DOI: 10.58895/ksp/1000150865 -10 erschienen in:
Forum Bildverarbeitung 2022 | Image Pocessing Forum 2022
DOI: 10.5445/KSP/1000150865 | https://www.ksp.kit.edu/site/books/m/10.58895/ksp/1000150865/

Machine learning-based multiobject tracking for sensor-based sorting

Georg Maier[1], Marcel Reith-Braun[2], Albert Bauer[3], Robin Gruna[1], Florian Pfaff[2], Harald Kruggel-Emden[3], Thomas Längle[1], Uwe D. Hanebeck[2], and Jürgen Beyerer[1,4]

[1] Fraunhofer IOSB, Institute of Optronics, System Technologies and Image Exploitation, Karlsruhe, Germany
[2] Intelligent Sensor-Actuator-Systems Laboratory, Karlsruhe Institute of Technology (KIT), Germany
[3] Mechanical Process Engineering and Solids Processing (MVTA), TU Berlin, Germany
[4] Vision and Fusion Laboratory (IES), Karlsruhe Institute of Technology (KIT), Germany

Abstract Sensor-based sorting provides state-of-the-art solutions for sorting of granular materials. Current systems use line-scanning sensors, which yields a single observation of each object only and no information about their movement. Recent works show that using an area-scan camera bears the potential to decrease both the error in characterization and separation. Using a multiobject tracking system, this enables an estimate of the followed paths as well as the parametrization of an individual motion model per object. While previous works focus on physically-motivated motion models, it has been shown that state-of-the-art machine learning methods achieve an increased prediction accuracy. In this paper, we present the development of a neural network-based multiobject tracking system and its integration into a laboratory-scale sorting system. Preliminary results show that the novel system achieves results comparable to a highly optimized Kalman filter-based one. A benefit lies in avoiding tiresome manual tuning of parameters of the motion model, as the novel approach allows learning its parameters by provided examples due to its data-driven nature.

Keywords Sensor-based sorting, machine learning, visual inspection, multiobject tracking

1 Introduction

Sensor-based sorting provides state-of-the-art solutions for sorting of granular materials. This umbrella term describes a family of systems that enable the physical separation of individual objects from a material stream on the basis of information acquired by one or multiple sensors. Among other fields of application, it is considered a key technology for achieving a circular economy. In distinction to mechanical sorting processes such as screening, wind sifting, or float/sink processes, the technology is sometimes also referred to as indirect sorting [1], since particle classification and separation are performed in separate steps. In theory, any number of classes can be recognized for sorting, and separation into multiple fractions is also possible in principle. In industrial applications, however, the task is preferably implemented as a binary sorting task, i. e., sorting into "product" and "residue", since multi-way sorting requires complex mechanical handling.

The functional principle can be summarized as follows. First, the material is fed into the system by means of a conveyor mechanism. Subsequently, the material is transported further via a transport medium. In the course of the transport, sensor-based data acquisition takes place. The data collected is evaluated with the goal to detect and classify individual particles in the material stream. The result of the classification is the basis for the sorting decision, which is executed by means of an actuator. A particular strength of the sorting technology lies in the variety of industrially available sensors that are suitable for use in sensor-based sorting systems. This results in great flexibility with regard to the detectable material properties and thus the sorting criteria to be applied. Due to their suitability for systems with high material throughputs, imaging sensors dominate at this point.

1.1 Motivation

Current systems use line-scanning sensors, which is convenient as the material is perceived during transportation. In case sorting criteria based on color, shape or texture suffice, line-scan cameras in the visible spectrum are used. However, this yields a single observation of each object only and no information about their movement. Due to a delay between localization and separation, assumptions regarding the velocity need to be made in order to calculate the location and point in time for separation [2,3]. Hence, it is necessary to ensure that all objects are transported at uniform velocities. This is often a complex task.

Recent works show that using an area-scan camera instead of a line-scanning one bears the potential to decrease both the error in characterization [4] and separation [5] in sensor-based sorting. Using a sufficiently high frame-rate, individual objects are observed at multiple time points. By employing a multiobject tracking system, this enables an estimate of the followed paths as well as the parametrization of an individual motion model per object. The latter allows for accurate predictions regarding which actuators need to be activated at what point in time such that an object is deflected and hence removed from the material stream. Therefore, the approach is also referred to as predictive tracking. Eventually, this results in an increased sorting quality.

