

# Deep learning-based instance segmentation on 3D laser triangulation data for inline monitoring of particle size distributions in construction and demolition waste recycling

Xiaoye Wu, Nils Kroell<sup>\*</sup>, Kathrin Greiff

Department of Anthropogenic Material Cycles, RWTH Aachen University, Wuellnerstr. 2, D-52062 Aachen, Germany

## ARTICLE INFO

### Keywords:

Sensor-based material flow characterization  
Digital image processing  
Deep learning  
Artificial intelligence  
Instance segmentation  
Circular economy

## ABSTRACT

Overlapping material flow presentations in construction and demolition waste (CDW) recycling make an inline particle size distribution (PSD) monitoring challenging. Here, we aim to build a deep-learning-based segmentation model for overlapping particles in 3D-laser-triangulation images of CDW. Our model was trained on three specially designed datasets with two transfer learning processes. U-net was employed as the backbone and Multi-Star algorithm was used to describe particle shapes. The final model demonstrated an impressive performance on test set, with a mean average precision (mAP) of 92.8% at IoU= 0.5. Comparing with the traditional segmentation algorithm based on image processing methods, the mAP can only reach to 27.4% on the same images. The shown model performance paves the way toward novel sensor technology applications for real-time PSD monitoring in CDW recycling.

## 1. Introduction

Construction and demolition waste (CDW) is the largest source of waste in Germany (Statistisches Bundesamt, 2021). The amount of the CDW generated in Germany has reached 230.9 million tons in 2019, which was 55.4 wt% of the total waste production (Statistisches Bundesamt, 2021). Despite the high amount and ecological relevance, the current CDW recycling shows considerable potential for optimization: Recovered aggregates are so far mainly used in applications with lower quality requirements such as road, earthworks, and landfill construction and only 12.5 wt% of the aggregate demand could be covered into recycled aggregates in 2018 (German Building Materials Association, 2020). In order to substitute more primary raw materials in high-quality applications and realize the associated environmental benefits, it is necessary for recycled materials to consistently meet given quality requirements (Da Leite et al., 2011).

### 1.1. Sensor-based particle size prediction

A decisive quality criterion for recycled CDW is the *particle size distribution* (PSD) (DIN 66165-1, 2022), which is currently determined by manual sieving analysis from taken samples. Sampling and manual

sieving analysis are time-, labor- and cost-intensive and the accuracy of the manually determined PSD is strongly limited by the representativeness of the sample (Khodier et al., 2020). Sensor-based particle size prediction could automate this process and enable an inline monitoring of the full material flow in real-time (Kroell et al., 2022). Based on the inline determined PSD novel sensor technology applications in CDW recycling such as an automated monitoring of product qualities and adaptive process control of, e.g., comminution units could be enabled. Based on these novel applications, the quality and acceptance of recycled aggregates from CDW recycling are believed to be significantly increased (Kroell et al., 2022).

While previous research (Section 1.3) has mostly focussed on determining PSDs at singled material flow presentations (i.e., particles do not touch or overlap each other), in most CDW processing plants, CDW material flows are transported as multilayered bulks due to the high throughput, i.e., particles overlap and touch each other (Kroell et al., 2022). To achieve a precise PSD prediction under overlapping material flow presentations, touching particles must therefore first be *segmented*, before a PSD prediction can be applied (Kroell et al., 2022). This problem can be described as an *instance segmentation* problem, where all instances (here: CDW particles) in an image shall be localized and segmented from each other at the pixel level (Minaee et al., 2022).

<sup>\*</sup> Corresponding author.

E-mail address: [nils.kroell@ants.rwth-aachen.de](mailto:nils.kroell@ants.rwth-aachen.de) (N. Kroell).

### Abbreviations

AP	Average precision
CDW	Construction and demolition waste
DL	Deep learning
IoU	Intersection over Union
mAP	Mean average precision
ML	Machine learning
NMS	Non-maximum suppression
PLSR	Partial Least Squares Regression
PSD	Particle size distribution
SAM	Segment Anything Model
3DLT	3D laser triangulation

## 1.2. Image segmentation without particle singulation

Traditional segmentation methods include thresholding, region growing, contour-based segmentation, and model-based segmentation (Jähne, 2012). These traditional methods can be applied to image segmentation tasks, where individual instances are distinguished from each other by, e.g., clear edges. In CDW recycling, however, such clear edges often do not exist since CDW particles often have complex particle shapes or overlap each other. As we will show in a later comparison (Section 3.4), traditional segmentation algorithms can reach their limits in such complex segmentation cases.

## 1.3. Related work

### 1.3.1. Sensor-based determination of PSDs

To obtain PSD of input materials in a fast and accurate way, many researchers studied on sensor-based material monitoring. For example, the method of Di Maria et al. (2016) used an image descriptor that translates an image of CDW into a feature vector, which can be interpreted into PSD prediction. Kandlbauer et al. (2021) used RGB images of singled particles of commercial waste by means of basic computer vision operations to obtain the characteristic geometric features. These features are then used as training data for a regression model to assign particle sizes to five classes. Zhang and Liang (2016) modeled an inline analysis for PSD prediction based on a soft sensor that trained a Support Vector Machine using parameters from the loop milling process for ores. However, the pretreatment of the materials and complexity in the recycling plants present difficulties for an inline PSD prediction.

### 1.3.2. Deep-learning-based instance segmentation

Deep learning (DL) segmentation models are becoming increasingly popular for instance segmentation, and have demonstrated high accuracies with good robustness across different use cases (Arnab and Torr, 2017). In medical science area, Allioui et al. (2023) employed a modified version of the U-Net for segmentation of chest computer tomography images, to support COVID-19 detection. To detect the wildfire, Qurratulain et al. (2023) applied a ResNet-50 model in instance segmentation for burnt area and achieved a greater accuracy compared with primitive techniques. Ni et al. (2021) developed a framework of 3D segmentation using mask R-CNN for individual blueberries as they develop in clusters and to extract blueberry cluster traits. In contrast to traditional segmentation approaches, however, DL segmentation models demand large amounts of labeled datasets for training. In the case of the instance segmentation for CDW, data labeling is associated with high effort since the contour of each particle in the dataset must be manually annotated. To demonstrate that DL based instance segmentation for CDW can be achieved with reasonable labeling effort, we apply transfer learning processes, as described in Section 2.3.

## 1.4. Research aim

The aim of this study is to demonstrate the technical feasibility of instance segmentation for sensor-based PSD predictions in CDW recycling using DL. To achieve this research aim, we aim at answering the following four research questions:

- RQ1: What are suitable hyperparameters for the given segmentation task?
- RQ2: How did the applied transfer learning process influence the segmentation performance of the model?
- RQ3: What are limitations of developed segmentation model?
- RQ4: How the achieved segmentation results using DL perform compared to traditional segmentation approaches?

## 2. Materials and method

To generate the DL model, CDW samples and specially designed training methods were applied. This chapter introduces the materials prepared for the experiment and the method used in this study.

