crystals

MDPI

Supporting Information

Investigation of AI Algorithms for Photometric Online Analysis in a Draft Tube Baffle Crystallizer

Laura Marsollek [1], Julius Lamprecht [1] and Norbert Kockmann [1,*]

Section S1: Key physical and chemical properties of used L-alanine

According to Wohlgemuth and Schembecker [1]:
CAS Number 56-41-7
Purity 98% (Carbolution)
Molar Mass 89.09 g/mol
Density 1.432 g/cm3 (at 22 °C)
Melting Point 314.6 °C (decomposition)
pH Value 5.5 - 7.0 (89.1 g/l, H2O, 25 °C)
Solubility >160 g/l

Table S2: Equipment and manufacturer information of all equipment used in the experimental setup.

Equipment	Title 2	Title 3
Thermostat	Ministat 230	Huber
Peristaltic pump	LabDos	HighTech Zang
Microscope	Advanced ICD	Bresser
Camera	Z6-II	Nikon
Stirrermotor	RZR 2102	Heidolph
Graphic card	GeForce RTX 4070 Super / RTX 5000	Nvidia

Section S3: Further information analytic setup

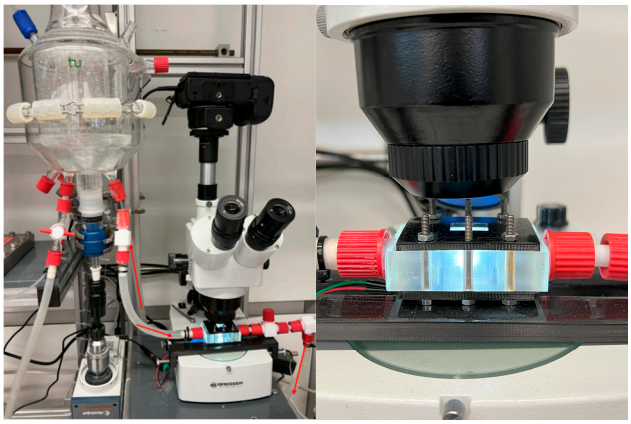
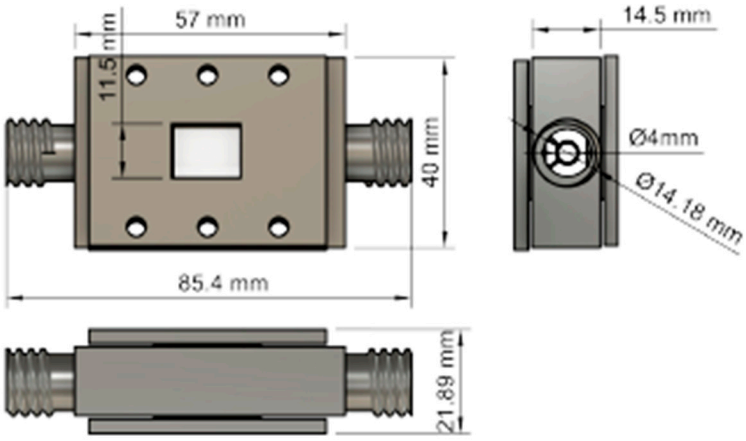


Figure S3: Picture of the photometric analytics in detail. Picture of the experimental setup with drawn flow direction. The temperated water from the thermostat follows the red arrows. The intern pump conveys the water upwards into the double jacket tube. There it flows through the analytics and the flow cell and leaves at the last t-piece. It then flows back into the thermostat (left), Close-up of the flow cell (right)

In addition to the temperature control of the L-alanine suspension, rinsing the FEP tube with water also has a positive effect on the generation of the training images. Due to the approximately equal refractive index of FEP ($n_{D,FEP}$= 1.344 (20°C) [2]) and water ($n_{D,W}$= 1.333 (20°C) [3]), the distortion otherwise caused by the curvature of the tube is almost

eliminated. Since the LED is wider than the crop of the microscope, the assumptions is made that the light going through the cell is parallel. So no correction factor needs to be considered. The camera was controlled via the DigiCamControl program. For 8x magnification, a 2x magnifying lens from Bresser was used, in addition to the 4x magnification of the objective.

Figure S4: Technical drawing developed flow cell. Created with Autodesk Fusion 360



Section S5: Calibration Camera

To calibrate the camera to the center of the tube and calculate the ratio of form pixels to µm, photos were taken of graph paper at tube height. To do this, the camera was first focused on the tube in the flow cell. The flow cell was then removed and replaced with a 2 mm high cuboid out of PLA with graph paper placed on top so that the graph paper was at the same height as the tube beforehand. A total of 5 photos were taken and 50 reference lines were drawn in ImageJ.

Table S6: Added L-alanine in performance tests

$m_{w,L-alanine}$ [%]	$m_{L-alanine}$ added [g]
1.00	0.31
2.50	0.78
5.00	1.57
7.50	2.32

Section S7: Definition of classes for YOLO models

The definition of the classes is one of the most challenging tasks during the implementation of an AI model for crystal analysis, as it has a major impact on the performance of the models. One of the most common errors in crystal analysis can be avoided by appropriate classification. Over- or undersegmentation can occur during crystal detection due to insufficient differentiation of individual crystals. In this case, crystals are incorrectly divided into several small parts or several crystals are detected as a single crystal. Therefore, blurred or not clearly separable crystals are not classified as crystals in YOLOv8 and YOLO8seg, but assigned to the *Hidden* class and thus not considered in the crystal size analysis. Figure S4 shows examples of crystals with bounding boxes and classifications.

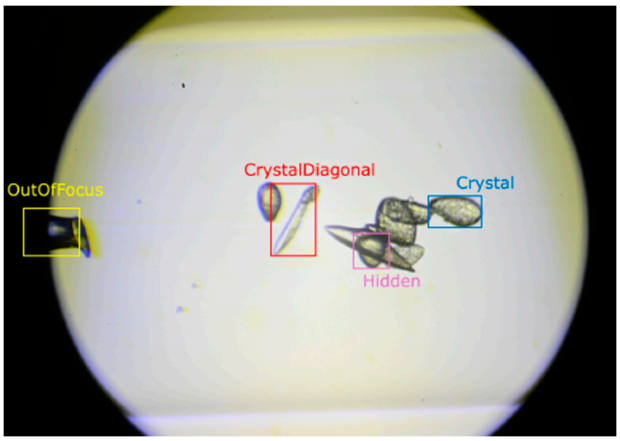


Figure S7: Example annotation of each of the four classes for YOLO.

Through the addition of the *OutOfFocus* class, a further source of error was eliminated. The class is assigned to all crystals that are located at the edge of the field of view in the microscope and are therefore not fully visible.

As the crystals in YOLOv8 are marked by bounding boxes, the longest side of the bounding box is determined as the Feret diameter. Due to the irregularly shaped crystals, a deviation from the actual length of approx. 8% occurs. By introducing the *CrystalDiagonal* class, a non-parallel orientation of the crystal to the flow direction was considered and the Feret diameter was calculated on the basis of the diagonal using Pythagoras' theorem. Thus, the systematic error could be reduced to 6%.

Section S8: Evaluation of the model training

The first step is to evaluate the models based on their training results. The box loss and class loss across various training epochs are used to observe the success of the training. An increasing discrepancy between training losses and validation losses enables identification of the first signs of over- or underfitting.

