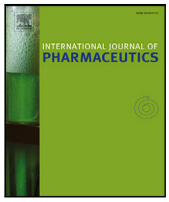
International Journal of Pharmaceutics 664 (2024) 124629

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm



Automated particle inspection of continuously freeze-dried products using computer vision

Quentin Herve [a], Nusret Ipek [b], Jan Verwaeren [b], Thomas De Beer [a,*]

[a] Laboratory of Pharmaceutical Process Analytical Technology, Department of Pharmaceutical Analysis, Ghent University, 9000 Gent, Belgium
[b] Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure Links 653 B-9000 Gent, Belgium

ARTICLE INFO

Keywords:
Freeze-drying
Continuous manufacturing
Quality analysis
Visual inspection
Deep learning
Computer vision

ABSTRACT

The pharmaceutical industry is progressing towards more continuous manufacturing techniques. To dry biopharmaceuticals, continuous freeze drying has several advantages on manufacturing and process analytical control compared to batch freeze-drying, including better visual inspection potential. Visual inspection of every freeze-dried product is a key quality assessment after the lyophilization process to ensure that freeze-dried products are free from foreign particles and defects. This quality assessment is labor-intensive for operators who need to assess thousands of samples for an extensive amount of time leading to certain drawbacks. Applying Artificial Intelligence, specifically computer vision, on high-resolution images from every freeze-dried product can quantitatively and qualitatively outperform human visual inspection. For this study, continuously freeze-dried samples were prepared based on a real-world pharmaceutical product using manually induced particles of different sizes and subsequently imaged using a tailor-made setup to develop an image dataset (with particle sizes from 50 μm to 1 mm) used to train multiple object detection models. You Only Look Once version 7 (YOLOv7) outperforms human inspection by a large margin, obtaining particle detection precision of up to 88.9% while controlling the recall at 81.2%, thus detecting most of the object present in the images, with an inference time of less than 1 s per vial.

1. Introduction

Freeze drying is a manufacturing process widely used in the pharmaceutical industry and especially in the biopharmaceutical sector. It addresses storage and distribution challenges, making it a leading edge in drug manufacturing by offering new opportunities and enhanced product stability. Lyophilization involves a well-defined sequence of three critical stages. At first, the aqueous solution is frozen on shelves to a temperature between −40 and −60 °C inside a vacuum chamber but at atmospheric pressure. The primary drying phase follows, which is initiated by reducing the chamber's pressure to a range of 50 to 200 μbar while gradually increasing the shelf temperature. This phase is essential for the sublimation of the ice crystals formed during freezing. Lastly, the process enters the secondary drying phase where the temperature of the product is raised further by increasing the shelf temperature, allowing the removal of remaining unfrozen water by desorption.

Following the incentive to move from a batch manufacturing approach to continuous manufacturing (Committee for Medicinal Products for Human Use, 2021), an alternative method for freeze drying was proposed by Corver (2013) namely continuous spin freeze-drying. The continuous freeze-drying technology starts with a continuous freezing step where the vials, filled with liquid product, are rapidly rotated along their longitudinal axis. The cooling and freezing is achieved by using a flow of sterile gas with a controlled temperature around the rotating vial. Consequently, the resulting frozen product is spread with a uniform thickness over the entire vial surface as shown in Fig. 1 (Van Bockstal, 2018). During the primary drying phase, sublimation removes the majority of the water content from the frozen product. Hereafter, low residual levels of bound water still present in the product will be removed by desorption in the secondary drying phase. This resulting freeze-dried product having a thin dried cake at the vial side wall allows greater surface inspection when compared to batch manufacturing where the cake is located at the bottom of the vial. With this approach, the inspection area increases by 30% when comparing area calculation on the part of the cake which is visible. Moreover, the layer thickness is reduced compared to the batch method, considering the lower filling volume used, which further enhances detection capabilities given appropriate lightning.

Ensuring the absence of visible particles stands as an essential requirement for parenteral products which demands strict control.

* Corresponding author.
E-mail addresses: Quentin.Herve@UGent.be (Q. Herve), Thomas.DeBeer@ugent.be (T. De Beer).

https://doi.org/10.1016/j.ijpharm.2024.124629
Received 23 June 2024; Received in revised form 14 August 2024; Accepted 21 August 2024
Available online 23 August 2024

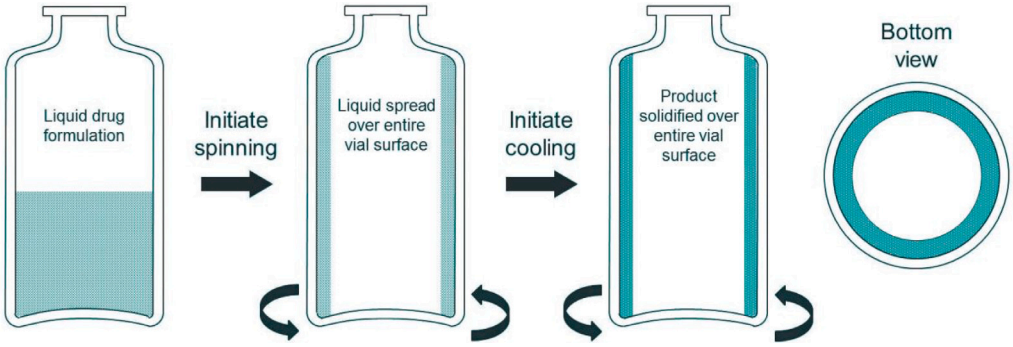


Fig. 1. Representation of the continuous spin freeze drying process.

Given the specific character of liquid injection, foreign substances can cause severe trauma if they are introduced into the patient's body (Langille, 2013). Particles can appear from different sources and origins. Extrinsic particles will originate unrelated to the manufacturing process while intrinsic ones will be related to it. As per the requirements defined in USP 790 (Pharmacopeia, 2021), all vials must undergo 100% visual inspection and ensure a particle-free product. The prevailing inspection technique continues to rely on manual visual inspection without magnification where trained operators conduct manual inspections of containers against contrasting white and black backgrounds. Particle detection being a probabilistic process (Aldrich et al., 2016), this method is increasingly subject to human fatigue due to the labor-intensive work leading to errors but also affected by naked eye limitations.

As the pharmaceutical industry is constantly struggling with rising complexity, cost and regulation, more of them are looking to Industry 4.0 manufacturing, considered as a new industrial revolution (Arden et al., 2021). Smart autonomous production lines that are managed using real-time obtained data and machine learning will lower pharmaceutical manufacturing costs, improve quality and reduce capacity constraints. To improve the efficiency, reproducibility and quality of human visual inspection, automated solutions are being considered in various fields of industry including the pharmaceutical one (Pokhriyal et al., 2023; Veillon and Shabushnig, 2020; Snyder, 1995). Automated visual inspection represents the use of computers and machine vision systems to automatically inspect and analyze the appearance of objects or surfaces. It is often used in manufacturing and quality control processes to ensure that products meet certain standards and specifications. Machine vision systems typically consist of a camera or other image capture device, a computer or embedded system to process the data, and algorithms to analyze and interpret the images. The system may also include additional hardware, such as lighting and specialized optics, to optimize the image capture and analysis process. Machine vision algorithms can be implemented using traditional computer vision techniques or using machine learning approaches, such as deep learning. Lately, the pharmaceutical industry has also started to explore the potential of detection algorithms for particles to achieve better process understanding (Péterfi et al., 2023; Ficzere et al., 2023).

In this study, an innovative method for visible particle detection in continuously lyophilized products is proposed. Particles from different origins and with different sizes starting from 50 μm to 150 μm were used to demonstrate the feasibility of the application. The size range chosen encompasses the lower limit for visible particles with the naked eye and the higher size which has according to the USP 1790 a probability of detection (PoD) of 70% with the naked eye. However, pharmaceutical manufacturing applications should value precise and safe procedures. While artificial intelligence (AI) strives to be a force of transformation in other domains, the application of computer vision remains limited in the realm of pharmaceutical quality control. Hence, in this study, the possibility of AI integration with a primary focus on computer vision-based object detection was explored to address the shortcomings of the current procedure of human-based inspection. On this basis, three different model architectures were proposed and tested to accurately and efficiently identify foreign objects in the vials; Faster RCNN, Cascade RCNN, and YOLOv7. Ren et al. (2015), Cai and Vasconcelos (2018) and Wang et al. (2023). Given the current standards in place and the potential of computer vision, a novel approach with high mean average precision (mAP) of foreign particle detection while keeping the lowest false reject rate possible was proposed.