### 2.1. Materials

The test material used in this study was recycled aggregate (particle size range 0 mm – 45 mm) from a CDW processing plant in Germany (MAV Krefeld GmbH [47809 Krefeld, Germany]). In the CDW processing plant, the CDW input material is first sieved and manually pre-sorted. Then, the coarser particles are comminuted using an impact crusher. From the comminuted material flow, ferrous metals, and films are removed using an overbelt magnet and windsifter, respectively, before a final screening is used to produce the recycled aggregate. The final recycled aggregate mainly consists of minerals, building stone, pottery (predominantly pieces of tile), concrete and brick debris (cf. Fig. 1a) and is currently used, e.g., for frost protection layers or gravel bearing layers according to TL SoB-StB 20 (Forschungsgesellschaft für Straßen- und Verkehrswesen, 2020) or TL Gestein-StB 04/23 (Forschungsgesellschaft für Straßen- und Verkehrswesen, 2023).

The samples were dried at 85°C until weight consistency and then sieved according to (DIN 66165-1, 2022) on the analytical sieve machine from Siebtechnik GmbH (Mühlheim [Ruhr], Germany) with sieve cuts 3.15 mm; 4.0 mm; 5.0 mm; 6.3 mm; 8.0 mm; 10.0 mm; 12.5 mm; 16.0 mm; 22.4 mm and 31.5 mm. The samples were sieved in every sieve cut for 90 s at a frequency of 1400 rpm. Under the constraints of working time and labor costs, the number of the experimental particles is decided to be  $n = 4352$ .  $n = 3432$  of them are used to simulate the real working conditions,  $n = 920$  of them are used to produce synthetic images. Considering the difficulty of the particle annotation, the particle size class of the CDW material used in this study were limited to 12.5 mm - 16.0 mm, 16.0 mm - 22.4 mm and 22.4 mm - 31.5 mm (Fig. 1a).

### 2.2. 3DLT measurement

3DLT is a sensor technology, which can be used to make measurements of 3D structures. With the laser system and triangulation method, the information about height and width of the object can be obtained. With the movement of the conveyor belt, the length of the objects can also be stitched together. In the end, the information about all three dimensions is put together by software.

The 3DLT measuring rig setup used in this work is located in the sensor laboratory of the Department of Anthropogenic Material Cycles at RWTH Aachen University (Fig. 1b). The setup consists of conveyor belt, laser emitter, reflect mirrors, and a 3DLT-sensor. The conveyor belt is 385 mm wide and operated a belt speed of 0.15 m/s. The 3DLT-sensor measures particles using the light section method. The light section is generated by two lasers, which project a light section centrally over the entire width of the conveyor belt. The camera can take pictures from

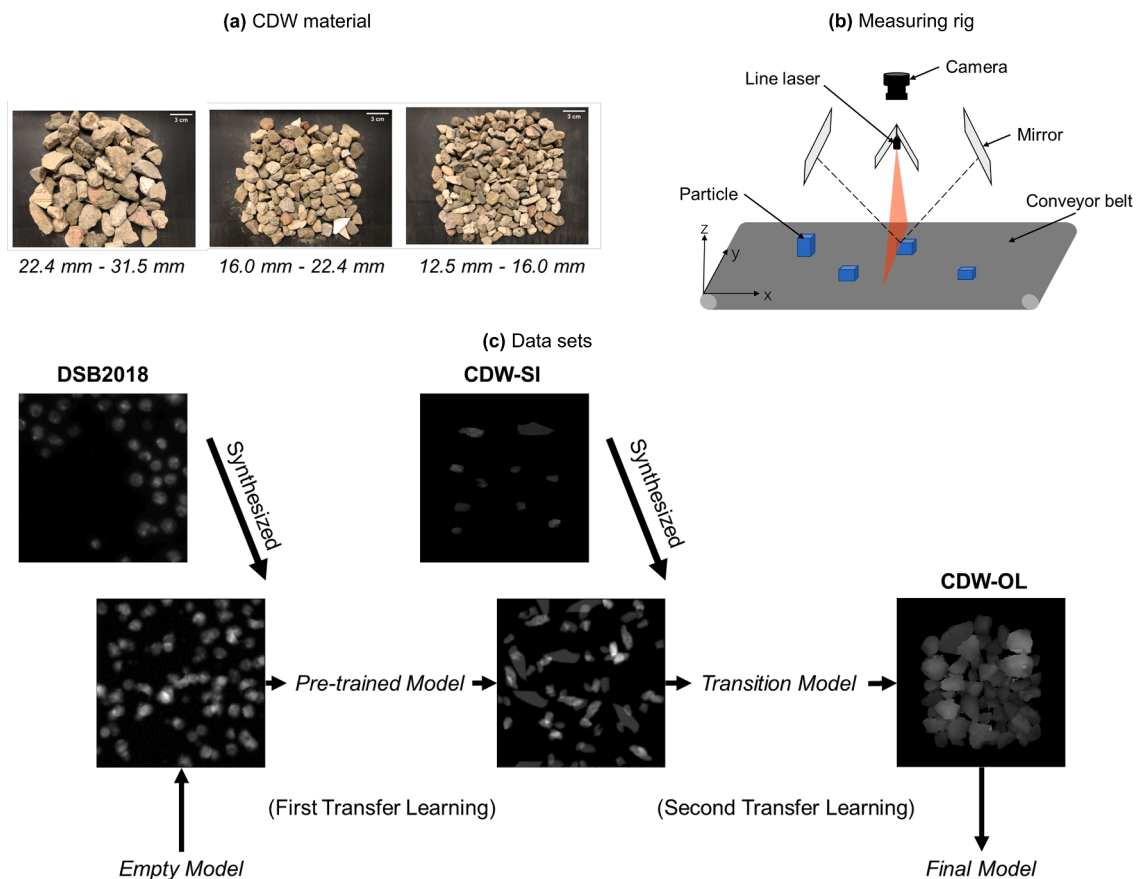


Fig. 1. Material and method. (a) CDW samples with different particle size; (b) 3DLT measuring rig used in this study; (c) Composition of the data sets and their relationships in model training.

both front and back of the particles with the help of the two reflection mirrors. The front and back recordings are saved independently. After calibration through a custom Python software for calibration developed by ANTS (see (Kroell et al., 2021) for details) the resulting 3DLT images have a spatial resolution of 0.331 mm/pixel in x- and y-directions (length and width) and a spatial resolution per gray value of 0.758 mm/pixel in z-direction (height).