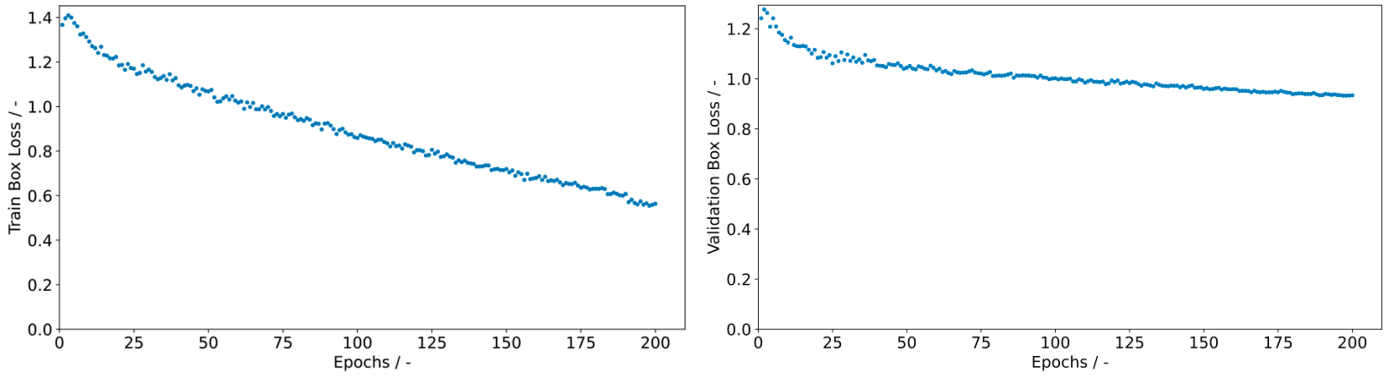
The Mean Average Precision (mAP_i) is calculated in particular for the evaluation of crystal detection. With the mean average precision at an intersection over union (IoU) threshold of 0.5 ($mAP50_i$) the performance is compared between the crystal size ranges and the solid contents. For further specification, images of only one class j can also be considered, resulting in the $mAP50_{i,j}$. For the assessment of the balance between precision and recall, the confusion matrix and the precision-recall curve are examined. To evaluate the performance of the model with these weights across different classes, the normalized confusion matrix is used. The precision-recall curve summarizes the metrics for precision and recall at all confidence levels and is determined using the validation dataset after training. The area under the curve indicates the AP_i.

Section S9: Training of YOLOv8

For the training, a pre-trained on the COCO dataset version of the YOLOv8 algorithm from Ultralytics was used. The training was based on 1372 photos (augmented). Since the computational effort increases significantly with the resolution [4], a compromise between accuracy and speed was found with an image size of 1280x840 pixels. The maximum duration of the training was 200 epochs, but the training was terminated as soon as no improvements could be detected within 50 epochs. In order to utilize the entire available VRAM memory without exceeding the capacity of the graphics card, the batch size was set to *autobatch*. In this state, the program automatically checks beforehand how many images can be processed simultaneously and sets the batch size accordingly. The training runs on a Nvidia GeForce RTX 4070 Super or a RTX 5000. Both graphic cards provide 16 GB of video RAM. The available RAM is 32 GB.

The YOLOv8 model's training showed decreasing box and class losses across 200 epochs, indicating improved object recognition and classification. This improvement was confirmed by the validation dataset, where the $mAP50_{BB}$ increased from 20% to 76% across the training epochs. For this reason, the training was considered successful, with no significant signs of overfitting, and the training duration was sufficient. More detailed information to the trainings results of YOLOv8 are given in Supporting Figure S10-S12.

Figure S10: YOLOv8 Box loss score on training dataset (a) and on validation dataset (b) over YOLO training epochs.



(a) (b)

Figure S11: YOLOv8 mAP50BB score over training epochs

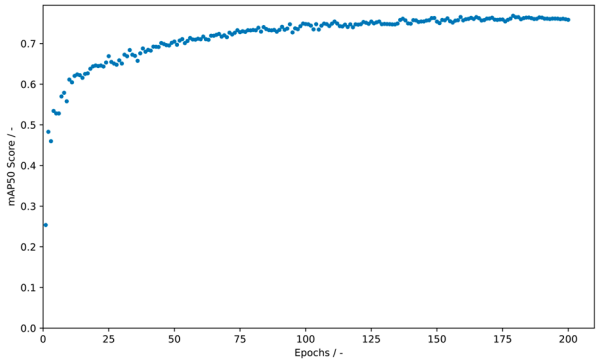


Figure S12: YOLOv8 Confusion matrix. Evaluated on validation dataset after training

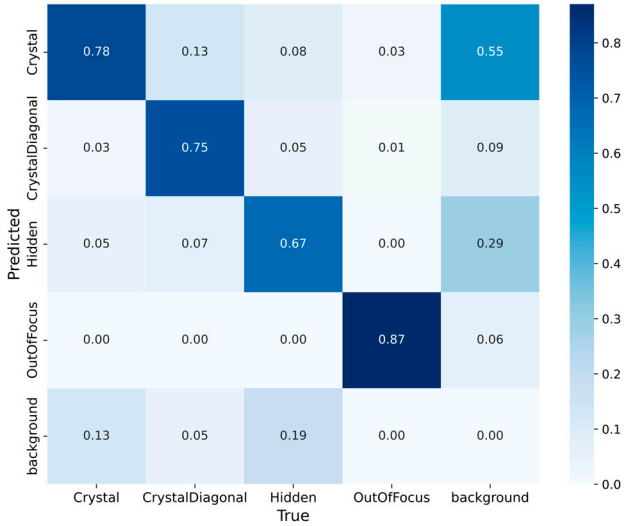
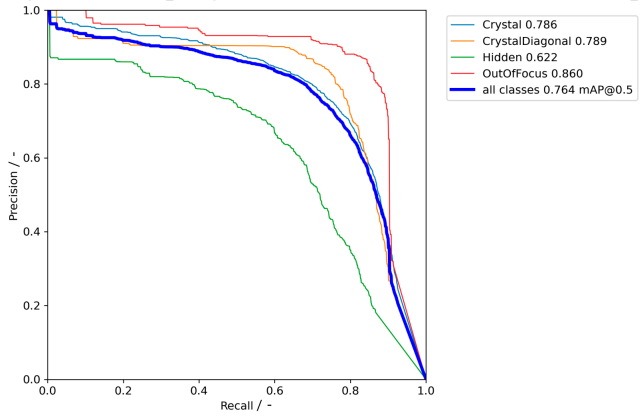


Figure S13: YOLOv8 Precision Recall curve. Evaluated on validation dataset after training. On the top right, the AP50$_{BB}$ and mAP50$_{BB}$ is displayed



Section S14: Crystal detection and classification performance of YOLOv8

After the training of the YOLOv8 model was concluded, its performance in the detection of crystals and its classification was evaluated. This step is important to identify strengths and weaknesses of the model and to show optimization potential. The results of this evaluation are presented below, the performance evaluation can be found in Figure S15-S17.

For the evaluation of the performance of the trained YOLOv8 model the normalized confusion matrix was used. Overall, most crystals were correctly identified. There is only a slight risk of confusion between the classes. The most relevant classes *Crystal* and *CrystalDiagonal* have a score of 0.78 and 0.75 respectively. For a further evaluation of the detection and classification, the precision recall curve for the four classes was considered. The model shows strong performance for the *Crystal* class, with an $AP50_{BB}$ of 76.6%, indicating stable precision and sensitivity, and a low false positive rate. Similar results are seen for the class *CrystalDiagonal* class. The class *Hidden* performs worst, with a PR value of 62.2%, high false positive (12%) and false negative (41%) rates, though this class is less critical. The *OutOfFocus* class performs best with an $AP50_{BB}$ of 86%. The overall $mAP50_{BB}$ across all classes is 76.4%, reflecting a good balance between precision and recall across the major classes.