This paper is divided into multiple parts where first the hardware used for the image acquisition, the approach used for the sample preparation and the pre-processing strategy used will be presented in the material and methods section. Subsequently, different computer vision model types will be showcased and compared to reach a conclusion defining the most promising one for the application. Finally, the best model candidate is compared with human visual inspection, which remains the benchmark method currently used in the field.

2. Material and methods

2.1. Continuous freeze drying samples preparation

Typical continuously freeze-dried samples rarely contain any kind of particles, especially in the quantities needed to build a deep learning-based computer vision model. Hence it was necessary to create sets of vials containing continuously freeze-dried product in which particles were artificially induced. To assemble a comprehensive dataset for particle detection analysis, five distinct particle categories were identified: two categories of intrinsic particles representative for those typically encountered in a manufacturing process and three extrinsic categories. The first intrinsic category consists of fibers. During preliminary experiments and in past work, fibers were detected in the vials and got trapped in the cake. It was assumed that the product used being present in large quantity was contaminated within our non-classified R&D laboratory over years of use resulting in the introduction of some impurities. As a result, fibers were considered a realistic class of intrinsic particles. Hence no particular measure were taken to tackle their presence in the end product. Secondly, particles originating from vial stoppers were considered another class of intrinsic particles. Particles originating from vial stoppers were prepared using a shredder to create a second example of intrinsic foreign matter for the analysis. Additional intrinsic particles, such as glass fibers, could have been considered but were left out due to the great difficulty of production. Furthermore, with the product tension more spread on the vial's wall, glass breakage is not considered a risk in Continuous Freeze Drying. The remaining three categories of particles were meticulously calibrated to possess precisely defined sizes, serving as informative benchmarks for evaluating size detection capabilities and enhancing the understanding of the detection performance. 50 μm black-dyed uniform polystyrene (Thermofisher scientific), 100 μm dark blue microparticles based on polystyrene (Sigman Aldrich) and 150 μm black-dyed uniform

Table 1
List of particles, their size ranges and origin.

Particle type	Particle size	Origin
Fibers	100 μm–1 mm	Intrinsic
Gray particles	100 μm–1 mm	Intrinsic
Black-dyed polystyrene beads	50 μm	Extrinsic
Dark blue polystyrene beads	100 μm	Extrinsic
Black-dyed polystyrene beads	150 μm	Extrinsic

polystyrene (Thermofisher scientific) were acquired to represent an increasing particle size library and satisfy the protocol envisioned for the detection capabilities.

An overview of particles and their origin are listed in Table 1. Formulations with 7.5% (w/v) of each Mannitol, Sucrose and Trehalose were prepared with all different particles. Some samples were made with individual categories but also mixed categories with different classes of particles present leading to an extensive number of instances presented in the first part of the results section in Table 2. Diverse particle concentrations were employed to adjust the quantity of particles in the produced samples. 10R vials (Schott, Müllheim, Germany), filled with 3 mL of the prepared solutions were spun using a WB600-D high-speed overhead stirrer (Wiggens, Beijing, China) with a 3000 RPM rotation. During the rotation, liquid nitrogen with a flow rate of 100 L/min was used to freeze the product. Once the outer part of the vials measured with a FLIR camera reached a temperature of −60 °C, they were all transferred to cylindrical aluminium holders in a −40 °C precooled chamber of a Martin Christ Freeze Dryer. All samples were prepared using specific freeze-drying process conditions to generate variability within the cosmetic output generated. After completion of the secondary drying step, samples were stoppered and cleaned for image acquisition.

2.2. Hardware selection

This study was encouraged by the improvement over human visual inspection capabilities for particle analysis. Research showed that detection probability by a qualified operator was only 4% for 50 μm particulate matter and increased to 90% for 200 μm particulate matter (Aldrich et al., 2016; Poms et al., 2019). More recent research even showed that these probabilities are lower in a typical production environment where particle morphological and optical properties have a significant impact (Mazaheri et al., 2023) on their detection rate. As part of an objective to improve these capabilities, the investigation aimed to explore smaller particle sizes to leverage the advancements in imaging technology. Detecting a particle typically requires it to span multiple pixels for distinction from background noise. The minimal number of pixels depends on factors like particle-to-background contrast and employed image processing techniques. In this particular case, to account for challenging conditions in the environment, a minimum of 10 pixels was conservatively selected for reliable detection. Leveraging a 16000-pixel resolution line scan camera with a 3.52 μm sensor size, the acquisition equipment should enable the detection of particles as small as 30 μm. Images were acquired using a DALSA LINEA LA-CM-16K05 A (Teledyne Technologies, Inc., Thousand Oaks, California, USA) with additional lens extension and SCHNEIDER COMPONON S 58–76 80/4,0 lens. CCS TH2-43/35-SW (STEMMER IMAGING AG, Switzerland) white-led lightning was provided below the object of interest. The camera used being linear, one Geared DC motor was used to allow vial rotation during the acquisition leading to a fully flattened representation of the vial's body. The settings of the camera are presented in the Supplementary Material section.

2.3. Data acquisition

All the images were acquired using a fixed setup presented in Fig. 2 to ensure high reproducibility between the different acquisitions. This setup was built using Paletti aluminium construction parts and an adjustable slide unit for the camera to assure precise measurements without variability for all the vials to be captured. For the 10R vials considered in this study, images were generated with a 23 600 × 10 900 pixels resolution. After conducting several trials to evaluate how much pixel were required to obtain a full vial body, it was determined that a complete vial representation only required a height of 10 900 pixels. Furthermore, by combining 23 600 lines, each consisting of 10 900 pixels in height, a full vial image could be obtained.

2.4. Data pre-processing

Images captured during the acquisition were initially saved in a large format exceeding the capacity limitations of the virtual memory allocated to the graphical processing unit (GPU). Consequently, adjustments were required to resize the images to meet the constraints imposed by the hardware resources. The collected data comprising more than 200 vials was divided randomly while keeping a comparable amount of each formulation and particle concentration into two sets: training and validation, with a distribution ratio of 80% and 20% respectively. The validation split was used during the training of the different models to fine-tune the different models considered and generate metrics for the model evaluation. In a later stage, a test (reference) data set was created and employed to assess the model performance on previously unseen data in order to study the generalizability of the model and to benchmark against human operators. To preserve the wealth of information in the high-resolution images and tackle the virtual memory limitations, it was decided not to lower the resolution but instead, to divide each image into 100 equally-sized patches, consisting of 2360 × 1090 pixels each. Following this slicing operation, instances of each class of particle within these patches were meticulously annotated by drawing bounding boxes around each particle using the COCO annotator (Brooks, 2019), a web-based tool dedicated to computer vision annotation. Each bounding box was given a label according to the type of particle. This annotation process allowed the export of the data in a standardized format widely used in various computer vision models, ensuring compatibility and ease of integration. This pre-processing task was applied to each split of data in order to obtain information on the ground truth regarding the localization and classification of the particles.

2.5. Data modeling

In this work, the problem of detecting intrinsic and extrinsic particles in vials is solved as a multi-class object detection problem, where object detectors are trained to identify and localize particles by constructing bounding boxes around them. Hence, the modeling involves automatic identification and localization of fibers, gray cap particles and beads of variable sizes (see Table 1). In this study, three object detection model architectures were considered and benchmarked; Faster R-CNN, Cascade R-CNN, and YOLOv7. These architectures were selected as they are well-studied and known to perform well in practice. Moreover, structurally, they differ in several aspects. Both Faster R-CNN and Cascade R-CNN are based on region proposal networks (RPN) and are known to have good detection capabilities for small-sized objects, so these model architectures were suitable for this application (Zhang et al., 2023). On top of that, YOLOv7 brought further accuracy improvements in the popular benchmarks (such as the well known COCO dataset) while sustaining a low inference detection time (Wang et al., 2023) which is an important aspect when considering its implementation for automated inspection on a continuous freeze-drying pipeline at an industrial scale.