### 2.3. Datasets and transfer learning approach

The necessary datasets to train the DL segmentation model consist of particle images and corresponding labels. Labeling the overlapped CDW particles is an expensive and laborious task. To train the model with relatively small data, we apply two transfer learning processes in this study (Fig. 1c). Transfer learning is a method that uses storing knowledge in an existing model while solving one problem and applying it to a different but related problem. The plan is to make two larger datasets (DSB2018 and CDW-SI), which are easier to be produced. After training the model on the two large datasets, the model is transferred to the last learning process based on the small dataset (CDW-OL), which consists of the overlapping CDW particles. Hence, the three investigated dataset are:

1. **DSB2018**: The first dataset was provided by (Data Science Bowl, 2018), which is a public dataset and consists of microscopy images of cells from different organisms. Since the shape and the size of the cells in the images are similar to the CDW particles in the 3DLT photos, this dataset was chosen to do the pre-training process of the DL model. There is no object overlap in these images, so synthetic images were created by randomly replicating, rotating, flipping, and

shifting the objects in the images to make sure that at least 15% of the object pixels were in the overlap of multiple cells (Lu et al., 2017). The DSB2018 dataset consists of  $n=447$  images as training set and  $n=50$  images as the validation set. There are between 60 and 70 cells in every image after synthesis.

2. **CDW-SI**: The second dataset was specifically created for this study and consists of singled CDW. There is still a gap between the DSB2018 dataset and the 3DLT image of the overlapping CDW particles in similarity. In order to improve the similarity of the datasets and at the same time enrich the existing data volume again, the second dataset was created.  $m=100$  g CDW materials were used for every image. Since these images of isolated particles can be automatically labeled by simple segmentation algorithm, they can be produced in large quantities. For example, Kronenwett et al. (2022) successfully trained a DL model for construction wastes classification with synthetic images. Inspired by this, these images were also made into synthetic images like DSB2018 to simulate the real working condition in the plants. Dataset CDW-SI includes 91 images (containing in total 910 particles) as training set and 12 images (containing in total 120 particles) as validation set.
3. **CDW-OL**: The third dataset specifically created for this study was created with overlapping CDW, which is the target condition for inline PSD monitoring in CDW recycling. Considering the image size limitation and the difficulty by labeling,  $m=500$  g CDW samples were used for one image whose composition simulated the real working condition in recycling plants. The samples were given a certain degree of overlap and touching. However, the way of the material accumulation adopted single layer accumulation to avoid the situation of complete coverage. There are a total of 52 particles in an image. The dataset CDW-OL based on real working condition in

plants includes  $n=52$  images (containing in total 2704 particles) as training set and  $n=14$  images (containing in total 728 particles) as validation set.

## 2.4. Segmentation model

### 2.4.1. Multi-Star

An algorithm called MultiStar (Walter et al., 2020) is used to detect the objects. MultiStar is a method for overlapping objects segmentation. It was originally designed to identify the cells in medical images. Walter et al. (2020) proposed the method based on the previous work of Schmidt et al. (2018). MultiStar uses three parameters, object probability  $p_{obj}$ , star distance  $r_k$  and overlap probability  $p_{over}$ , to predict the shape of the object.

Object probability  $p_{obj}$  of a pixel is defined as the normalized Euclidean distance to the nearest background pixel. The value of the  $p_{obj}$  can be regarded as the confidence of object detection, because typically the pixels near the object center are more likely to be a part of an object. This parameter can roughly present the shape of the objects in the image. The star distance refers to the Euclidean distance  $r_k$  to the object boundary from each pixel in the object. The pixels with high  $p_{obj}$  are chosen to be the starting points of star distance calculation. The star distance can be calculated by following each radial direction  $k$  until a pixel with different object identity is encountered (Walter et al., 2020). In this study 32 directions were investigated. These 32 directions of a pixel led to 32 vertexes and built a star-convex polygon, which was the prediction of the instance. The overlap probability  $p_{over}$  is designed to exclude the overlapping region from two objects (Schmidt et al., 2018). The value of  $p_{over}$  is 1 at pixels where at least two objects overlap and 0 elsewhere. The value can also be between 0 and 1, to show the different degrees of certainty. Non-maximum suppression (NMS) algorithm was used to avoid multiple detection of the same object but sometimes also made wrong suppression because of the overlap (Schmidt et al., 2018). With the help of  $p_{over}$ , the algorithm can avoid conflicts between two object predictions by overlap extraction. After giving  $p_{obj}, r_k$  and  $p_{over}$  to NMS, the prediction of the segmentation result can be generated.

### 2.4.2. DL model implementation

The DL model and the production of synthetic images in this study was based on the work of Walter et al. (2020). The DL model was implemented using Python v3.9.12 and executed on the high-performance computer of Department of Anthropogenic Material Cycles at RWTH Aachen University with an AMD Ryzen Threadripper PRO 3995WX 64-Cores 2.70 GHz CPU, two NVIDIA RTX A6000 GPUs with 48 GB memory and 256 GB RAM. Anaconda was used for packages and environments management. The model was implemented on the open-source deep learning framework PyTorch package 1.12.0 and accelerated by CUDA v11.3 and cuDNN v8.3.2. The software used for labeling was Labkit, which is integrated in imageJ from Fiji.

### 2.4.3. Model training

The DL network used in this study was U-Net (Ronneberger et al., 2015). It is a reliable and stable DL architecture for image segmentation. Besides, it is also suitable for the task with a limited number of images and annotations (Minaee et al., 2022). Therefore, U-Net was chosen to be the DL model backbone because it seems most likely to succeed with the task in this study.

As input to the model,  $256 \times 256$  images were fed in, whose resolutions were adjusted to reduce the computational complexity and save the memory. For DSB2018 and CDW-SI, random flips, rotations, and elastic deformation were applied to make synthetic images. The output of the model included three branches for the three prediction features. Each branch consisted of a single convolutional layer. For the object probability and overlap probability branches, a single output channel

with sigmoid activations was used. For the star distances branch, 32 output channels for the 32 radial directions and ReLU activations were used. The output parameter branches were used in subsequent NMS algorithm to compute the final segmentation.

Each input image was given 400 random pixel proposals to make predictions of object probability  $p_{obj}$ , star distance  $r_k$  and overlap probability  $p_{over}$ , as mentioned in Section 2.4.1. The ground truth values of the three parameters were computed from the label of the corresponding image. The prediction value was then compared with the ground truth values and evaluated by loss function. The loss function of the model follows:

$$L(\theta, \sigma_i) = \frac{1}{\sigma_{over}^2} L_{over}(\theta) + \frac{1}{\sigma_{obj}^2} L_{obj}(\theta) + \frac{1}{\sigma_{dist}^2} L_{dist}(\theta) + \log(\sigma_{over} \sigma_{obj} \sigma_{dist}) \quad (1)$$

where the  $\theta$  is the network parameter and the  $\sigma_i$  is the task uncertainty.  $\sigma_i$  is designed for multi-task learning, which weighs the different losses in loss computation (Kendall et al., 2017). The  $L_{over}$  and  $L_{obj}$  are binary cross-entropy losses.  $L_{dist}$  is the mean absolute difference between the predicted and true star distances with every contribution of the pixel weighted by its true object probability. Pixels in overlapping regions are excluded from  $L_{obj}$  and  $L_{dist}$ .