For a specific evaluation of the trained model performance in crystal detection and classification, the $AP50_{BB,Crystal}$ values of the YOLOv8 model are evaluated across different crystal sizes. The YOLOv8 model performs well for larger crystal sizes (180–630 µm), with high $AP50_{BB,Crystal}$ scores of 0.702 for the crystal range 180-250 µm and 0.972 for 250-630 µm. However, performance significantly declines as crystal size decreases. By considering the normalized confusion matrix for the crystal size range of 0-90 µm, it is apparent that 90 % of the *Crystal* class are classified as *Background*, which indicates a low recall (many false-negative results). Furthermore, 58% of the class *Background* are recognized as *Crystal*, meaning precision (high false positives) is also low. Overall, the model is balanced for the upper crystal size ranges in terms of precision and recall, but has significant weaknesses in recognizing small crystals.

Another important factor that influences recognition is the number of crystals per image. It is expected that the YOLO model will perform better if there are fewer crystals per image, as there will be less overlapping. With overlapping crystals, it is more difficult for the models to distinguish the crystals from each other.

The $AP50_{BB,Crystal}$ score is analyzed at different mass fractions. At a low w_{ala} of 1%, the score starts at 77.1% and peaks at 90.3% for a w_{ala} of 2.5%. With a further increase in w_{ala}, $AP50_{BB,Crystal}$ decreases sharply to 53.3 % at 5 % and 44.9 % at 7.5 %. This decrease is due to increasing overlap of crystals, where densely filled images lead to misclassification to the *Hidden* class, which reduces detection performance. At lower mass fractions, only a

few crystals are present, making it more difficult for the model to identify positive examples. In addition, the influence of the background increases, leading to false detections. Furthermore, a low number of crystals increases the statistical impact of errors, which further decreases the $AP50_{\mathrm{BB,Crystal}}$ value.

Figure S15: YOLOv8 Influence of crystal size on performance. AP50$_{\mathrm{BB,Crystal}}$ from YOLO8seg model on 20 images in each crystal size class with example images

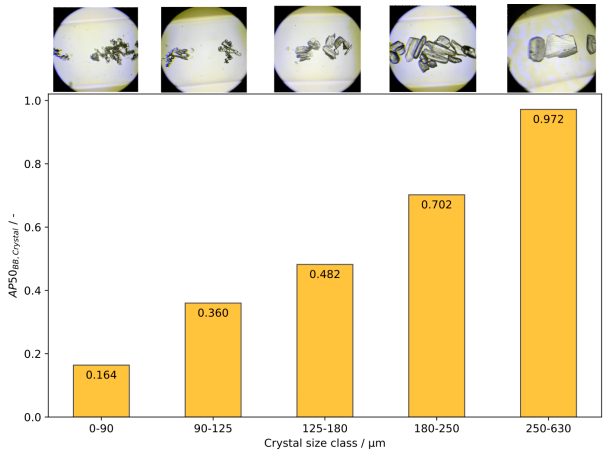


Figure S16: - YOLOv8 Influence of crystal size on performance. Normalized confusion matrix from YOLO model. Evaluated on 20 images in the crystal size class 0-90 µm

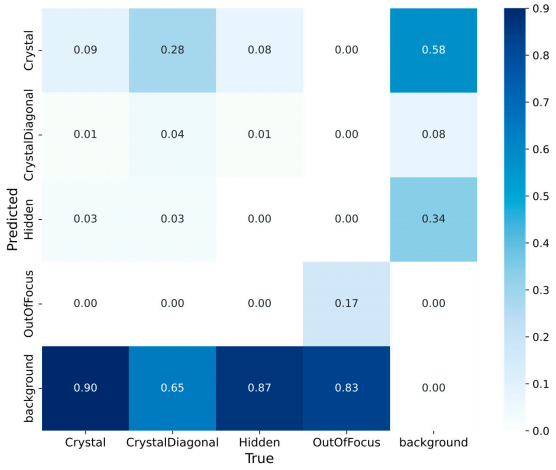
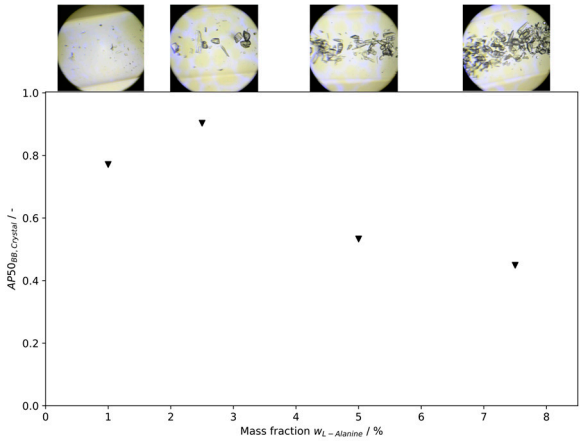


Figure S17: - YOLOv8 Influence of crystal number on performance. AP50$_{\mathrm{BB,Crystal}}$ from YOLO model on 20 images in $w_{\mathrm{L-alanine}}$ from 1% to 7.5%



Section S18 - Training of YOLO8seg

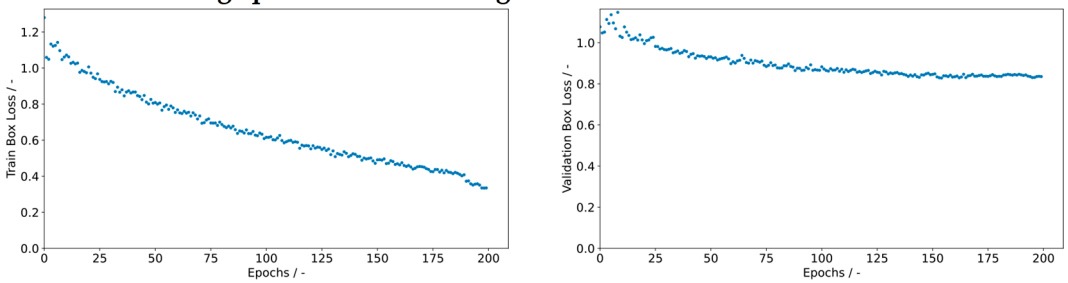
The training of the YOLO8seg model is based on a model pre-trained on the COCO dataset for segmentation tasks. For the implementation, the Ultralytics library version 8.0.28 was used [5]. Since the models do not differ fundamentally, the same training parameters were applied as for YOLOv8. The batchsize is set to *autobatch* and the model is trained for 200 epochs with an image size of 1280x840 pixels. The training lasted 22 hours for all 200 epochs and included 1110 labeled images (augmented) with a 70/30 (training/validation) split.

For crystal detection and classification, bounding boxes and masks in the form of polygons are recognized by YOLO8seg. The Feret diameter is determined in the same way as for manual image evaluation using the rotating calipers method, as described in Chapter 3.1. Similar to YOLOv8, the performance is evaluated by calculating *mAP50* and *mAP50-95*. Since YOLO8seg recognizes both bounding boxes and masks, a separate *mAP50* score, $mAP50_{BB}$ for bounding boxes and $mAP50_{M}$ for masks, is calculated. This enables a separate evaluation of general detection and precise segmentation

For the Bounding Boxes (BB) loss curves a decreasing box and class losses was shown during the training of YOLO8seg across the 200 epochs, indicating improved object recognition and classification. This improvement was confirmed by the validation dataset, where the $mAP50_{BB}$ increased to a score of 83.8% across all classes on the 278 validation images. The $mAP50-95_{BB}$ score reached a value of 63.8%, with the $mAP50-95_{BB,Crystal}$ being higher at 72.4%. For this reason, the training was considered successful, with no significant signs of overfitting, and the training duration was sufficient.