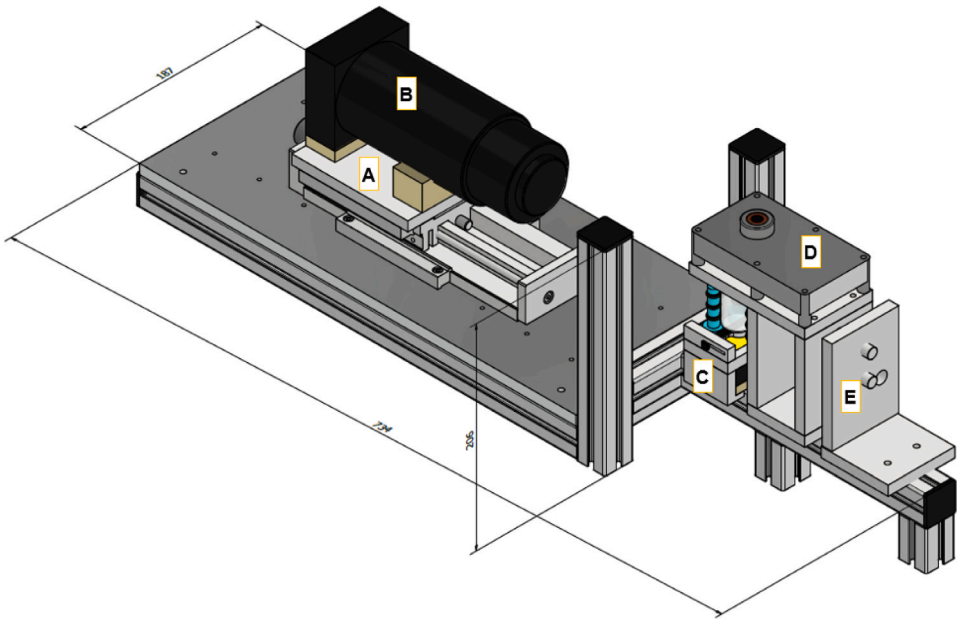


Fig. 2. Schematic representation of the acquisition setup with sliding united (A), line scan camera (B), led lightning (C), geared DC motor (D), spring system to keep the vial in a fixed position on the horizontal axis (E).

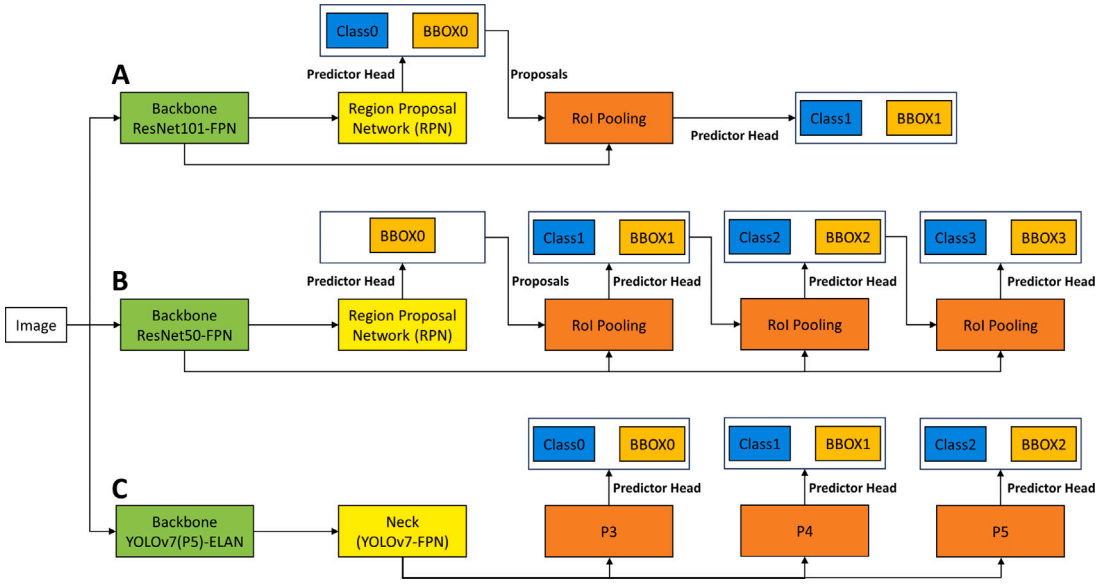


Fig. 3. The drawing shows the main components of three model architectures used in this study; Faster R-CNN (A), Cascade R-CNN (B), and YOLOv7 (C). BBOX stands for bounding box regressor, Class stands for the classifier. The fully connected network heads are denoted as prediction heads.

Region-based Convolutional Neural Network (Faster R-CNN) is an established object detection model and has demonstrated accurate localization capability (Ren et al., 2015). It follows a two-stage approach where the first stage is the region proposal network (1) to generate potential object locations in the image by dividing the image into small regions. Then, it evaluates whether the proposed regions contain an object of interest for further analysis. In the second step, a region-based CNN (2) is used to predict the object class and refine the boundaries for all the proposed regions from step 1. Following a convolutional backbone (ResNet101) for feature extraction, a feature pyramid network (FPN) brings in the ability to identify objects at different scales. FPN creates a pyramid of feature maps at various scales which are derived from the layers of the ResNet101 backbone. This allows a scalable model to capture both coarse and fine-grained object properties for more accurate detection capabilities. Hence, the Faster R-CNN model with ResNet101 backbone and FPN is suitable for robust object detection tasks across various image scales where precision and scalability are critical components.

As the second model, the Cascade R-CNN model as an advanced variant of the Faster R-CNN model is used with the capability of mask prediction (Cai and Vasconcelos, 2018). This model architecture not only extends the accuracy of the object detection/classification but also allows pixel-wise segmentation of the object mask. Comparably to the Faster R-CNN model, it follows a two-stage approach where the RPN is the same whereas the second step implements a cascade head instead of region-based CNN as detailed in Fig. 3. This cascade of stages refines the object detection capabilities progressively. In each cascade stage, a classifier and regressor are present, whereas the later stages in the cascade focus on the false positives of the preceding stages of the model. Therefore, in each stage, the model aims to improve the

precision of the object localization and filter the background regions. In this case, the backbone (ResNet50) has fewer layers in order to balance the computational efficiency and the detection capabilities of the model. Additionally, the FPN is also present as described above for scalable detection of the objects. Overall the cascaded structure allows the model to progressively refine predictions, addressing the challenge of scale variation by selectively filtering proposals in each stage, ultimately leading to overall precision and robustness improvements in object detection.

Lastly, YOLOv7 is a recent, single-stage and real-time object detector and is one of the most accurate real-time object detectors to date as shown in recent studies (Wang et al., 2023; Sun et al., 2024; Morita et al., 2023). YOLOv7 is based on the efficient and accurate network architecture called the Efficient Layer Aggregation Network (ELAN). This network structure is derived from the gradient path analysis (Wang et al., 2022). Moreover, compound model scaling and anchor-free detection provide efficiency without sacrificing loss in detection capability. The YOLOv7 model consists of three main components: a backbone (for feature extraction), a neck (fusion of features from different layers, FPN), and a head (prediction of bounding boxes and classification). Additionally, YOLOv7 is capable of performing pixel-wise segmentation, similar to the Cascade R-CNN model presented beforehand. Fig. 3 shows the YOLOv7 architecture and components in detail. Interestingly, the YOLOv7 model achieves good detection capabilities while being computationally more efficient than the region-based CNN models. Due to the continuous production approach of freeze-dried vials, the particle inspection procedure must adhere to the speed of the pipeline. In this case, computational efficiency is important as the image acquisition and pre-processing already take a considerable amount of time. Therefore, YOLOv7 as a real-time object detector has an advantage over the other models proposed which are known to be less computationally efficient during inference.