### 2.5. Sobel-Watershed segmentation method

To evaluate the performance of the DL model, a segmentation method based on traditional image processing was built as reference. The basic principle of Sobel-Watershed method is to find the object edges with the help of Sobel filter (Shapiro and Stockman, 2001) and then separate all objects by finite erosion operation. After that, Watershed algorithm (Roerdink and Meijster, 2000) can locate and label each instance in the image. In the end, finite dilation operation is applied, to restore objects to their original size. This algorithm is designed to deal with some simple overlapping problems.

### 2.6. Evaluation metric

The results of the segmentation need to be evaluated to show the performance of the model. They were evaluated on the test set according to average precision (AP) (Mu Zhu, 2004).

The test set includes 5 images and corresponding labels, which were not used in former training or validation process. AP is a metric of object detection accuracy. It indicates the proportion of the correct prediction to all instances in an image. The Intersection over Union (IoU) threshold that evaluates whether the prediction is correct is  $\tau$ . Mean average precision (mAP) is the mean value of the AP values from all test images. It was used to describe the performance of the model over the whole test dataset.

## 3. Results and discussion

### 3.1. Hyperparameter optimization (RQ1)

As mentioned in Section 2.4.3, the model is trained on the training sets, which include  $n=587$  images in total, and verified on the validation sets, which include  $n=76$  labeled images. After the two transfer learning processes, the model for CDW segmentation under missing particle singulation condition can be generated. In this process, many hyperparameters can be changed, which have influence on the final results of the model. After determining the structure of the model, the most important hyperparameters to be set were batch size, the number of epochs  $Epoch$ , learning rate  $lr$ , threshold on  $p_{obj}$   $\rho$  and IoU threshold in NMS  $\nu$ . For the pre-trained model based on DSB2018, the best value of some hyperparameters have already been settled in the previous work of Walter et al. (2020), in which the batch size was 4,  $Epoch$  was 40,  $lr$  was  $10^{-4}$ ,  $\rho$  was 0.3 and  $\nu$  was 0.1. This study adopted these values in the

pre-trained model. Considering the similarity of *DSB2018* and *CDW-SI*, these values were also retained in the first transfer learning process. However, for the second transfer learning, the images in *CDW-OL* show a significant difference from the previous two datasets. Therefore, the hyperparameters of the final model need to be redetermined.

The value of batch size and  $lr$  were the same as the former two training processes, because the dataset was small enough and the computing power was sufficient. The value of *Epoch* can be set at where the loss starts to converge. There were three different prediction values in this study, which were object probability  $p_{obj}$ , overlap probability  $p_{over}$  and star distance  $r_k$ . As a result, there were three loss values during the training process. Fig. S1 in Supplementary Material shows the loss change of the three prediction values with the growth of *Epoch*. The loss of all three parameters have already converged, when the value of epoch reaches at 300.

To find the best performance of the final model, the segmentation result of the model was evaluated and compared under different hyperparameters. The evaluation was completed on the test set, which consists of  $n=5$  images and  $n=5$  corresponding labeled images. Each image includes  $n=52$  particles under missing singulation conditions. The performance comparison of the candidate models is shown in Fig. S2 (Supplementary Materials).

After comparing the mAP values of the model output under different IoU threshold  $\tau$ , the model with  $\rho = 0.5$ ,  $\nu = 0.2$  had the best performance, whose mAP values at different evaluation levels were the highest among the four candidate models (Table S4). The average mAP of this model under four evaluation standards is higher than the other three models by 8.62%, 8.26% and 4.71% respectively. Therefore, the model with  $\rho = 0.5$ ,  $\nu = 0.2$  is chosen to be the final model for the segmentation task.

### 3.2. Influence of transfer learning (RQ2)

#### 3.2.1. Models produced with transfer learning processes

In this study, the first training process together with the two transfer learning processes produced three models. The pre-trained model is the output of the zero-initialized model trained by *DSB2018*. Then the first transfer learning brings the knowledge from the first model to the second training process and the output is the transition model, which is trained by *CDW-SI*. Similarly, the final model based on *CDW-OL* is produced by the second transfer learning process.

Table 1 presents the mAP values on test images of the three models under different IoU thresholds. After the first transfer learning process, that is, from pre-trained model to the transition model, the mAP with the IoU threshold  $\tau = 0.5$ ,  $\tau = 0.6$  and  $\tau = 0.7$  increase by 24.2%, 26.8%, and 22.4% respectively. After the second transfer learning process, that is, from transition model to the final model, the mAP with the IoU threshold  $\tau = 0.5$ ,  $\tau = 0.6$  and  $\tau = 0.7$  increase by 51.7%, 74.8% and 216.0% respectively. It is interesting that the mAP values with  $\tau = 0.8$  of the first two models are almost the same but the value of the final model increases to 0.4349. This indicates that for transfer learning process, the target dataset is the most important dataset, which can provide the details of the images and lead to a higher precision of the prediction. The complete AP values of the three models are included in Table S1, S2 and S3 in Supplementary Materials.

The findings described above show that transfer learning can progressively improve the segmentation performance of the model. As the

**Table 1**

mAP values of the models produced with transfer learning (best mAP values highlighted in bold font).

IoU Threshold $\tau$	0.5	0.6	0.7	0.8
Pe-trained model	0.4924	0.3691	0.1971	0.0599
Transition model	0.6115	0.4681	0.2413	0.0565
Final model	<b>0.9276</b>	<b>0.8783</b>	<b>0.7625</b>	<b>0.4349</b>

dataset gets closer to the target image, the model performs better and better. Fig. 2 presents the segmentation performance of the three models on the test images.

#### 3.2.2. Loss reduction

Fig. 3a shows the total loss of the training process with and without transfer learning. The blue line represents the training process of the transition model with *CDW-OL*, where two transfer learning processes were already applied. It is also the training method used in this study. The red line represents the training process of a zero-initialized DL model with *CDW-OL*, where there is no transfer learning process. All other parameters of the two models are the same. The model with transfer learning starts with a loss of 1.006 while the zero-initialized model without transfer learning starts at 3.752. After the same training steps, the loss of the model with transfer learning decreases to  $-1.365$ , while the loss of the model without transfer learning stops at  $-1.057$ . During the whole training steps of the two models, the loss of the model with transfer learning is always lower than the loss of the zero-initialized model.

Fig. 3b shows the total loss of the whole training process. The training steps axis is compressed to conveniently present the loss change. The loss of the first training phase with *DSB2018* starts at 2.966 and then rapidly decreases. It ends at  $-0.651$  and the pre-trained model is handed over to the second training phase. The second phase of training based on *CDW-SI* starts with a loss of 0.536, which is much lower than the starting loss of the first training phase. The loss stops at  $-0.0175$  and then the last training phase based on *CDW-OL* begins. Starting at 1.006, the total loss declines rapidly. After 300 epochs, the final model is generated with the loss of  $-1.365$ . Fig. 3b shows that both transfer learning processes make the next training phase starts at a low loss value, in other words, they make the training process faster.