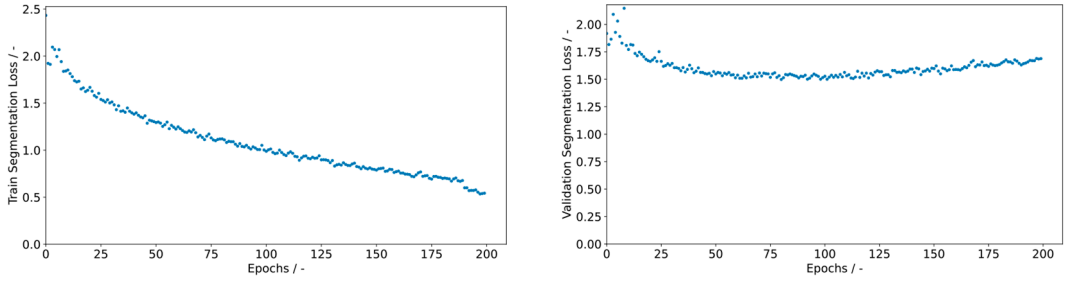
For the loss curves of the instance segmentation (Masks) a different trend was visible. Although training success was initially apparent, the curve of validation losses began to increase after epoch 100. Along with the significant difference between training and validation losses, this is an indication of overfitting. The training was therefore terminated earlier. For the $mAP50_{M}$ the same trend can be observed as for the $mAP50_{BB}$. Despite the premature termination of training, a value of 82.3% was achieved. According to the Pascal VOC standard, a good range was thus reached [6]. Further information to the trainings results of YOLO8seg are given in Figure S19-S23.

Figure S19: YOLO8seg Loss curve of the Bounding Boxes over the course of the training epochs on the training dataset (a) and the validation datatset (b)



(**a**) (**b**)

Figure S20: YOLO8seg Loss curve of the instance segmentations (masks) over the course of the training epochs on the training dataset (a) and the validation datatset (b)



(**a**) (**b**)

Figure S21: YOLO8seg mAP_{BB} score. Measured on the validation dataset over the course of each training epoch

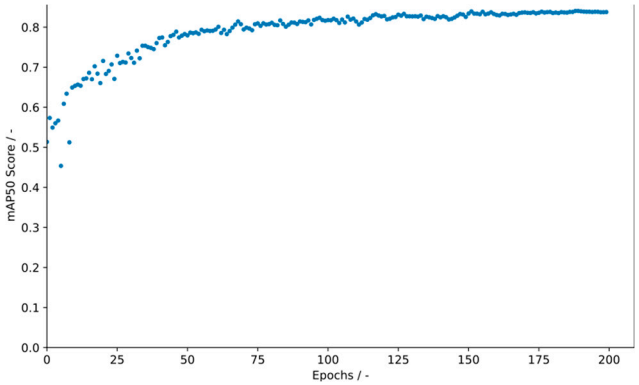


Figure S22: YOLO8seg Mask Precision-Recall curve. Evaluated on the validation dataset. The values in the top right represent the AP_{Mask} of each class

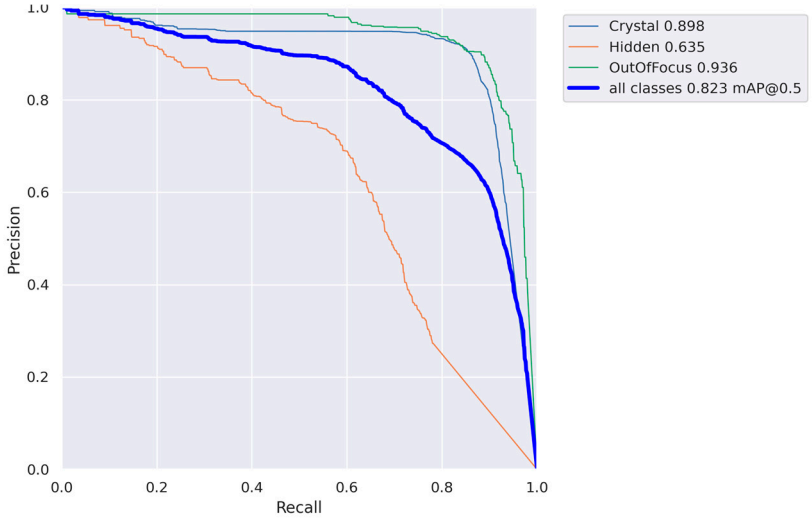
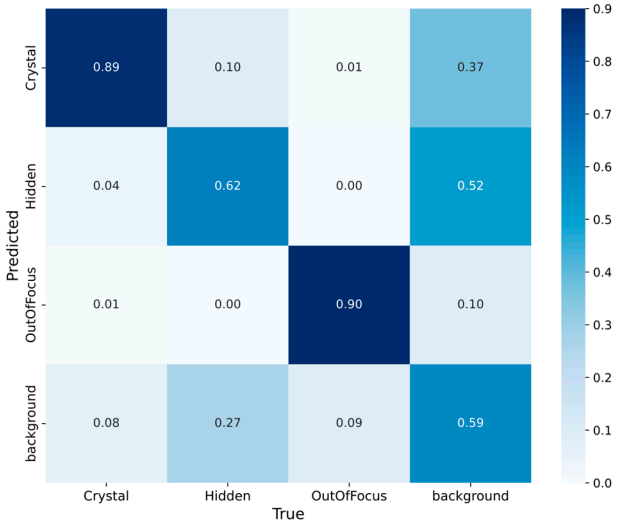


Figure S23: YOLO8seg Confusion matrix of the trained YOLO8seg model. Measured with the best performing model on the validation dataset. The values are normalized on the total number of true detections.



Section S24 - Crystal detection and classification performance of YOLO8seg

The performance of the trained YOLO8seg Model is also evaluated for the detection and classification of crystals. First, the precision-recall-curve with the resulting $AP50_{M}$ value for all classes is examined. The diagrams to this chapter are given in the Figure S25-S27.

The YOLO8seg model achieves a high $AP50_{M}$ of 89.8% for the *Crystal* class, indicating strong stability with high precision and recall. For the class O*utOfFocus* an $AP50_{M}$ of 93.6% was achieved, consistent with standard YOLO results. The precisision-recall curves also shows, that as confidence increases, precision increases while recall decreases, highlighting the need for a balance. The confusion matrix confirms high true positive rates and low false negative rates for the *Crystal* (89%) and *OutOfFocus* (90%) classes, with only 8% of true crystals missed. The *Hidden* class shows poor performance for both metrics, with frequent misclassifications. Since these errors are not included in the crystal size analysis, this class prevents distortion of the crystal size distribution by crystals that are difficult to recognize. To further assess model performance, an analysis across crystal size ranges and w_{ala} will be conducted.

The analysis of the $AP50_{BB,Crystal}$ confirms that YOLO8seg performs better with larger crystal sizes, achieving an $AP50_{BB,Crystal}$ of 90.6% in these groups. For smaller crystals, the score decreases on the one hand due to a higher number of false positives at a lower confidence level and on the other hand due to increased false negatives, which lead to a lower recall value. The confusion matrix for the crystal size range 0-90 µm shows that 60% of the crystals are misclassified as *Background*, with 32% as false positives, indicating a bias toward detecting larger crystals. However, the YOLO8seg model still outperforms the standard YOLOv8 model on smaller crystals, with a higher number of true positive *Crystal* detections.

For the analysis of the impact of the number of crystals on the performance of YOLO8seg the $AP50_{BB,Crystal}$ was examined. As seen before for YOLOv8, $AP50_{BB,Crystal}$ decreases both with very low and very high crystal counts. At low w_{ala}, there are too few crystals, so that misidentification significantly impairs detection accuracy. At high mass fractions, the large number of crystals causes increased overlap, which also reduces detection accuracy. This trend mirrors the performance observed with the standard YOLOv8 model, indicating that the YOLO8seg faces similar challenges when dealing with images containing either very few or very many crystals.