The region-based CNN models were initialized using pre-trained weights obtained from pretraining on ImageNet, and the YOLOv7 was pre-trained the MS COCO dataset (Deng et al., 2009; Lin et al., 2014). In the implementation of the models, both Faster R-CNN and Cascade R-CNN models utilized the Detectron 2 framework with PyTorch backend while YOLOv7 was trained with the implementation attached to the seminal article (Wu et al., 2019; Wang et al., 2023) (see supplementary information for more details on hardware/software). Generally, using, pre-trained weights are used in order to transfer the knowledge obtained in related tasks (in this case, object detection tasks on more general but very large image databases) to the task at hand. It is commonly assumed that learning can benefit from this transfer, resulting in a reduced need for annotated data, improved training efficiency, and better generalization capabilities.

2.5.1. Hyperparameters & Data augmentation

During the training of Faster RCNN and Cascade RCNN models using the Detectron 2 framework, 16 images per minibatch were found to be optimal considering the GPU constraints. Both models were trained with 1e−3 learning rate and linear warm-up with a cosine annealing scheduler. Subsequently, they were both trained for 6000 iterations. Specifically for the Cascade RCNN heads, cascade stage Intersection over Union threshold values were defined as 0.35, 0.5 and 0.65.

For the YOLOv7 model, the batch size chosen was 20 images considering the GPU constraints. The default learning rate (1e−2), momentum (9.37e−1) and warm-up epochs (3.0) were kept. In total, the YOLOv7 model trained 400 epochs. The following image augmentation techniques were applied for Faster R-CNN and Cascade R-CNN during training phase; random brightness adjustment, random contrast adjustment and random flipping (both horizontal and vertical). During YOLOv7 training; random HSV adjustment, image translation, random scaling, random flip (both horizontal and vertical), mosaic, mix-up and paste-in were applied. All the code is provided in the GitHub repository.[1] For all models, the best model weights on the validation set were used for the reference evaluation.

2.6. Model evaluation

This section describes how the obtained models were evaluated. The performance metrics: precision, recall, F1 scores, mean average precision, and Intersection over Union (IoU) are defined in Eqs. (1)–(6).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}} \tag{3}$$

$$\text{IoU}(\text{bbox}_i, \text{bbox}_j) = \frac{\text{Area of Overlap}\,(\text{bbox}_i, \text{bbox}_j)}{\text{Area of Union}\,(\text{bbox}_i, \text{bbox}_j)} \tag{4}$$

$$\text{mAP} = \frac{\sum_{n=1}^{K} \text{AP}_n}{\text{K}} \tag{5}$$

$$\text{mAR} = \frac{\sum_{n=1}^{K} \text{AR}_n}{\text{K}} \tag{6}$$

Firstly, precision (Eq. (1)) is the ratio of true positive predictions to the sum of true positives and false positives and thus measures the fraction of the detected regions in the image that contain a particle. Recall (Eq. (2)) is the ratio of the true positive predictions to the sum of true positives and false negatives and thus measures the fraction of particles that were detected by the model. F1 scores (3) represent the balance between precision and recall combined into a single value and allow the assessment of the overall performance of a model. The intersection over union (IoU) given by Eq. (4) computes the overlap between the predicted bounding box (bbox_i) and the ground truth (bbox_j), providing information on the ability of a model to accurately localize particles. Mean average precision (mAP) is given in Eq. (5) and quantifies the exactness of the detection among all classes K. It is calculated by averaging the average precision (AP) values over all classes indexed by n present in the dataset. In accordance with mAP, the mean average recall (mAR) demonstrated in Eq. (6) is calculated by considering all the average recall values per class K. For detailed information on average precision (AP), average recall (AR) calculations in this context and all other formulas presented above, refer to the (Padilla et al., 2021) paper.

2.7. Reference analytics with visual assessment

In the final assessment, the models developed were subjected to a comparison against human visual inspection. In this phase, a group of five individual operators was responsible for conducting quality control by meticulously examining an independent reference test set created for this specific task. With respect to regulatory requirements governing inspection processes, an inspection protocol was defined and followed by the inspectors. To train the operators for their inspection task, they were given several examples of each particle category. Diverse background conditions and varying levels of illumination were employed during the evaluation. All vials were then inspected without any magnification devices for approximately 45 to 60 s to replicate the standard human visual inspection method. This reference set was thoughtfully composed of 30 vials, which each operator will inspect including vials without any particles to ensure that no bias would be induced during the inspection.

The same formulations chosen to create the training and validation datasets were used here to obtain a similar visual output. Each of these mannitol, sucrose and trehalose based formulations were used for one-third of the 30 prepared vials. Several samples were generated, each

[1] https://github.com/Quentinh29/particleDetection/.

Table 2
Composition of the different datasets with their corresponding number of annotations per class.

Particle type	Training annotations	Validation annotations	Test annotations
50 μm	6662	3857	478
100 μm	2593	1362	43
150 μm	1726	494	32
Fibers	2491	1235	549
Gray particles	1059	582	78

vial containing multiple examples from each particle. Finally, samples were prepared using relatively low numbers of particles in order to create a more challenging detection. Ultimately, the average detection numbers for each particle class and overall detection were calculated, serving as a benchmark for model comparison. In some cases, when the number of particles was too large to be counted reliably, it was decided to give an attribute "High" to avoid any miscounting and faulty comparison later on. Expert annotations were also counted and reported as the ground truth through the use of Coco Annotator (Brooks, 2019) following the same approach used for the training and validation sets.

3. Results and discussion

3.1. Results

Particle Dataset Generation

For each dataset generated and annotated, numbers presented in Table 2 were obtained. The number of instances per category was clearly imbalanced, but the initial goal of reaching a minimum of 1000 examples per category for the training set was reached. During the annotation process, it could already be noted that some misclassification between the different sizes of beads could be expected. Depending on the depth within the lyophilized cake, the contrast differences sometimes led to more difficult and less accurate annotation process. In other words, it was difficult to precisely set the bounding box around the object in consideration. Additionally, it was considered challenging to differentiate 50 μm with 100 μm and 100 μm with 150 μm when both examples were not displayed on the same image for direct comparison by the annotator. Ultimately, size determination being out of scope, this will not affect the expectations of this study. All the data generated is made available on the Zenodo[2] platform. In Table 2, it can also be pointed out that the number of test annotations (reference data) is significantly lower compared to the validation set. This arose from the need to use human inspectors as benchmarks for the method validation. For a realistic model evaluation, it was decided to define this reference set with as few particles as possible per vial.

Some examples of the different particles in scope for this detection project are presented in Fig. 4, each image represents 1/100 of an entire vial. It is also worth mentioning that, unlike the calibrated size of bead particles which were perfect spheres, the gray particles from the stopper, generated with the shredder, exhibited entirely different shapes, sizes and characteristics.

Model Metrics Evaluation

Following the completion of both training and validation steps for all three models, standard metrics (Padilla et al., 2021) are presented in Table 3 for comparison.

The results show that YOLOv7 consistently outperforms Faster R-CNN and Cascade R-CNN in terms of mean Average Precision (mAP) at IoU threshold ranging from 0.5 to 0.95. Results also indicate superior performance at lower IoU threshold with validation and a test score of 0.806 and 0.696 respectively surpassing Faster R-CNN (0.771 and 0.682) and Cascade R-CNN (0.762 and 0.677). For specific categories such as polystyrene beads and fibers, YOLOv7 demonstrate competitive performance across almost all size ranges for both datasets. Finally, in regards to gray particles, YOLOv7 yielded better numbers at both validation (0.3222) and test (0.247). In summary, this last model exhibits superior overall performance for the task in consideration by demonstrating consistently higher mAP scores for the majority of the categories. This confirmed that YOLOv7 is a robust candidate for this detection study.

