungsansätze für das Change-Management. pp. 23-25. Plattform Lernender Systeme, München. (2020).
18. Abdelkafi N. et al.: Künstliche Intelligenz (KI) im Unternehmenskontext: Literaturanalyse und Thesenpapier. pp. 19-24. Apress Inc, New York City (2019).
19. Wolf A. und Schunk H.: Greifer in Bewegung: Faszination der Automatisierung von Handhabungsaufgaben. pp. 109. Carl Hanser Verlag, München (2016).
20. Yang Q. et al.: Transfer Learning. pp. 13. University Press, Cambridge (2020).
21. Intel Corporation: Intel RealSense Depth Camera D435. [Status: 12th August 2021]. Available from: https://www.intelrealsense.com/depth-camera-d435/.
22. Ukil A.: Intelligent Systems and Signal Processing in Power Engineering. pp. 111-115. Springer Verlag, Berlin (2007).
23. Koonce B.: Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization. pp. 36-38. Apress, New York City (2021).
24. Jadon S., Garg A.: Hands-On One-shot Learning with Python: Learn to implement fast and accurate deep learning models with fewer training samples using PyTorch. pp. 27-30. Packt Publishing Ltd., Birmingham (2020).
25. Haase T.: Greifplanung und Greifskills für reaktives Greifen. pp. 10-15. KIT Scientific Publishing, Karlsruhe (2011).
26. Ronneberger O., Fischer P., Brox T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. University of Freiburg, Freiburg (2015). arXiv:1505.04597v1 [cs.CV].
27. Hinton G., Vinyals O., Dean J.: Distilling the Knowledge in a Neural Network. (2015) arXiv:1503.02531v1 [stat.ML].