The findings above show that transfer learning can accelerate the training process by using the prior knowledge. The training based on *DSB2018* and *CDW-SI* provide the knowledge for the last training process, to help decreasing the starting loss of the model.

### 3.3. Error analysis (RQ3)

The DL model seems successful on the segmentation task, but there are still some drawbacks in this method that have a negative influence on the accuracy of the results.

#### 3.3.1. Annotation of overlapping particles

The first problem is the creation of the *CDW-OL*, in which the particles are not singled. The labels (ground truth) of these images do not include the overlapping parts of the particles, because the overlapping regions cannot simply be detected from the images by manual labeling. As a result, the capability of overlap prediction obtained from the first two datasets may be lost during the final training process based on dataset *CDW-OL*. At the same time, the shapes of some overlapping objects in the ground truth are not the same as in reality. In this situation, the AP value would not be influenced but errors may occur in further applications.

#### 3.3.2. Error of random sampling

Another problem is the random sampling of the starting points. The starting points were selected from the pixels with high  $p_{obj}$  and used to calculate the other parameters for segmentation prediction. The problem is that there are usually multiple qualified pixels. The selection of these starting points is completed by random sampling. Because of this randomness, the outputs of the models after every training process are different. The AP difference caused by this problem is  $\pm 3\%$  to  $\pm 8\%$ .

#### 3.3.3. Error of concave polygon prediction

Fig. 4 shows a common segmentation mistake of the model. The particle with the shape of concave polygon cannot be segmented

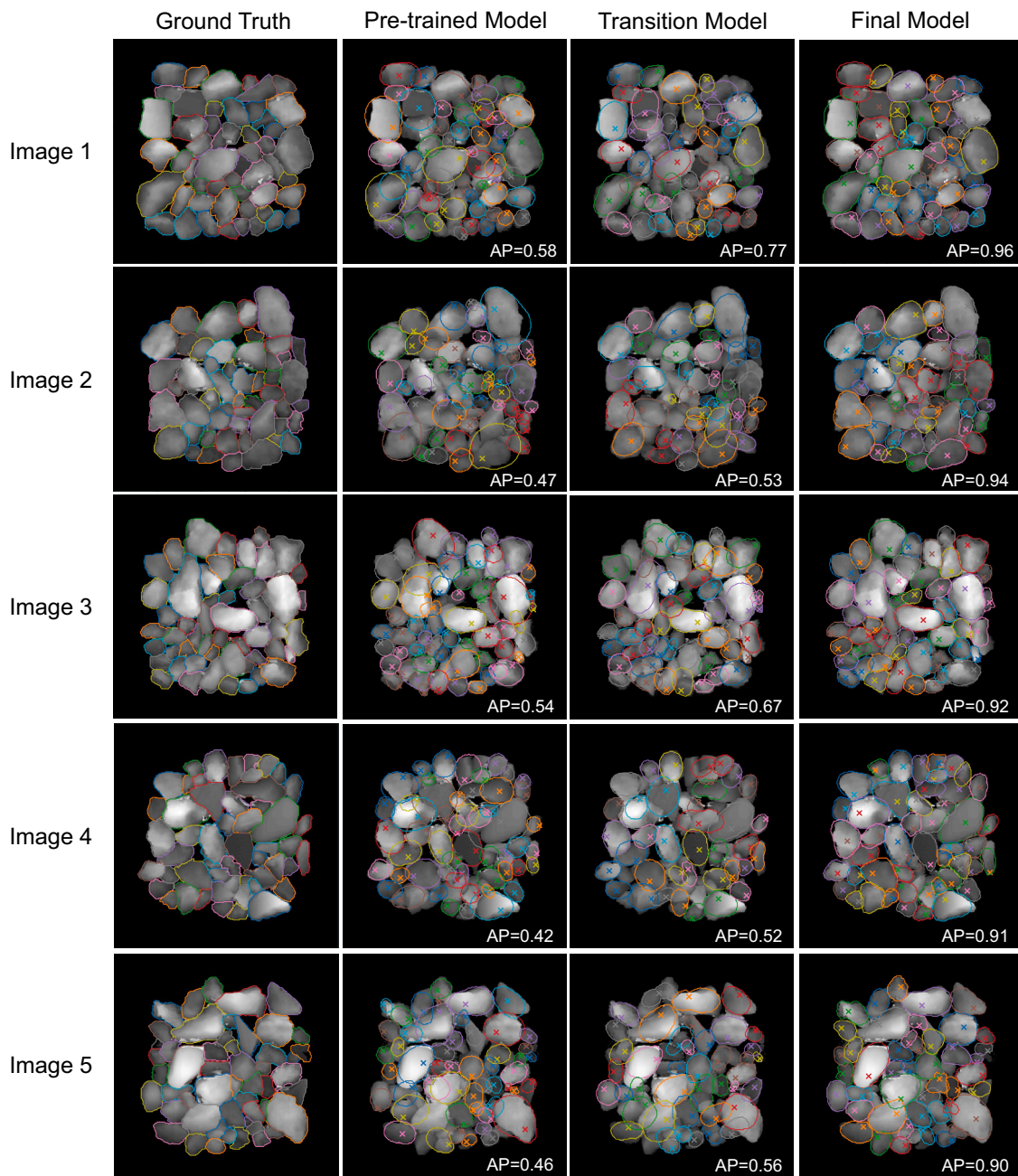


Fig. 2. Segmentation results of the models produced with the transfer learning on test images. AP: average precision at IoU = 0.5.

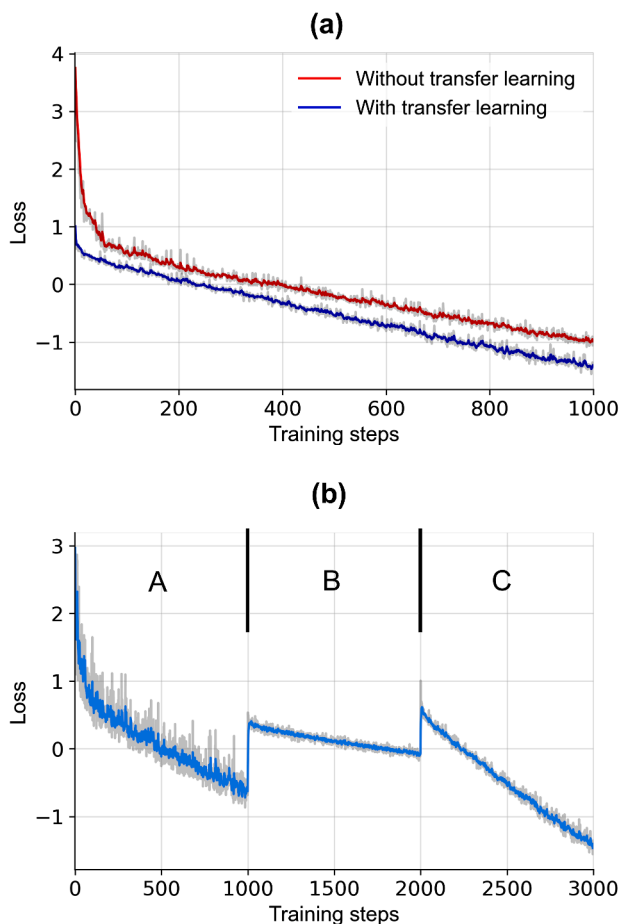
correctly because of a limitation of MultiStar algorithm. As mentioned in Section 2.4.1, there are three important parameters which were used to describe the object. Among them, star distance gives the basic shape of the particle. 32  $r_k$ s are calculated from the starting point and the contours of the objects are formed by connecting the end of the  $r_k$ s. However, this method cannot be applied to describe the concave polygons. As shown in Fig. 4c, when the “corner” of the concave polygon is not covered by the  $r_k$ , the particle shape can be misdescribed. Compared with the ground truth, the prediction of the star distances often lacks a part of the particle. As a result, the particle in the segmentation output of the DL model was always smaller than the real one.