To further investigate the detection capabilities of particles with this model, mean Average Precision per category at lower IoU thresholds 0.65 and 0.5 were calculated and presented in Table 4. Considering that the presence of a single particle would be a reject criterion for any type of product, the exact localization of a bounding box was considered less important than the detection itself. Hence, these lower Intersection over Union thresholds would provide more information on the detection capabilities of the model.

As demonstrated in Table 4, considering lower Threshold values significantly increased the mean average precision values for all individual categories. 0.65 threshold gave clear improvement with a mean average precision score of 0.67 for the 50 μm against 0.425 when considering the varying IoU between 0.5–0.95. This was also observed for the gray particles and the 100 μm particles which presented the lowest precision values on this reference data set. Further lowering the threshold to 0.5 resulted in clear improvement for the precision. Therefore, based on these results, it was decided to opt for an IoU threshold of 0.5 for the prediction on reference data, thus allowing some error margin on the localization. In Fig. 5, a full raw image is provided and additionally in Fig. 6, six individual patches with different categories predicted are reported to exhibit the efficacy of detection with YOLOv7 on these images.

To further evaluate the model, a confusion matrix is provided in Fig. 7. A confusion matrix is a representation widely used in machine learning and statistical classification to assess the performance of a classification model. The columns represent the true classes of the objects and the rows represent the predicted classes. The confusion matrix in object detection is built differently from the ones in the simpler classification tasks, here the sum of the percentage in one row can exceed 100%, which reflects that some instance of particle may have been detected by multiple bounding boxes. This event can occur from object overlap or ambiguity within the images, leading to particles being counted multiple times. Additionally, this matrix presents one additional row and column which correspond to the background false negative (FN) and background false positive (FP). While the first one indicates the occurrence of a particle of a particular category that the model failed to detect, the other represents the detection of an object while none is present.

According to the results presented in this confusion matrix, it was possible to determine where misclassification happens in the predicted reference set for each particle category. This confusion matrix was built using an IoU threshold of 0.5 and a confidence threshold of 0.001. First, for the 50 μm particles class, 83.1% of the predictions were true positives, meaning they were correctly identified as 50 μm particles. However, 27.6% of the predictions were false positives, where an object was detected as a 50 μm particle but no such particle was actually present. Additionally, the model failed to detect 11.5% of the 50 μm particles. Secondly, regarding the 100 μm particles, 62.5% of the predictions were true positives; nevertheless, 12% of the predictions were misclassified as 150 μm. A false positive rate of 5.7% was observed for this class, and the model failed to detect 10.4% of this particle size. For 150 μm particles, the model achieved a true positive rate of 80%, but 20.8% of these particles were misclassified as 100 μm particles. The missed detection rate for 150 μm particles was 4%. Subsequently, the fiber class exhibited a true positive rate of 90.6% but revealed a high false positive rate of 50.5% and 9% of detection failure by the model. Ultimately, gray particles presented 78.3% of true positive predictions

[2] https://zenodo.org/records/10640247.

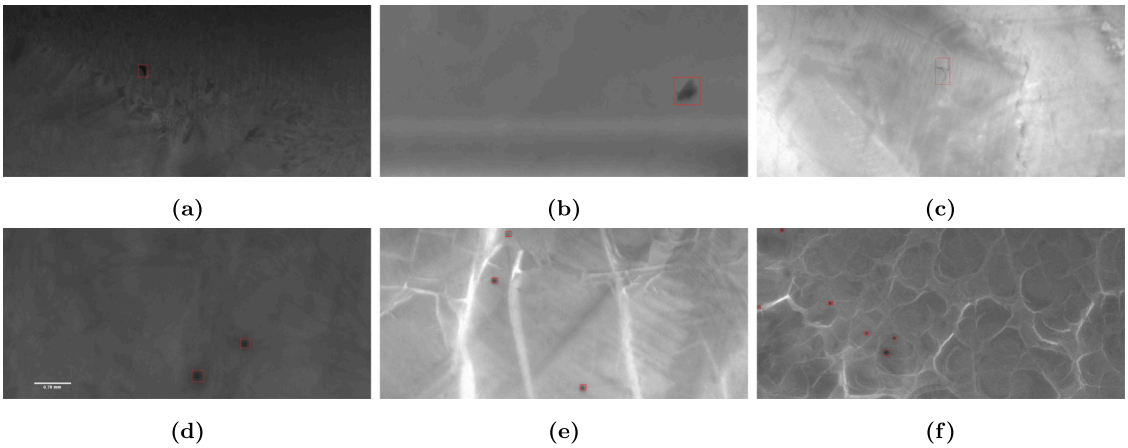


Fig. 4. Examples of different particles, gray particle from stopper (a), gray particle from stopper (b), fiber (c), 150 μm dark polystyrene beads (d), 100 μm blue polystyrene beads (e), 150 μm and 50 μm dark polystyrene beads (f).

Table 3
Validation and test results: mAP and per-category AP.

Category	Faster R-CNN		Cascade R-CNN		YOLOv7	
	Validation	Test	Validation	Test	Validation	Test
Mean average recall (IoU @0.5:0.95)	0.555	0.518	0.570	0.529	0.576	0.544
Mean average precision (IoU @0.5:0.95)	0.435	0.41	0.448	0.424	0.465	0.43
Mean average precision (IoU @0.5)	0.771	0.657	0.762	0.677	0.806	0.696
50 μm precision (IoU @0.5:0.95)	0.389	0.386	0.408	0.405	0.402	0.425
100 μm precision (IoU @0.5:0.95)	0.501	0.361	0.514	0.379	0.55	0.38
150 μm precision (IoU @0.5:0.95)	0.487	0.555	0.518	0.573	0.528	0.538
Fibers precision (IoU @0.5:0.95)	0.52	0.547	0.527	0.566	0.525	0.561
Gray precision (IoU @0.5:0.95)	0.28	0.202	0.271	0.196	0.322	0.247

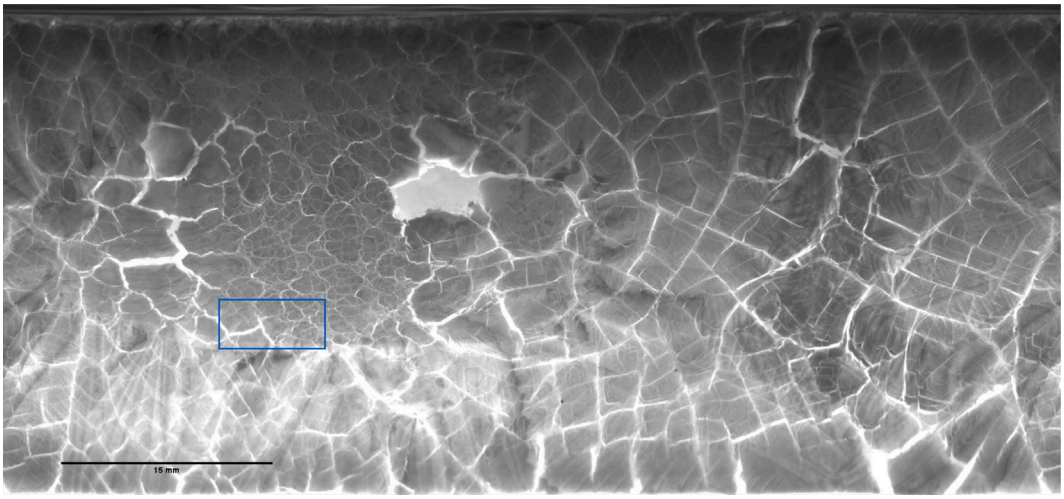


Fig. 5. Full representation using 23600 × 10900 pixels of one 7.5% (w/v) Trehalose vial with a rectangle representing the size of one individual patch.

Table 4
mean Average Precision (mAP) per category compared in function of varying IoU threshold values.