Figure S25: YOLO8seg Influence of crystal size on performance. $AP50_{BB,Crystal}$ from YOLO8seg model on 20 images in each crystal size class with example images

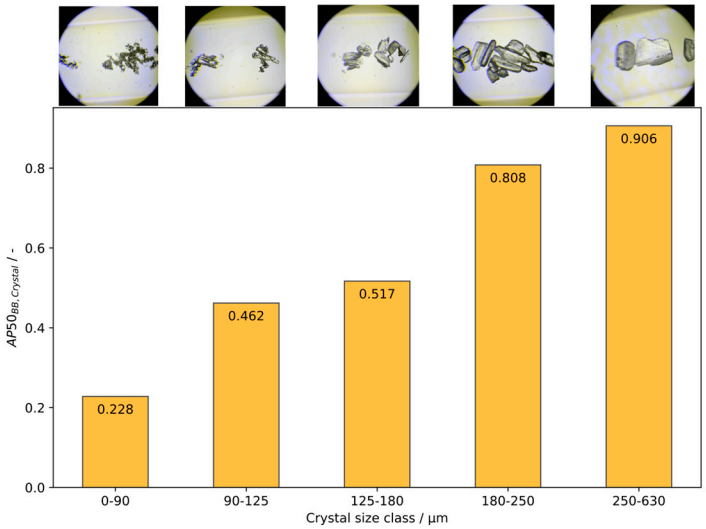


Figure S26: YOLO8seg Influence of crystal size on performance. Confusion matrix from YOLO8seg model. Evaluated on 150 images form the crystal size class 0-90 µm

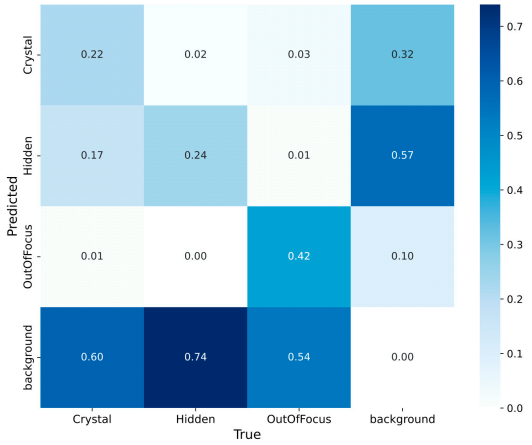


Figure S27: YOLO8seg Influence of crystal number on performance. $AP50_{BB,Crystal}$ from YOLO8seg model on 20 images in $w_{L-alanine}$ from 1% to 7.5%

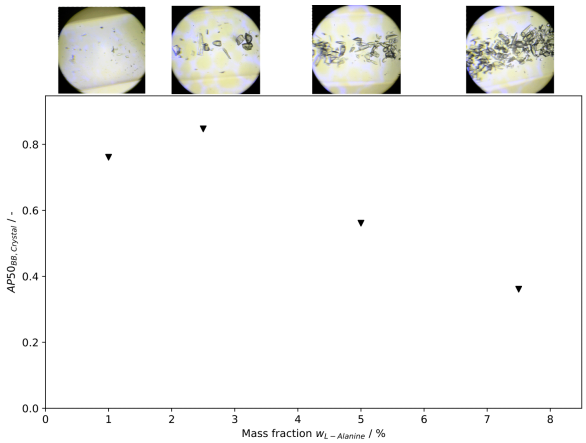


Figure S28: Example Images of YOLO8seg

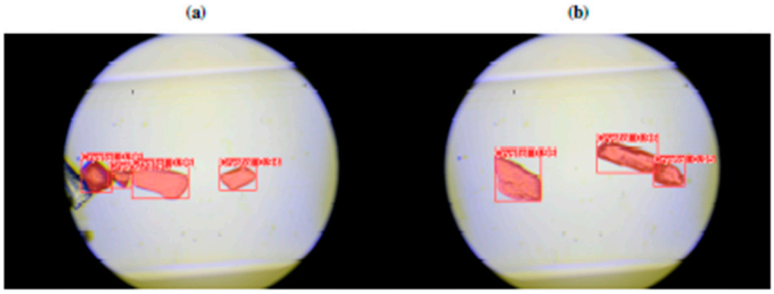


Figure 61: Example images of YOLO Segmentation masks on crystal size class 250-630μm

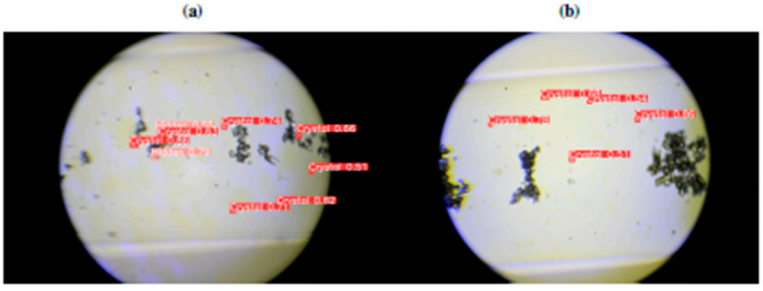


Figure 62: Example images of YOLO Segmentation masks on crystal size class 0-90μm

Section S29 - Training U-Net

The training of the U-Net model is more complex than that of the YOLO model. Since there is no existing library, the training model must be prepared manually before training can take place. In order not to exceed the calculation limit, the maximum image size is 1024 x 1024 pixels. The aim of the training is to define the optimization algorithm, the loss function and to set the training parameters.

Since the U-Net is merely capable of segmentation rather than classification, the loss function of the U-Net model is a binary cross-entropy loss function for binary classification. The Adam optimizer from TensorFlow [7] is used to optimize the model weights. Adam offers advantages such as adaptive learning rates and exponential moving averages of the gradients. It requires less memory and less hyperparameter tuning than other optimizers despite faster convergence. The image masks are not provided as polygons in the form of a text file, but as binary masks that are read in during training. For this purpose, the polygons provided are first read in and translated into masks. These are then saved as tif files.

A total of 347 images are used for training the U-Net, which are divided in a ratio of 70:30 (train-validation). The training runs for a total of 100 epochs with an unadjusted learning rate of lr = 1e-4. The stack size is set to 4 due to the increased computational load. The training time of the model is about 30 hours.

Table S30:U-Net processing steps and required process time. Captured on average in Python.

Step	Process time [s]
Cut	< 1
U-Net detection	19
Reconstruction	< 1
Edge detection and CSD determination	15

Section S31 – Comparison Training Results

After the training, the YOLOv8 model achieves a value of 76 % for the cross-class $mAP50_{BB}$, while YOLO8seg reaches a value of 82.3 %. For the most important class *Crystal*, YOLOv8 achieves an $AP50-95_{BB,Crystal}$ of 53.3 % and is outperformed by YOLO8seg with an $AP50-95_{BB,Crystal}$ of 63.8 %. Throughout the training, YOLO8seg showed higher precision and recall, resulting in better overall performance. The poorer performance of YOLOv8 could be due to problems in differentiating between the *Crystal* and *CrystalDiagonal* classes, as some of them differ only minimally in the orientation of the crystals (see Chapter 3.1). The class *CrystalDiagonal* is not required in YOLO8seg.