These errors limit the application of the model in certain specific situations. They are the directions for future model improvement, to achieve a higher precision in recycling plants.

#### 3.4. Comparison with traditional segmentation process (RQ4)

To show the advantage of DL based segmentation model, a method based on traditional image processing called Sobel-Watershed method was built in this study. It can be applied to some simple overlapping objects segmentation tasks. However, as the complexity of the object overlap increased, the accuracy of the segmentation result starts to drop down. Besides, the erosion process also changes the original forms of the particles.

The segmentation results of Sobel-Watershed method on the test set were compared with the results of the DL based model. Fig. 5a presents the AP values of the two methods on test images based on the results in Table S5 and S6. As shown in Fig. 5a, the AP values of DL model on all 5 test images are significantly higher than the AP values of Sobel-Watershed method, especially for image 2, the AP value rises from 0.1779 to 0.9412. The segmentation accuracy improves by 429.1%. The



**Fig. 3.** Loss change during the training process. (a) Comparison of the zero-initialized model and the Pre-trained model in loss reduction; (b) Loss change in three different training phases, which are A: DSB2018-based training process, B: CDW-SI-based training process, C: CDW-OL-based training process. (Grey lines: corresponding raw data, colored lines: smoothed data [exponential moving average]).

lowest AP value of DL model occurs in image 5 with 0.9038. However, it is still higher than the best result of the Sobel-Watershed method in image 3, which is 0.4660. The mAP of the DL model on the test set is also 238.4% higher than the mAP of the traditional method. The AP curve of the DL model changes smoothly without noticeable fluctuations when  $\tau = 0.5$ , which proves the DL model maintains a stable performance with different segmentation tasks. For traditional segmentation method, the AP values fluctuate significantly on the test images. When the evaluation

standard  $\tau$  rises to 0.7, the AP value of DL model is still acceptable, but the Sobel-Watershed method can hardly be applied to the segmentation tasks.

Fig. 5b shows the segmentation output of the two methods. Sobel-Watershed method separated the objects on the edge of the material stack well. However, the particles in the center of the material stack, which were contacted or overlapped with other particles, were more likely to be incorrectly recognized. These particles usually had blurred edges and thus the responses of the Sobel operator were small, which resulted in multiple particles being incorrectly identified as one. At the same time, the particles in the center lacked background signals, which play an important role in Watershed segmentation algorithms.

For DL based model, the position of the particles had no influence on the segmentation results. Some particles with specific shapes may cause segmentation errors (see Section 3.3.3), but most of the particles were correctly separated, no matter they were in the middle or on the edge of the material stack. Furthermore, DL model also had the capability to give a prediction of particle shape for overlapping areas, which was obtained in the first two learning processes.

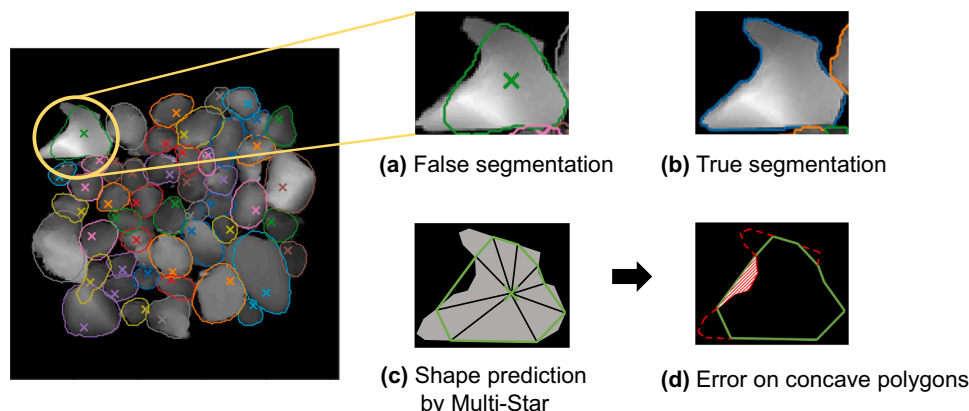
### 3.5. Comparison with other DL-based segmentation model

The *Segment Anything Model* (SAM) is a deep-learning segmentation model developed and recently published by Meta AI (Kirillov et al., 2023). This model is available online for different segmentation tasks, so it was used to make a comparison with the DL model in this study. Fig. 6 shows the segmentation results on the same test image from both models.