Category	IoU values		
	0.5–0.95	0.5	0.65
50 μm	0.425	0.842	0.67
100 μm	0.38	0.606	0.543
150 μm	0.538	0.825	0.81
Fibers	0.561	0.785	0.74
Gray particles	0.247	0.42	0.381

with 12.7% of false positive instances. Lastly, the model failed to detect 6.7% of the particles present in the dataset.

The F1 score is a commonly used evaluation metric for the classification of objects. It is a single metric that combines precision and recall to measure a model's performance. Therefore, it represents the harmonic mean of precision and recall, providing a balanced assessment of a model's ability to make accurate predictions while minimizing false positives and false negatives. In Fig. 8, the F1 curve visualizes how precision and recall trade off against each other across different confidence thresholds ranging from 0.001 to 1 considering all object classes. Following the assessment of the overall F1 curve, a new confidence threshold corresponding to the optimum found on the curve was chosen for the rest of the study. This higher confidence threshold will create

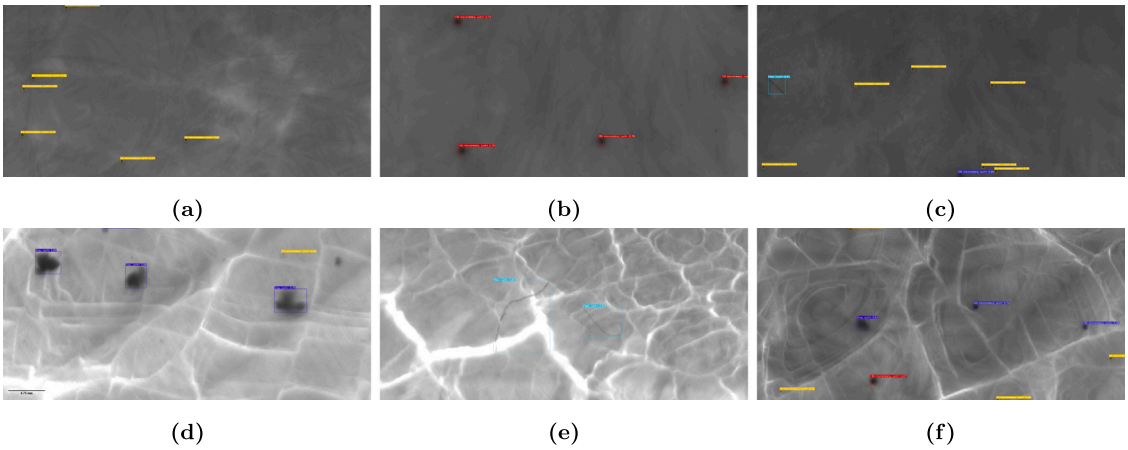


Fig. 6. Sample of images and prediction by the YOLOv7 model with their corresponding confidence value, 50 μm particles (a), 150 μm particles (b), 50,100 μm particles and fibers (c), gray particles (d), fibers (e), 50,100,150 μm and gray (f). Specific colors were used for the bounding box of each particle type, yellow for 50 μm particles, red for 150 μm particles, light blue for 100 μm particle and dark blue for gray particles.

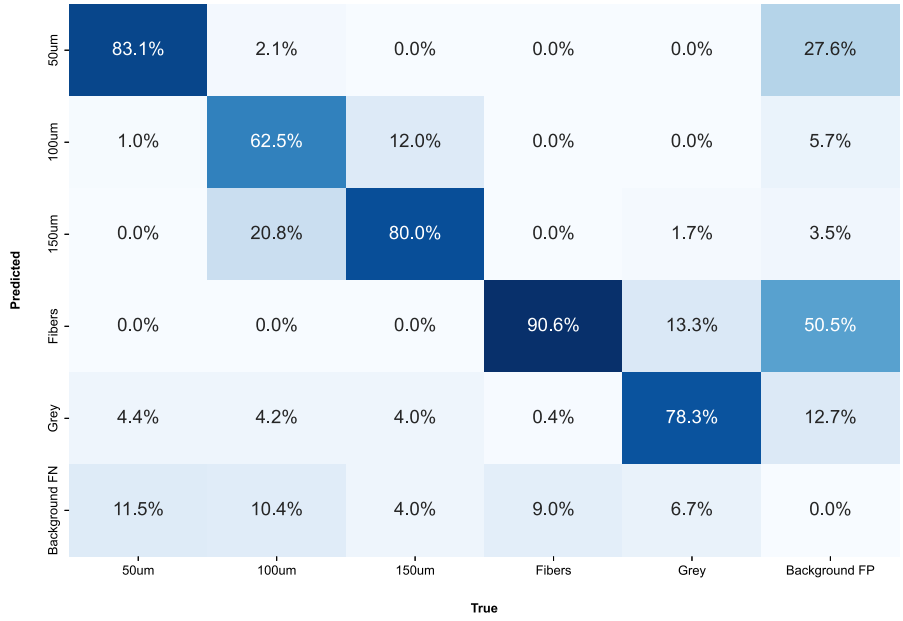


Fig. 7. Confusion matrix for all categories on the test dataset. The diagonal values represent the true positive values per category. Background false negative (FN) indicates a particle of a particular category present that is missed by the model. On the other hand, the background false positive (FP) represents a particle of a particular category is detected while there is no object annotated in that region.

a bigger amount of false negative but will be better at minimizing the false positive detection i.e. a particle is detected where in reality there is none present. Depending on the application and the goal of the study, this threshold could be optimized to a greater value such as 0.63 which correspond to the point where the curve sharply start to fall off hence giving more importance to the precision of the model compared to the recall. This would imply a reduction of the false reject rate of the prediction but would induce more cases of false negative i.e. a particle was present but was not detected.

Following this new confidence threshold determination, precision and recall per category were calculated using Eqs. (1), (2) and presented in Table 5. An overall precision and recall regrouping all classes were also calculated using the total number of true positive (TP), false positive (FP) and false negative (FN) to obtain information on how well the model would perform if the detection was considered on the presence of particle or not. A clear decreased in the recall values were noticed for all classes while better precision values were obtained.

Table 5
Comparison of the precision and recall values with the different confidence threshold considered.

	Conf 0.001		Conf 0.43	
	Precision	Recall	Precision	Recall
50 μm	82.8%	87.8%	92.4%	79.5%
100 μm	62.5%	85.7%	90.3%	75.7%
150 μm	64.5%	95.2%	87.0%	95.2%
Fibers	75.6%	90.9%	85.6%	83.3%
Gray	54.0%	92.2%	85.0%	60.7%
Overall	77.1%	90.1%	88.9%	81.2%

Model Prediction Compared to Reference Method

In adherence to regulatory requirements governing the evaluation of a novel inspection method, equivalence to the established reference