In terms of crystal size, a strong improvement in performance was observed for both models with larger crystals in the form of an increasing mAP50BB. The maximum mAP50BB score is achieved by YOLO for crystal sizes between 250-630 µm, with an almost perfect detection at an $mAP50_{BB}$ of 96.5 %. YOLO8seg also achieves the best results here, with an $mAP50_{BB}$ of 90.6 %. Both models thus detect an almost identical crystal size distribution as in the manual analysis. For the medium crystal size range, both models show good results at a confidence level of 0.7, which were confirmed by the two-sided t-test with non-significantly differing mean values. The crystal size range of 0-90 µm is the most difficult for the detection of crystals, as a distinction must be made on the basis of a few pixels. A physical limit for crystal detection was determined at 10 µm. With increasing confidence level, the detection rate in this crystal size range decreased significantly for both YOLO models. YOLOv8 was no longer able to detect any crystals at a confidence level of 0.9. However, re-presentative crystal size distributions could still be determined for average confidence levels of 0.7

During the extensive testing of the developed AI models on generated crystal images, it became evident that the U-Net model could not fulfill the desired expectations for online

crystal sizing, as it had too high computational complexity and very inaccurate CSDs. Therefore, it could not compete with the other models and was not further considered. In contrast the YOLOv8 and YOLO8seg models exceeded the expectations. YOLOv8 proved to be efficient for execution of CSD determination, but was surpassed by YOLO8seg in terms of precision and accuracy. Thus, the developed models were applicable within specified limits. While crystal detection works excellently for larger crystals, the detection of smaller crystals remains an area for improvement. The developed YOLO8seg model achieves $mAP50_{BB}$ of 82.3% on the validation dataset. YOLOv8 also performs well, achieving $mAP50_{BB}$ of 76%. Both models perform worse when crystals are too small < 90 µm. Here the $mAP50_{BB}$ drops to 22.8% for YOLO8seg or 16.4% for YOLOv8. Additionally when the number of crystal becomes too high, w_{ala} >= 7.5%, YOLO8seg drops to an $mAP50_{BB}$ 36.1%, while YOLO reaches 44.9%.

Figure S32: Manual Image Analysis Results for crystal size range 250-630 µm

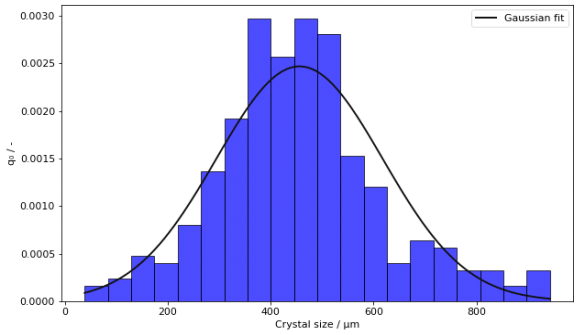


Figure S33: Manual Image Analysis Results for crystal size range 0-90 µm

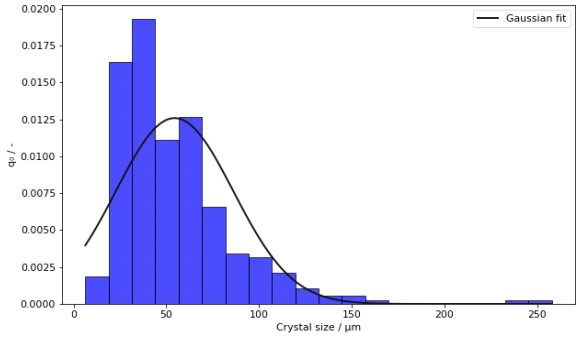


Figure S34: YOLOv8 Boxplot comparing CSD for crystal size range 0-90 µm using ImageJ and YOLOv8 at different confidence levels

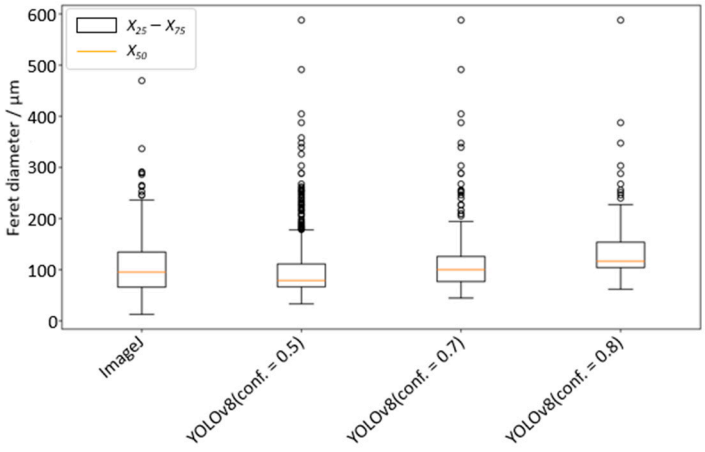


Table S35: YOLOv8 CSD statistics for crystal size range 0-90 µm for ImageJ and different confidence levels

Confidence level	Count	Mean [µm]	Median [µm]	IQR [µm]	Std. Dev [µm]
Image J	301	54.47	47.66	34.32	31.67
0.5	1132	47.52	39.45	22.36	24.43
0.7	402	56.04	50.02	24.52	29.37
0.8	133	69.20	58.26	24.83	33.85

Table S36: YOLOv8 Error measures for crystal size range 0-90 µm and different confidence levels

Confidence level	Mean abs error [µm]	Mean rel error [%]	Median abs error [µm]	Median abs error [%]	Q1 rel. error [%]	Q3 rel. error [%]
Image J	0	0	0	0	0	0
0.5	-6.95	-12.76	-8.21	-17.23	1.00	-17.30
0.7	1.57	2.88	2.35	4.94	16.68	-6.40
0.8	14.73	27.05	10.60	22.23	58.12	14.34

Figure S37: YOLOv8 Boxplot comparing CSD for crystal size range 90-125 µm using ImageJ and YOLOv8 at different confidence levels

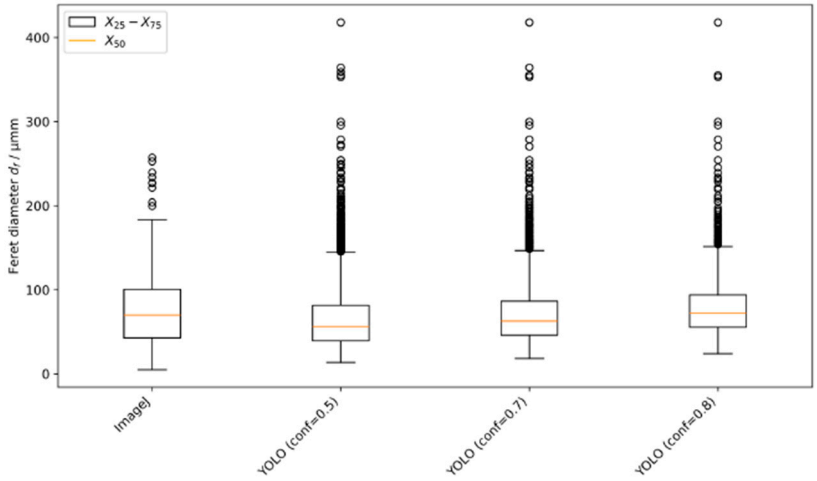


Figure S38: YOLOv8 Boxplot comparing CSD for crystal size range 125-180 µm using ImageJ and YOLOv8 at different confidence levels