While previous works focus on physically-motivated motion models, it is shown in [6] that state-of-the-art machine learning methods provide a powerful tool for achieving an increased prediction accuracy, particularly in complex sorting scenarios. However, the approach has not been evaluated in real sorting experiments yet, but rather using pre-recorded image data and a simulated separation.

1.2 Contribution

In this paper, we present the development of a neural network-based multiobject tracking system and its integration into a laboratory-scale sorting system with an area-scan camera. This is the first time that the complete development cycle required to make such machine learning-based methods applicable in an industrial sorting setting is considered. With respect to the data processing model itself, we consider the multilayer perceptron from [6]. This model takes observation coordinates of individual objects, which in our case are determined by means of real-time image processing, as an input and generates the predictions for future time points, in our case for the separation stage, as an output. Eventually, actual sorting experiments using the neural network-based multiobject tracking system are conducted.

2 Materials and Methods

In the following, we provide details on the experimental setup, e. g., the exemplary sorting scenario and the considered sorting system, the different prediction models that are compared experimentally as well as the implementation of the real-time inference engine.

2.1 Experimental Setup

We choose an exemplary sorting scenario from the field of construction waste recycling. By generating pure fractions from construction and demolition waste, the material is prepared for the production of recycled construction materials [7]. In our scenario, we consider an input stream consisting of sand-lime brick and brick, see Figure 1. The task is to remove brick from the waste stream. The material is crushed to a grain size of 4 to 6 mm prior to sorting.

(a) Sand-lime brick **(b)** Brick **(c)** Mixed material

Figure 1: Photos of the materials used for the exemplary sorting task.

Both for the acquisition of training data as well as the experimental validation, we use the lab-scale sorting system shown in Figure 2. A detailed description of the system is provided in [5]. A vibrating feeder is used to feed the material in the system. For transportation, a conveyor belt with a width of 140 mm is used. At the end of the belt, right before discharge, the material stream is recorded using an area-scan camera in combination with a ring light. After discharge and during a flight phase, separation is performed using fast switching pneumatic valves.

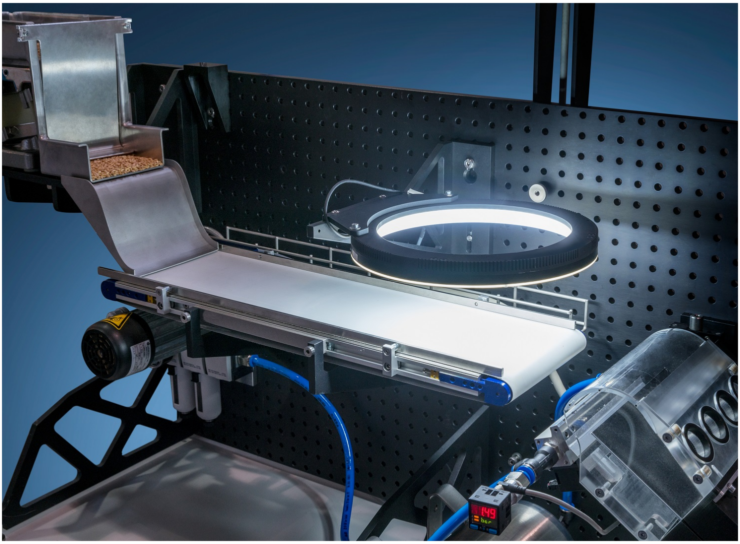


Figure 2: Photo of the lab-scale sorting system used in this study.

The acquired image data is processed with the aim of localizing and classifying individual particles. Based on the classification, a sorting decision is calculated. In case a particle is to be removed from the material stream, a control signal is calculated and transmitted describing the time as well as the valves to be activated in the array. Exactly this calculation, referred to as the prediction model in the following, is the subject of the present study.

2.2 Prediction Models

We validate the proposed approach comparatively. Hence, we also consider two established prediction models for the calculation of the control signals for separation.