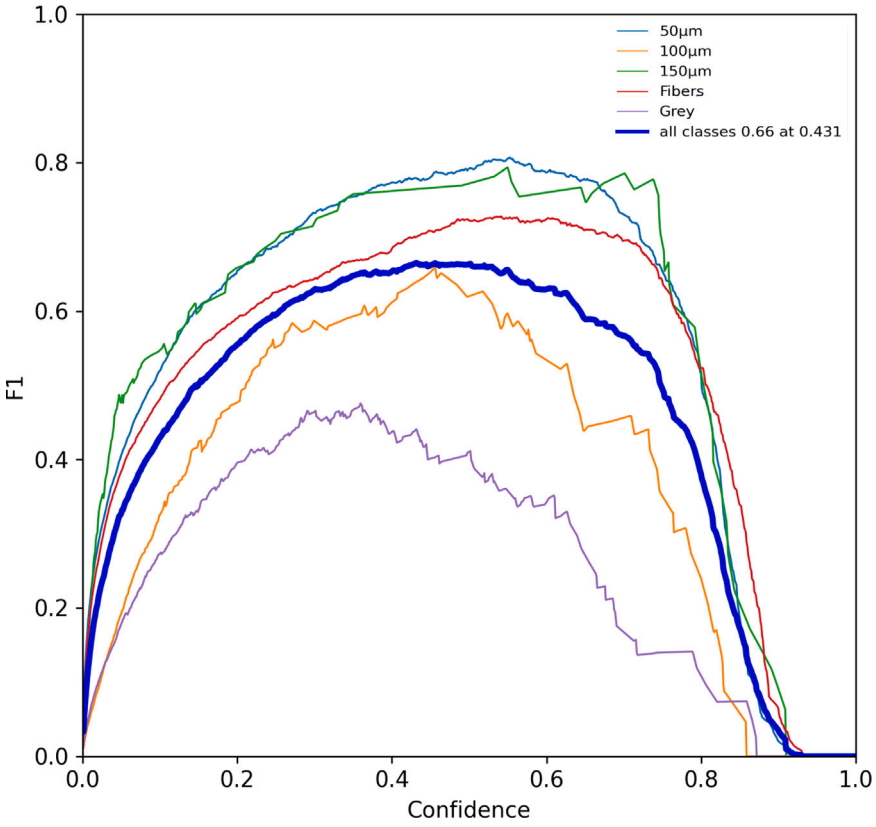


Fig. 8. Illustration of F1 curve for all categories with separate lines over all confidence levels between 0.001 and 1 on the test dataset. F1 curves are indicators for optimal cut-off levels that represent the precision and recall metrics harmoniously. An overall F1 curve representing all classes is given as a blue line.

procedure, namely visual inspection by human operators, is imperative. Consequently, the task of inspecting a predefined reference set, as outlined in the material and methods section, was assigned to five distinct operators. Each operator was given the task of writing down the number of pre-defined particles that she could see during the inspection process. Following this quality analysis, aggregated results from all operators were averaged and compared against the "expert annotations". These annotations were generated through the use of Coco Annotator (Brooks, 2019), employing the same predefined reference set comprising images captured with a high-definition camera. The comparison process extended to evaluating predictions made by the selected model YOLOv7. While half of the reference set was indented to be perfectly clean, some fibers inherent to the drug product used were still present in relatively high quantities compared to the other particle classes considered in this data set. It was clearly noted that the detection difference for this category was significant between the human inspector and the expert annotations. In most cases, at least one operator could not detect a single fiber while many more were present within the product. Given the complexity of size determination for the polystyrene beads, it was decided for the operators to not classify them according to size categories but to retain them as a unique class during the visual inspection task. However, the model being trained with the different sizes available, particles with different size are still predicted by the model. The aggregated human visual inspection results of the five different operators, the expert annotations numbers and the model annotation are presented in Table 6. Ultimately, in addition to the averaged value for each particle detection, minimum and maximum values are provided for the human inspection inside the brackets in the Human Inspector column.

From the results, a noteworthy and consistent disparity in the detection of fibers between the human inspector and YOLOv7 across all vials is visible. This difference is already noticeable with the expert annotations numbers for this category and can be explained by the increased resolution and ideal settings that the acquisition setup presents. The expert annotator was not subject to naked eye limitations and could leverage the high-resolution capabilities of the camera for the annotation task. The expert annotation revealed a significant presence of fibers which was expected to pose challenges for all human inspectors, with instances where some operators could not detect any fibers while the model could successfully identify the count close to the truth.

Additionally, YOLOv7 appeared to perform well in detecting beads, the detection results were often in agreement with the expert annotations. Some wrong classification between size determination can be explained by the high probability of misclassification if the beads were located deeper within the cake. This could also be a consequence of some mistakes made during the annotation process. It can still be noted that despite the fact that not all particles were correctly classified, the model captured efficiently all foreign objects with comparison to human inspection. Moreover, the model predictions demonstrated a level of consistency across various vials with different formulations. These differences usually led to different visual outputs with the presence of cracks as an example. On the other hand, human inspection showed more variability within its results from vial to vial but also within operators. Ultimately, some vials presented a too large amount of particles to be counted reliably by the operators hence no specific number was attributed but the label "High" was attributed. However, YOLOv7 showed some reliability in this task, high number of particles were detected and counted efficiently.

3.2. Discussion

In this work, an automated visual inspection solution based on the use of a machine vision camera and deep learning algorithm was presented. This was encouraged by the limitations that the normal procedure brings to the general quality analysis. As demonstrated in Poms

Table 6
Comparison of the particles detection and annotations on the reference set. Vial names provide information on the formulation used during the sample generation with Mannitol (M), Sucrose (S) and Trehalose (T). Expert annotations depicts the particle count made during the annotation process. Human inspector columns exhibits the numbers of instances per category counted during the inspection process with Beads corresponding to all size of particles together. YOLOv7 presents the prediction count made by the model using an IoU threshold of 0.5 and a confidence score of 0.43.

Vial name	Expert annotations					Human inspector			YOLOv7				
	50 μm	100 μm	150 μm	Fibers	Gray	Beads	Fibers	Gray	50 μm	100 μm	150 μm	Fibers	Gray
M1	1	0	0	41	5	0.8 [0;1]	1 [0;3]	0 [0;0]	4	0	0	33	2
M1b	0	0	0	12	2	0 [0;0]	0 [0;0]	0 [0;0]	1	0	0	12	0
M2b	2	0	0	25	2	0.2 [0;1]	0 [0;0]	0 [0;0]	4	1	0	26	0
M3	2	0	0	31	5	0 [0;1]	0.8 [0;3]	0 [0;0]	3	0	0	30	2
M4	0	8	0	27	3	3.2 [3;4]	0 [0;0]	0 [0;0]	4	5	0	29	3
M5	6	0	0	31	1	0 [0;0]	1 [0;2]	0 [0;0]	3	0	0	23	3
M6	2	0	6	26	3	5.8 [5;7]	0.2 [0;1]	0.2 [0;1]	4	4	2	32	5
M7	5	0	0	27	3	0 [0;0]	1.4 [1;2]	0 [0;0]	4	0	0	33	1
M8	1	0	0	15	4	0 [0;0]	0 [0;0]	0.8 [0;1]	3	0	0	17	5
M9	3	0	0	27	2	0 [0;0]	0.4 [0;1]	0 [0;0]	1	0	0	23	1
M10b	0	5	3	18	0	4.8 [4;6]	0 [0;0]	0 [0;0]	2	6	0	13	0
S1	4	0	0	25	1	0 [0;0]	0.2 [0;1]	0 [0;0]	5	0	0	34	0
S3	0	0	0	12	6	0 [0;0]	0 [0;0]	0.8 [0;2]	3	0	0	17	1
S4	0	14	0	9	2	High	0 [0;0]	0 [0;0]	1	11	3	13	3
S5	3	0	0	26	5	0 [0;0]	0.2 [0;1]	0 [0;0]	6	0	0	34	3
S6b	1	0	4	39	1	4.4 [4;5]	0 [0;0]	0 [0;0]	2	2	4	43	1
S7	0	0	0	12	0	0 [0;0]	0 [0;0]	0 [0;0]	4	1	0	17	3
S8	2	0	0	13	8	2 [0;3]	0.6 [0;1]	0.4 [0;2]	5	2	0	17	4
S9	3	0	0	12	4	0 [0;0]	0.4 [0;1]	0 [0;0]	2	2	0	23	1
S10	183	0	12	7	2	High	0 [0;0]	0.4 [0;1]	123	3	10	17	2
T1	0	0	0	10	3	0 [0;0]	0.2 [0;1]	0 [0;0]	5	0	0	22	0
T2b	116	0	0	16	0	High	0 [0;0]	0 [0;0]	113	2	0	22	0
T3	1	0	0	5	4	0 [0;0]	0 [0;0]	0 [0;0]	7	0	0	11	2
T4bb	0	12	0	17	2	10 [10;10]	0 [0;0]	0 [0;0]	1	12	2	23	2
T5	2	0	0	12	1	0 [0;0]	0 [0;0]	0 [0;0]	4	2	0	13	2
T6	0	0	7	5	0	4.8 [4;6]	0 [0;0]	0 [0;0]	0	3	8	11	1
T7	0	0	0	7	1	0 [0;0]	0 [0;0]	0 [0;0]	2	0	0	9	1
T8	1	0	0	4	2	0 [0;0]	0 [0;0]	0 [0;0]	1	0	0	5	1
T9	0	0	0	18	4	0 [0;0]	0 [0;0]	0 [0;0]	3	1	0	25	10
T10	140	4	0	20	2	High	0 [0;0]	0 [0;0]	138	4	0	24	3