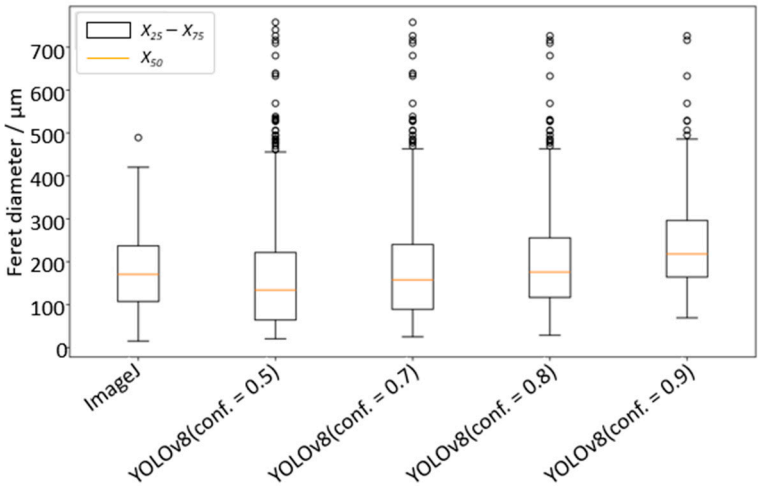


Table S39: YOLOv8 CSD statistics for crystal size range 125-180 for ImageJ and different confidence levels

Confidence level	Count	Mean [µm]	Median [µm]	IQR [µm]	Std. Dev [µm]
Image J	500	179.78	170.62	129.62	97.10
0.5	2110	158.23	134.08	157.15	112.85
0.7	1410	176.81	157.80	151.17	113.02
0.8	966	198.07	175.99	138.76	108.97
0.9	379	241.26	218.10	131.77	105.74

Table S40: YOLOv8 Error measures for crystal size range 125-180 µm and different confidence levels

Confidence level	Mean abs error [µm]	Mean rel error [%]	Median abs error [µm]	Median abs error [%]	Q1 rel. error [%]	Q3 rel. error [%]
Image J	0	0	0	0	0	0
0.5	-21.56	-11.99	-36.54	-21.42	-40.07	-6.55
0.7	-2.97	-1.65	-12.82	-7.51	-17.02	1.38
0.8	18.29	10.17	5.37	3.15	8.79	7.84
0.9	61.48	34.19	47.48	27.83	52.99	24.93

Figure S41: YOLOv8 Boxplot comparing CSD for crystal size range 180-250 µm using ImageJ and YOLOv8 at different confidence levels

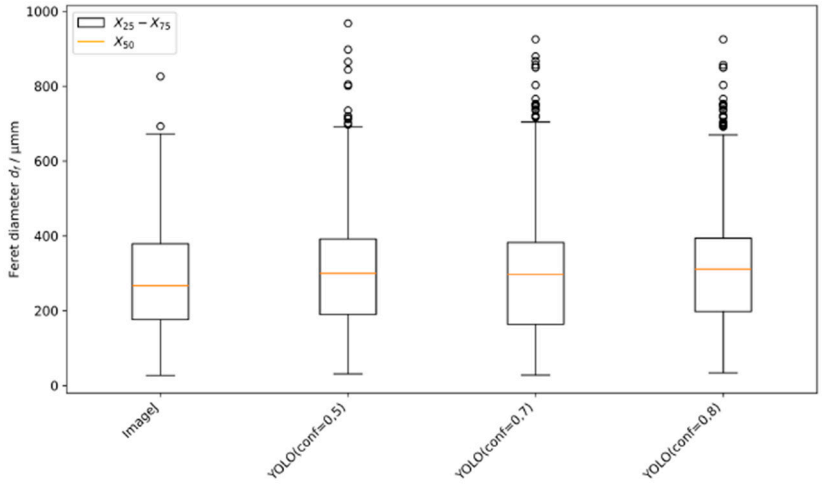


Figure S42: YOLOv8 Boxplot comparing CSD for crystal size range 250-630 µm using ImageJ and YOLOv8 at different confidence levels

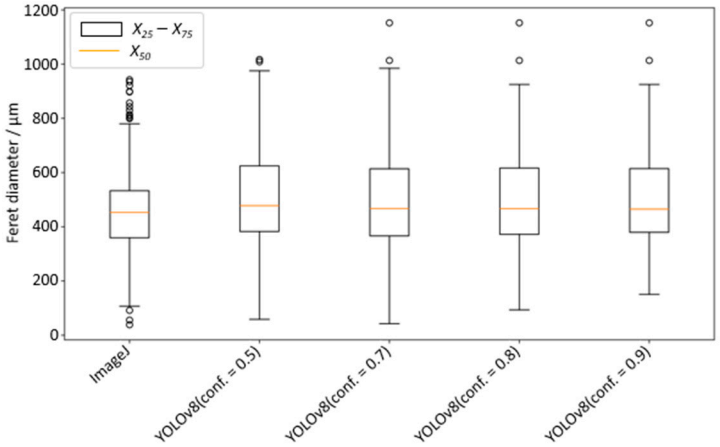


Table S43: YOLOv8 CSD statistics for crystal size range 250-630 for ImageJ and different confidence levels

Confidence level	Count	Mean [µm]	Median [µm]	IQR [µm]	Std. Dev [µm]
Image J	276	456.02	452.84	173.60	161.68
0.5	302	504.27	477.86	242.07	185.51
0.7	293	489.86	467.31	247.57	189.31
0.8	272	494.49	466.82	244.42	182.87
0.9	240	498.01	465.08	234.93	171.53

Table S44: YOLOv8 Error measures for crystal size range 250-630 µm and different confidence levels

Confidence level	Mean abs error [µm]	Mean rel error [%]	Median abs error [µm]	Median abs error [%]	Q1 rel. error [%]	Q3 rel. error [%]
Image J	0	0	0	0	0	0
0.5	48.25	10.58	25.02	5.53	6.48	17.21
0.7	33.84	7.42	14.47	3.20	1.95	15.20
0.8	38.47	8.44	13.98	3.09	3.56	15.69
0.9	41.99	3.21	12.24	2.70	5.64	15.31

Figure S45: YOLO8seg Boxplot comparing CSD for crystal size range 0-90 µm using ImageJ and YOLO8seg at different confidence levels

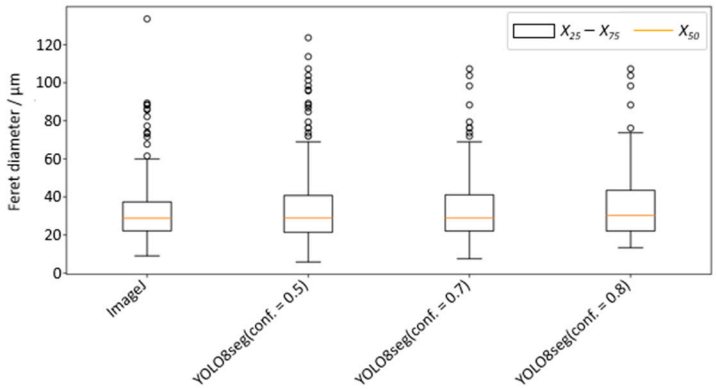


Table S46: YOLO8seg CSD statistics for crystal size range 0-90 µm for ImageJ and different confidence levels

Confidence level	Count	Mean [µm]	Median [µm]	IQR [µm]	Std. Dev [µm]
Image J	300	32.09	28.81	15.22	15.22
0.5	314	33.19	28.92	19.36	19.08
0.7	211	33.14	28.92	19.00	17.07
0.8	137	35.52	30.27	21.37	18.09

Table S47: YOLO8seg Error measures for crystal size range 0-90 µm and different confidence levels

Confidence level	Mean abs error [µm]	Mean rel error [%]	Median abs error [µm]	Median abs error [%]	Q1 rel. error [%]	Q3 rel. error [%]
Image J	0	0	0	0	0	0
0.5	1.10	3.44	0.11	0.38	-3.34	9.12
0.7	1.05	3.28	0.11	0.38	-0.22	9.99
0.8	3.44	10.71	1.47	5.09	-0.22	16.34