First, as a base-line, we consider the system to be equipped with a line-scan camera instead of an area-scanning one. This corresponds to a setup as used in the industry at the time of writing. In this case, no information regarding particles' motion is known. Consequently, a uniform transport velocity has to be assumed. A fixed, typically experimentally determined delay is added to the point in time of observation of a particle in order to calculate the temporal component of the prediction. Furthermore, it is assumed that no velocity perpendicular to transport direction exists. Hence, the valves to be activated correspond to the lateral position of the particle as seen by the camera.

Second, we consider the approach originally proposed in [8] and experimentally validated in [5]. By using a high-speed area-scan camera, particles contained in the material stream are observed at multiple points in time and tracked via a multiobject tracking system. This way, motion parameters, e. g., the velocity in and perpendicular to transport direction, can be determined individually for each particle. In combination with a motion model, these parameters are used to precisely estimate the control signal for separation. The approach focuses on applying Kalman filters on the centroid of the particles for predictive tracking. In this course, linear, physically motivated models, such as constant velocity (CV), are used.

The novel data-driven approach experimentally validated in this paper takes the last five captured position measurements of each particle as input and directly outputs the control signal for separation, i. e., the estimated arrival time and location of the particle at the separation bar. This is opposed to the original predictive tracking algorithm, which for this purpose uses the estimated positions and velocities from the underlying Kalman filter. The input measurements are provided by the exact same multiobject tracking system employed in the original predictive tracking setup. The approach uses a multilayer perceptron with four hidden layers as a predictor, where each hidden layer consists of 16 neurons. Further details on the architecture and training procedure are given in [6].

While numerous tools and software frameworks are now established for model development, the use of neural networks in production systems and, in the present case, under real-time conditions still represents a very current research topic. In the course of this study, various frameworks for integration into the sorting system were investigated in a first step. A technical constraint was the use of the programming language C++. After a first research, the frameworks *TensorRT* from NVIDIA and *OpenVino* from Intel were chosen. These frameworks differ fundamentally in the target hardware on which the inference is executed. *TensorRT* allows the execution of the inference on dedicated NVIDIA graphics cards, *OpenVino* on Intel CPUs as well as integrated Intel GPUs. In both cases, conversion of the model was necessary prior to any potential application. *Onnx* was identified as the current supposedly universal format for this purpose.

In addition to training the developed model on the basis of the generated image sequences, it was also necessary to take knowledge about the system structure into account in the implementation, see Figure 3. Here, parameters relating to the separation, such as the distance between the camera observation area and the separation bar, were primarily decisive. To compensate for errors potentially arising due to measurement inaccuracies, parameters for manual configuration of an offset, e. g., with regard to the distance, were implemented.

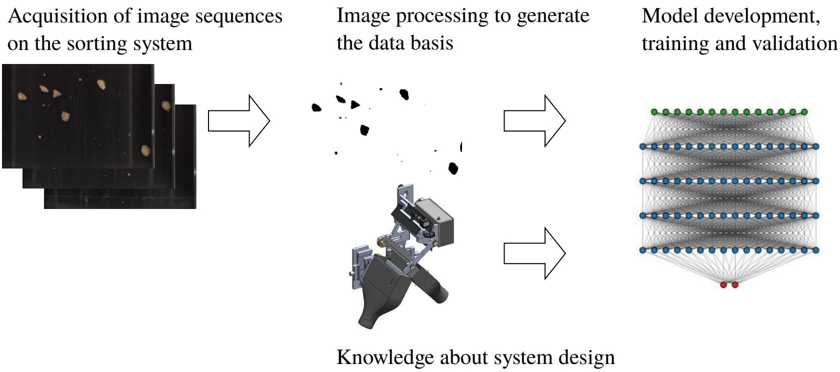


Figure 3: Schematic illustration of the development process of the machine-learning based multiobject tracking.

3 Experimental Validation

We conduct sorting experiments using the methods and materials described in Section 2. One experiment corresponds to sorting 200 g of the material in a batch manner. Additionally to the three prediction models described in Section 2.2, three different mixing ratios are investigated. More precisely, we consider ratios of *residue*, i. e., brick, of

10 %, 25 % and 50 %. Furthermore, we conduct experiments with a mass flow of 10 g/s and 20 g/s.