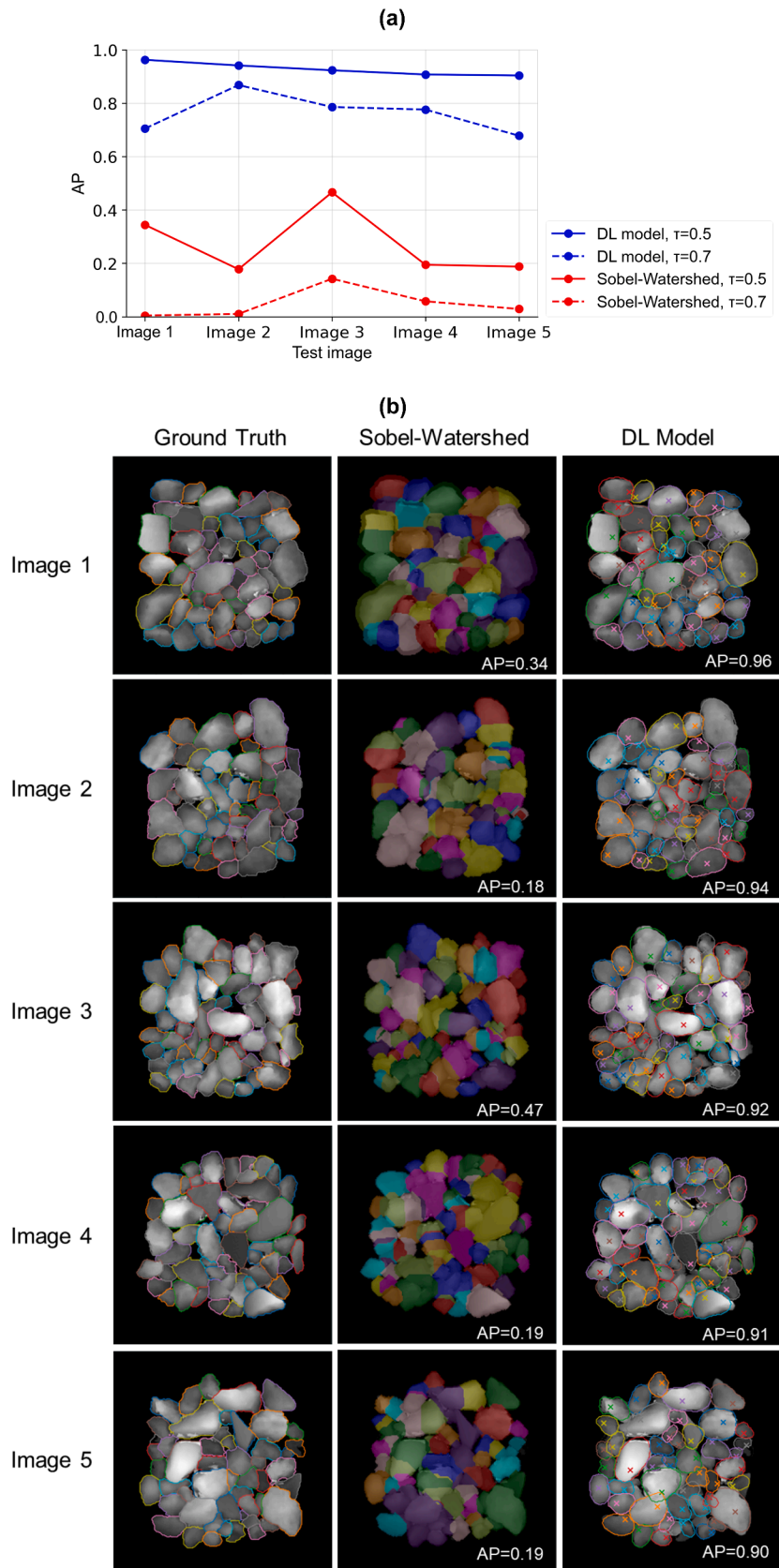
As shown in Fig. 6, the segmentation result from SAM is also very promising. All particles are detected, and the edges of the particles are precisely segmented. However, it also shows disadvantages, which make it less suitable for the task of particle size prediction. First, SAM incorrectly identifies the background signal among the materials as a particle. This is because SAM was not trained on the dataset corresponding to the task of this study. Second, benefitting from the training on DSB-2018 and CDW-SI, MultiStar model remains the ability to predict the overlapping areas between the objects. This ability can be essential in the task of particle size prediction. In future work, the finetuning of foundation segmentation models could thus be an interesting approach to combine the advantages of foundation models with domain-specific datasets and context.

## 4. Conclusion

This work aimed to develop an inline segmentation method for CDW images based on 3DLT measurements under missing particle singulation conditions, to make the recycling procedure more effective. A DL algorithm was chosen to be the segmentation method in this study, since the



**Fig. 4.** Segmentation mistake of concave polygons.



**Fig. 5.** Performance comparison between traditional segmentation method and DL model. (a) Comparison of the AP values of DL model and Sobel-Watershed method on test images with IoU threshold  $\tau = 0.5$  and  $\tau = 0.7$ ; (b) Segmentation results of Sobel-Watershed method and DL model on test images.



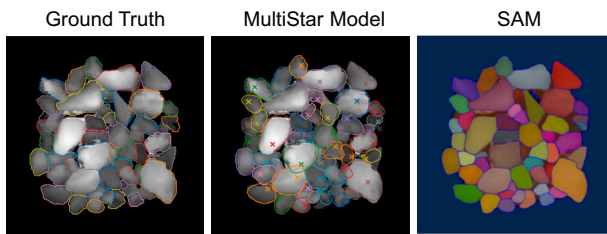


Fig. 6. Segmentation results of two DL-based models on test image.

effect of traditional segmentation methods are limited, in the case of complex particles accumulation conditions. The architecture of the DL model is U-Net (Ronneberger et al., 2015) and the segmentation method used to predict the shape of the particle is MultiStar (Walter et al., 2020). To enrich the data for model training and accelerate the training process, two transfer learning processes based on three datasets were applied. The training process was divided into three phases by transfer learning processes, in which the model was trained sequentially on the three datasets.

The models with different hyperparameters were evaluated and compared with each other. The model with the best performance was selected as the final segmentation model. The mAP of the model is 0.9276 with  $\tau = 0.5$  and 0.4349 with  $\tau = 0.8$ . [RQ1] The transfer learning operations reduced the difficulty during the training process. [RQ2]

However, this model still has some limitations. The creation of the overlapping CDW labels for training is time-consuming, thus this model was not trained with a large amount of data. Random sampling in the algorithm led to instability of segmentation results and made the evaluation difficult. Furthermore, MultiStar method cannot detect the concave polygons correctly. There is still room for improvement in future research. The most important optimization is that the model can be trained on a larger dataset, which consists of more CDW images with different overlapping scenarios. In addition, the DL model could be used to predict the overlapping parts of CDW particles, which would be very beneficial for PSD predictions and would require alternatives techniques (e.g., synthetic data) to label overlapping particle contours. Although the algorithm in this study achieved a promising performance, it is not the only solution for overlapping object segmentation. Besides Multistar and U-Net, other methods can also be combined with other DL models to predict the shape of the objects (Al Arif et al., 2018). [RQ3]

Compared with traditional segmentation methods, the DL based model showed a better performance in both accuracy and robustness. The mAP value (IoU = 0.5) of the investigated traditional segmentation model (Sobel-Watershed) is only 0.2741, while that of DL based model is 0.9276. Besides, the model can also maintain a good performance under stricter evaluation criteria. When the IoU threshold of the evaluation rose to 0.7, the mAP of traditional model declined to 0.0372, while the mAP of DL based model was 0.7625. [RQ4]

The model proposed in this study was trained with two transfer learning processes and three datasets to solve the problem of overlapping materials segmentation. It achieved a high accuracy in the experiments and showed a great potential for CDW segmentation in real working condition. This result provides prospect for the wide application of PSD prediction in waste recycling plants.

## Funding

This work was funded by the National Austrian Research Promotion Agency (FFG) within the program "Production of the Future" under the project EsKorte (grant no. 877341) and the German Federal Ministry of Education and Research (BMBF) within the program "Digital GreenTech" under the project KIMBA (grant no. 2WDG1693D). Open access funding was provided by the RWTH Aachen University Open Access Publication

Fund. The responsibility for the content of this publication lies with the authors.

## CRedit authorship contribution statement

**Xiaoye Wu:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Nils Kroell:** Conceptualization, Data curation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Kathrin Greiff:** Resources, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgements

We would like to thank MAV Krefeld GmbH for providing the CDW sample material.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.resconrec.2024.107541](https://doi.org/10.1016/j.resconrec.2024.107541).