Figure S48: YOLO8seg Boxplot comparing CSD for crystal size range 125-180 µm using ImageJ and YOLO8seg at different confidence levels

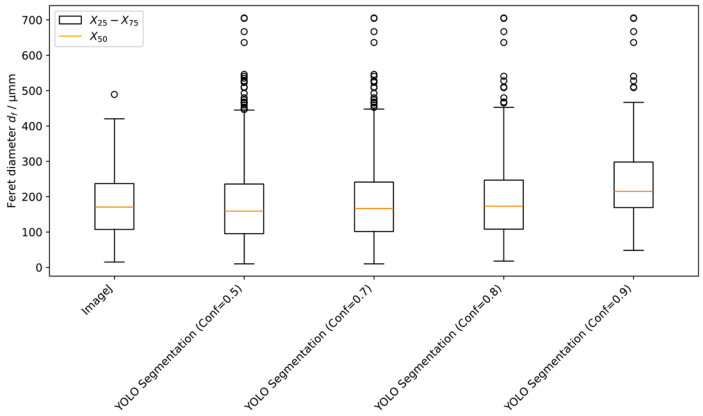


Table S49: YOLO8seg CSD statistics for crystal size range 125-180 for ImageJ and different confidence levels

Confidence level	Count	Mean [µm]	Median [µm]	IQR [µm]	Std. Dev [µm]
Image J	500	179.78	170.62	129.62	97.10
0.5	1717	174.98	159.13	140.35	104.60
0.7	1397	181.92	166.71	139.59	104.82
0.8	1176	187.65	173.21	138.64	102.99
0.9	581	233.55	214.76	128.91	96.15

Table S50: YOLO8seg Error measures for crystal size range 125-180 µm and different confidence levels

Confidence level	Mean abs error [µm]	Mean rel error [%]	Median abs error [µm]	Median abs error [%]	Q1 rel. error [%]	Q3 rel. error [%]
Image J	0	0	0	0	0	0
0.5	-4.80	-2.67	-11.49	-6.73	-11.24	-0.57
0.7	2.13	1.19	-3.91	-2.29	-5.41	1.75
0.8	7.87	4.37	2.59	1.52	0.67	4.11
0.9	53.76	29.90	44.14	25.87	57.36	25.70

Figure S51: YOLO8seg Boxplot comparing CSD for crystal size range 250-630 µm using ImageJ and YOLO8seg at different confidence levels

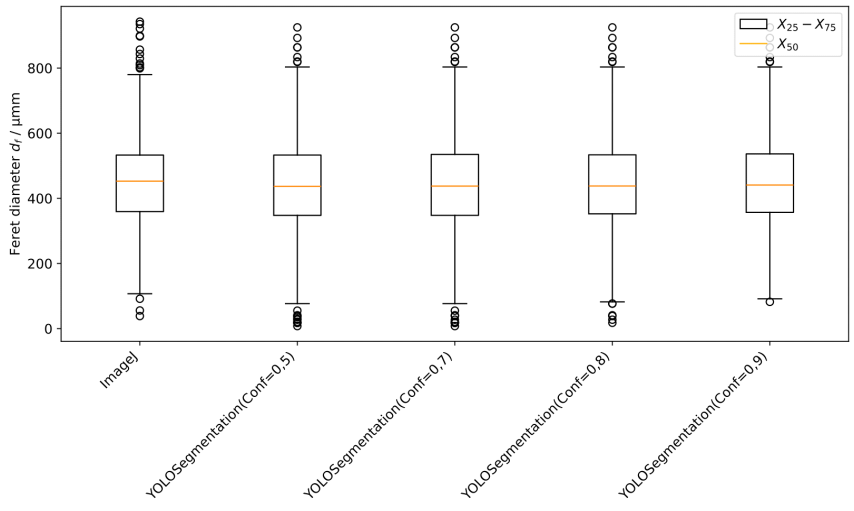


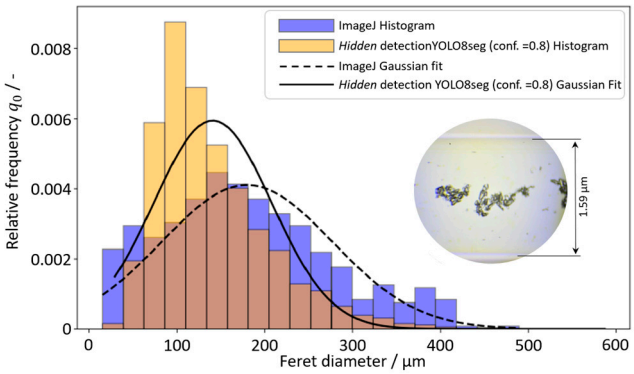
Table S52: YOLO8seg CSD statistics for crystal size range 250-630 for ImageJ and different confidence levels

Confidence level	Count	Mean [µm]	Median [µm]	IQR [µm]	Std. Dev [µm]
Image J	276	456.02	452.84	173.60	161.68
0.5	305	435.18	436.86	185.25	169.98
0.7	298	437.85	437.65	187.18	167.52
0.8	292	442.03	437.92	181.12	161.16
0.9	282	450.97	440.89	179.52	153.07

Table S53: YOLO8seg Error measures for crystal size range 250-630 µm and different confidence levels

Confidence level	Mean abs error [µm]	Mean rel error [%]	Median abs error [µm]	Median abs error [%]	Q1 rel. error [%]	Q3 rel. error [%]
Image J	0	0	0	0	0	0
0.5	-20.84	-4.57	-15.98	-3.53	-3.23	0.01
0.7	-18.18	-3.99	-15.20	-3.36	-3.22	0.38
0.8	-13.99	-3.07	-14.92	-3.30	-1.90	0.13
0.9	-5.05	-1.11	-11.96	-2.64	-0.66	0.66

Figure S54: YOLO8seg Relative frequency q_0 of the detected *Hidden* instances compared to ImageJ. This indicates that most detected *Hidden* instances are of lower size. The orange area represents shared columns



References

1. Wohlgemuth, K.; Schembecker, G. Modeling induced nucleation processes during batch cooling crystallization: A sequential parameter determination procedure. *Computers & Chemical Engineering* **2013**, *52*, 216–229, doi:10.1016/j.compchemeng.2012.12.001.
2. Holscot Advanced Polymers Ltd. Fluorinated Ethylene Propylene: Technical Data for FEP. Available online: https://holscot.com/glossary/fep-is-a-type-of-fluoropolymer-that-shares-similarities-with-other-fluoropolymers-like-pfa/ (accessed on 24 October 2024).
3. Schiebener, P.; Straub, J.; Levelt Sengers, J.M.H.; Gallagher, J.S. Refractive index of water and steam as function of wavelength, temperature and density. *Journal of Physical and Chemical Reference Data* **1990**, *19*, 677–717, doi:10.1063/1.555859.
4. Wang, C.; Wang, Q.; Qian, Y.; Hu, Y.; Xue, Y.; Wang, H. DP-YOLO: Effective Improvement Based on YOLO Detector. *Applied Sciences* **2023**, *13*, 11676, doi:10.3390/app132111676.

5. Jocher, G. Ultralytics YOLOv8. Available online: https://docs.ultralytics.com/models/yolov8/ (accessed on 20 October 2024).
6. Everingham, M.; van Gool, L.; Williams, C.K.I.; Winn, J.; Zisserman, A. The Pascal Visual Object Classes (VOC) Challenge. *Int J Comput Vis* **2010**, *88*, 303–338, doi:10.1007/s11263-009-0275-4.
7. TensorFlow. tf.keras.optimizers.Adam: TensorFlow v2.16.1. Available online: https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/Adam (accessed on 24 October 2024).