3.1 Model Training

The multilayer perceptron was trained on a data set of particle tracks recorded on the lab-scale sorting system described in Section 2, with tracks obtained by a preceding offline run of the multiobject tracking algorithm. Although we test the novel approach on several mass flows and mixing ratios in this paper, the multilayer perceptron was trained on only one specification, a mass flow of 20 g/s with a ratio of brick of 25 %, where we used the tracks of both brick and sand-lime brick for training. Images were captured at a frame rate of 100 Hz. The belt velocity was approximately 1 m/s.

The ground truth for the particle's arrival time and location was generated using the concept of a *virtual separation bar* (see [6,8]), since their exact values are not accessible due to the lack of a camera capturing the scene at the separation bar and the limited temporal resolution of most cameras. For this reason, only the images of the area-scan camera are used for training. Therefore, the prediction is performed with respect to a specific pixel row in the camera image corresponding to the virtual separation bar and the tracking phase is shortened accordingly. In addition, the coordinate system for the measurements is shifted so that the virtual separation bar coincides with the real one. The ground truth is then obtained by linear interpolation between the last measurement before and the first measurement after the virtual separation bar. For deployment, the trained network is applied to the original configuration and fed with non-shifted measurements. Although this concept introduces some inaccuracies due to interpolation errors and the assumption of similar particle motion on the belt and in the flight phase, it offers the benefit of not requiring additional sensors and allowing the network to be trained in an unsupervised fashion without additional costs for manually labeling the ground truth.

3.2 Experimental Results

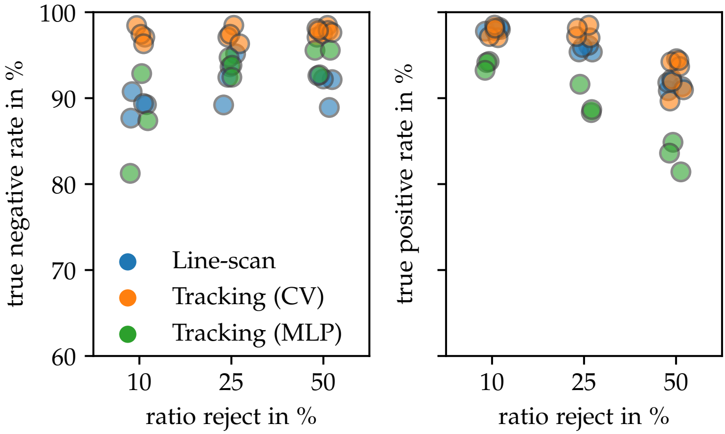
The *true negative rate* (TNR) and *true positive rate* (TPR) were determined as performance indicators for the sorting quality. The TNR refers to the proportion of *residue* material that has been successfully removed, and the TPR to the proportion of *product* material that has successfully been accepted, i. e., not been removed. A selection of the results obtained is shown in Figure 4. The individual markers represent the result of an individual experiment.

As can be seen from Figure 4, the preliminary results show that the novel system achieves results comparable to a highly optimized Kalman filter-based one, although it does not outperform it. However, considering the early stage of development and the opportunities for increasing performance, e. g., by means of training data, we consider it a promising future research direction. An already gained benefit lies in avoiding tiresome manual tuning of parameters of the motion model, as the novel approach allows learning its parameters by provided examples due to its data-driven nature.