## References

- Al Arif, S.M.M.R., Knapp, K., Slabaugh, G., 2018. SPNet: shape prediction using a fully convolutional neural network. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. Springer International Publishing, Cham, pp. 430–439.
- Alloui, H., Mourdi, Y., Sadgal, M., 2023. Strong semantic segmentation for Covid-19 detection: evaluating the use of deep learning models as a performant tool in radiography. *Radiography* 29 (1), 109–118. <https://doi.org/10.1016/j.radi.2022.10.010>.
- Arnab, A., Torr, P.H.S., 2017. Pixelwise instance segmentation with a dynamically instantiated network. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017/7/21–2017/7/26. IEEE, Honolulu, HI.
- Da Leite, F.C., RdS, Motta, Vasconcelos, K.L., Bernucci, L., 2011. Laboratory evaluation of recycled construction and demolition waste for pavements. *Constr. Build. Mater.* 25 (6), 2972–2979. <https://doi.org/10.1016/j.conbuildmat.2010.11.105>.
- Data Science Bowl, 2018. Find the nuclei in divergent images to advance medical discovery. Dataset. <https://www.kaggle.com/competitions/data-science-bowl-2018/data>.
- Di Maria, F., Bianconi, F., Micale, C., Baglioni, S., Marianni, M., 2016. Quality assessment for recycling aggregates from construction and demolition waste: an image-based approach for particle size estimation. *Waste Manag.* 48, 344–352. <https://doi.org/10.1016/j.wasman.2015.12.005>.
- Kronenwett, F., Maier, G., Leiss, N., Gruna, R., Thome, V., Längle, T., 2022. Sensorgestützte Charakterisierung von Bauschutt bei hohen Belegungs dichten mittels deep learning. In: Pomberger, R., Adam, J., Altendorfer, M., Bouvier-Schwarz, T. (Eds.), *VORTRÄGE-Konferenzband Zur 16. Recy & DepoTech-Konferenz*, pp. 293–298.
- DIN, 66165-1, 2022. Particle size analysis - Sieving analysis - Part 1: Fundamentals. Deutsches Institut für Normung. Beuth Verlag, Berlin.
- Forschungsgesellschaft für Straßen- und Verkehrswesen, 2020. Technische Lieferbedingungen für Baustoffgemische Zur Herstellung von Schichten ohne Bindemittel im Straßenbau: TL SoB-StB 20, 2020th ed. FGSV Verlag GmbH, Köln.
- Forschungsgesellschaft für Straßen- und Verkehrswesen, 2023. Technische Lieferbedingungen für Gesteinskörnungen im Straßenbau: TL Gestein-StB 04/23, 2004th ed. FGSV Der Verlag, Köln.
- German Building Materials Association. Daten und Fakten zur Baustoff-Steine-ErdenIndustrie; 2020.
- Jähne, B., 2012. *Digitale Bildverarbeitung*. Springer Berlin Heidelberg, Berlin, Heidelberg.

- Kandlbauer, L., Khodier, K., Ninevski, D., Sarc, R., 2021. Sensor-based particle size determination of shredded mixed commercial waste based on two-dimensional images. *Waste Manag.* 120, 784–794. <https://doi.org/10.1016/j.wasman.2020.11.003>.
- Kendall A., Gal Y., Cipolla R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics; 2017.
- Khodier, K., Viczek, S.A., Curtis, A., Aldrian, A., O'Leary, P., Lehner, M., et al., 2020. Sampling and analysis of coarsely shredded mixed commercial waste. Part I: procedure, particle size and sorting analysis. *Int. J. Environ. Sci. Technol.* 17 (2), 959–972. <https://doi.org/10.1007/s13762-019-02526-w>.
- Kirillov A., Mintun E., Ravi N., Mao H., Rolland C., Gustafson L., et al. Segment Anything; 2023.
- Kroell, N., Chen, X., Greiff, K., Feil, A., 2022. Optical sensors and machine learning algorithms in sensor-based material flow characterization for mechanical recycling processes: a systematic literature review. *Waste Manag.* 149, 259–290. <https://doi.org/10.1016/j.wasman.2022.05.015>.
- Kroell, N., Chen, X., Maghmoumi, A., Koenig, M., Feil, A., Greiff, K., 2021. Sensor-based particle mass prediction of lightweight packaging waste using machine learning algorithms. *Waste Manag.* 136, 253–265. <https://doi.org/10.1016/j.wasman.2021.10.017>.
- Lu, Z., Carneiro, G., Bradley, A.P., Ushizima, D., Nosrati, M.S., Bianchi, A.G.C., et al., 2017. Evaluation of three algorithms for the segmentation of overlapping cervical cells. *IEEE J. Biomed. Health Inform.* 21 (2), 441–450. <https://doi.org/10.1109/JBHI.2016.2519686>.
- Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., Terzopoulos, D., 2022. Image segmentation using deep learning: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 44 (7), 3523–3542. <https://doi.org/10.1109/TPAMI.2021.3059968>.
- Mu Zhu. Recall, Precision and Average Precision 2004.
- Ni, X., Li, C., Jiang, H., Takeda, F., 2021. Three-dimensional photogrammetry with deep learning instance segmentation to extract berry fruit harvestability traits. *ISPRS J. Photogramm. Remote Sens.* 171, 297–309. <https://doi.org/10.1016/j.isprsjprs.2020.11.010>.
- Qurratulain, S., Zheng, Z., Xia, J., Ma, Y., Zhou, F., 2023. Deep learning instance segmentation framework for burnt area instances characterization. *Int. J. Appl. Earth Obs. Geoinf.* 116, 103146 <https://doi.org/10.1016/j.jag.2022.103146>.
- Roerdink, J.B., Meijster, A., 2000. The watershed transform: definitions, algorithms and parallelization strategies. *Fundam. Inform.* 41 (1,2), 187–228. <https://doi.org/10.3233/FI-2000-411207>.
- Ronneberger O., Fischer P., Brox T.. U-Net: convolutional networks for biomedical image segmentation; 2015.
- Schmidt, U., Weigert, M., Broaddus, C., Myers, G., 2018. Cell detection with star-convex polygons. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. Springer International Publishing, Cham, pp. 265–273.
- Shapiro, L.G., Stockman, G.C., 2001. *Computer Vision*. Prentice Hall, Upper Saddle River, NJ.
- Statistisches Bundesamt. Abfallbilanz (Abfallaufkommen/-verbleib, Abfallintensität, Abfallaufkommen nach Wirtschaftszweigen), 2021.
- Walter F.C., Damrich S., Hamprecht F.A. MultiStar: instance segmentation of overlapping objects with star-convex polygons; 2020.
- Zhang, Y., Liang, Y., 2016. Research on soft-sensor based on support vector regression for particle size of grinding and classification process. 2016 Chinese Control and Decision Conference (CCDC); 2016/5/28 - 2016/5/30. IEEE, Yinchuan, China.