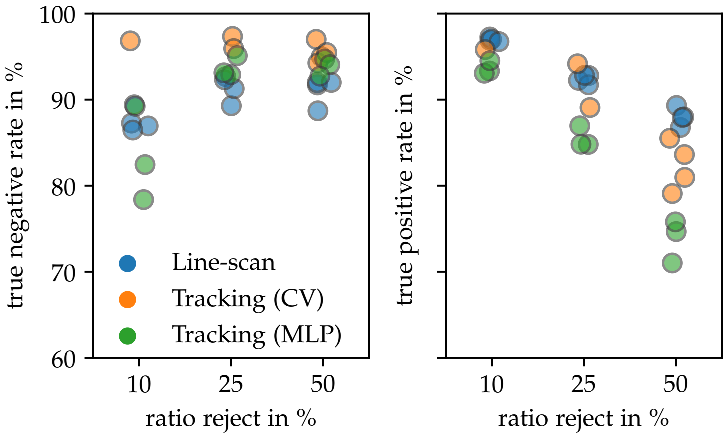
4 Conclusion

In this paper, we presented the experimental validation of a novel neural network-based multiobject tracking system. For this paper, we implemented and integrated the system for use with a laboratory-scale sorting system that was equipped with an area-scan camera. We compared the performance to ones achieved using a line-scan-based system as well as a multiobject tracking system with physically-motivated motion models. Preliminary results show that the novel system achieves results comparable to a highly optimized Kalman filter-based one, although it does not outperform it yet. However, an advantage of the novel system lies in avoiding tiresome manual tuning of parameters of the motion model.

Considering the early stage of development of the system, we believe there exist various interesting research directions to boost its performance. Great potential is believed to lie in the expansion and systematic selection of training data. Furthermore, a system combining physically-motivated as well as machine learning-based models as described in [6] is of great interest.



(a) Mass flow 10 g/s.



(b) Mass flow 20 g/s.

Figure 4: Results of the sorting experiments using the three different prediction models in terms of TNR and TPR. The individual markers represent the result of an individual experiment.

Acknowledgement

IGF project 20354 N of research association Forschungs-Gesellschaft Verfahrens-Technik e.V. (GVT) was supported by the AiF under a program for promoting the Industrial Community Research and Development (IGF) by the Federal Ministry for Economic Affairs and Climate Action on the basis of a resolution of the German Bundestag.

References

1. S. P. Gundupalli, S. Hait, and A. Thakur, "A review on automated sorting of source-separated municipal solid waste for recycling," *Waste Management*, vol. 60, pp. 56–74, Feb. 2017.
2. N. Dias, I. Garrinhas, A. Maximo, N. Belo, P. Roque, and M. T. Carvalho, "Recovery of glass from the inert fraction refused by MBT plants in a pilot plant," *Waste Management*, vol. 46, pp. 201–211, Dec. 2015. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0956053X15300684
3. B. Küppers, S. Schloegl, G. Oreski, R. Pomberger, and D. Vollprecht, "Influence of surface roughness and surface moisture of plastics on sensor-based sorting in the near infrared range," *Waste Management & Research*, vol. 37, no. 8, pp. 843–850, Jun. 2019. [Online]. Available: https://doi.org/10.1177/0734242X19855433
4. G. Maier, A. Shevchyk, M. Flitter, R. Gruna, T. Längle, U. D. Hanebeck, and J. Beyerer, "Motion-based visual inspection of optically indiscernible defects on the example of hazelnuts," *Computers and Electronics in Agriculture*, vol. 185, 2021.
5. G. Maier, F. Pfaff, C. Pieper, R. Gruna, B. Noack, H. Kruggel-Emden, T. Längle, U. D. Hanebeck, S. Wirtz, V. Scherer, and J. Beyerer, "Experimental evaluation of a novel sensor-based sorting approach featuring predictive real-time multiobject tracking," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 2, pp. 1548–1559, 2021.
6. J. Thumm, M. Reith-Braun, F. Pfaff, U. D. Hanebeck, M. Flitter, G. Maier, R. Gruna, T. Längle, A. Bauer, and H. Kruggel-Emden, "Mixture of experts of neural networks and kalman filters for optical belt sorting," *IEEE Transactions on Industrial Informatics*, 2021.
7. S. Dittrich, V. Thome, J. Nühlen, R. Gruna, and J. Dörmann, "Baucycle–verwertungsstrategie für feinkörnigen bauschutt," *Bauphysik*, vol. 40, no. 5, pp. 379–388, 2018.

8. F. Pfaff, M. Baum, B. Noack, U. D. Hanebeck, R. Gruna, T. Längle, and J. Beyerer, "Tracksort: Predictive tracking for sorting uncooperative bulk materials," in *2015 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2015, pp. 7–12.