et al. (2019), several challenges need to be addressed with particle analysis in parenteral products. These challenges are even enhanced when it comes to lyophilized products as particles are present in a stationary state, making their detection more difficult. Considering these challenges, the spin freeze-drying approach developed by Corver (2013) presented some clear advantages and opportunities for this visual inspection task. This procedure outcome notably expanded the surface of inspection by augmenting the accessible cake area surface while concurrently reducing layer thickness as outlined in the introduction section. Moreover, progress made in machine vision technology, coupled with the introduction of innovative high-resolution cameras, is fundamentally transforming inspection capabilities within the manufacturing environment. Notably, in the context of quality analysis where the object under interest is a cylinder, the utilization of line scan camera emerged as a strategic choice. This technology not only presents superior resolution but also operates at higher frequencies making them suitable for deployment within a continuous manufacturing line and thereby facilitating a transition towards a fully automated process aligned with the industry 4.0 vision (Arden et al., 2021).

In this study, employing a proof of concept approach under sub-optimal lightning conditions, it was possible to generate high-resolution images within 35 to 40 s utilizing long exposure times to compensate for any insufficiency in illumination caused by the use of different formulations. Typically, in the case of formulations based on Mannitol, the images generated presented less light intensity while formulations based on Sucrose or Trehalose showed presence of cracks within the cake thus leading to more contrast differences on the particle potentially in presence with more light coming through the product. Additional front lightning could allow a shorter exposure time and therefore increase the frequency of each line acquisition leading to lower acquisition time when foreseeing a full implementation in line. Following the acquisition step, specific model evaluation showcased the potential of computer vision to assist to a precise and accurate quality analysis. By leveraging the computational efficiency of YOLOv7 and its robustness in localizing small objects, it was possible to demonstrate a clear improvement in detection rate compared to human operators.

In Fig. 9, a histogram with 4 different bars per category corresponding respectively to, the amount of instances of one class, the amount of correct predictions in that class, the amount of instances non-detected for that class and the amount of false positive in that class is provided. Despite the strong imbalance of the reference data, it can clearly be noted that particles from various sizes are well handled by the model. Some misclassification are present but can in most cases be attributed to the adjacent size category as shown in the confusion matrix see Fig. 7. The fiber category also exhibited promising results as it completely outperformed human detection with a precision of 83.3% and a recall of 85.6%. However, this class also presented a high percentage of the total amount of false positive compared to the other categories. Some artifacts present within the product or thin lines non-related to any foreign matter were detected as fibers. This could still be decreased with further optimization on, the annotation, the acquisition process and model parameters such as confidence level etc. Similarly, the class 50 μm also presented a large number of false positive compare to the remaining classes. It was noticed that small artifact or contrast difference within the cake were detected as particle as well. In some cases, oversights during the annotation process were the issue, while in others, it was a false positive. Ultimately, regardless of their high percentage, this only represented an amount of instances which was significantly lower than the total number of objects present in the whole dataset.

The class gray particle presented a typical example of what can be encountered in a normal process environment i.e. intrinsic particles. This category raised typical challenges such as the varying sizes that the particle generation created, but also the varying and uncharacteristic shapes. Whilst beads and fibers presented more constant shapes and sizes, gray particles exhibited completely different sizes from comparable to the 50 μm to parts bigger than 200 μm. This led in the first

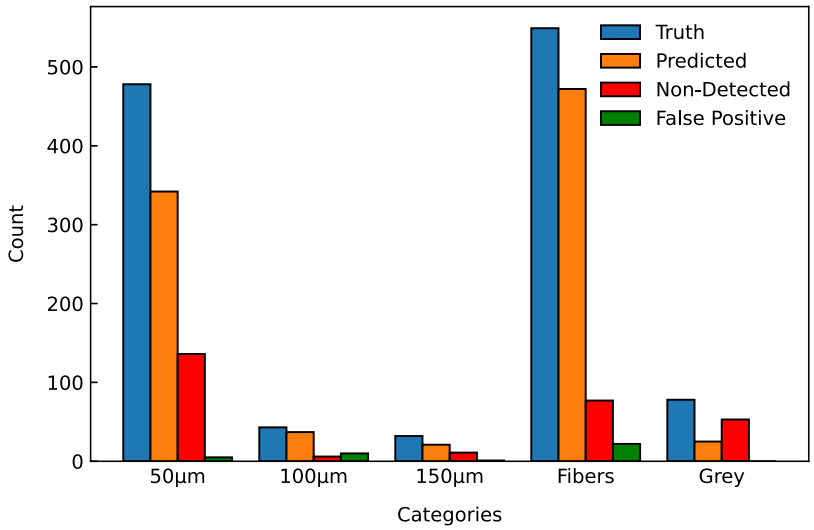


Fig. 9. Histogram summary of all vials comparing the number of instances counted by the expert (Truth), the number of instances predicted by YOLOv7 (Predicted), the number of instances non detected by the model (Non-Detected) and the number of instances wrongly predicted by the model (False Positive).

place to more difficulty during the annotation process as some instances were similar to the beads in size and shape, but also in the detection part where misclassification happened. As depicted in Fig. 7, 13% of the gray particles were detected as fibers. Both classes representing the intrinsic particles exhibited promising results but showcased the need of further development to create a more robust model. In the case of gray particles, multiple sizes and shapes could be defined as additional categories depending on their occurrence. Further optimization on the confidence threshold could be performed to decrease the number of false positive at the expense of more false negative for example.

In Table 2, all the data generated and used for the method comparison is presented, including the number of instances per individual vial in the reference dataset. Despite the absence of specificity regarding bead size for the human operator, a clear improvement in detection was demonstrated by the model YOLOv7. However, only a few examples of vials without any particles were included in the reference data, making it difficult to assess how well the model would perform when the product is free of any foreign matter. Specific categories, such as Fibers, presented promising results with clear outperformance compared to the reference method. Despite several misclassifications, YOLOv7 captured a wide majority of foreign matter in the drug, outperforming the human inspector in both efficiency and efficacy, thus answering the requirements imposed by the regulatory regarding new methods for visual inspection (Pharmacopeia, 2021).

To conclude, considering both acquisition and prediction time, only 45 s were needed to fully inspect a vial with the hardware presented in the Appendix section. This confirmed the potential that this approach presents for a continuous manufacturing line. By optimizing the setup and the acquisition conditions, a shorter time for full inspection would be achievable and would present a clear viable alternative to the human inspection with enhanced results. It is also worth mentioning that this performance difference between YOLOv7 and a typical expert operator would be further increased if the analysis was performed over a longer period as a typical visual inspection expert would be subject to a stronger error rate overtime (Poms et al., 2019). Besides the fatigue that operators experience after extended periods of performing this task, prolonged inspection work can also negatively affect their vision. As explained in the introduction, human fatigue would lead to a sharp decrease in precision whereas computer vision and deep learning can ensure a strong reproducibility across samples for a longer period without being subject to any kind of fatigue.

4. Conclusion

This paper discusses the implementation of an automated vision-based quality inspection system developed for quality analysis of continuously freeze-dried products. Through offline analysis using a dedicated setup, it was demonstrated that particle inspection for freeze-dried products is feasible. Specific particles with calibrated sizes were chosen to act as a benchmark and demonstrate the detection capabilities with size information which can be related to regulation expectations. Deep learning model and computer vision showcased the potential of an improved method in both efficiency and accuracy by providing precise measurement where the standard human inspection was struggling but also a fast computational analysis time with a full analysis available in less than 45 s.

Funding

This study was partially funded by Ghent University special research funds (BOF), by the Flemish Government under the 'Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen' program.

CRediT authorship contribution statement

Quentin Herve: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Nusret Ipek:** Writing – review & editing, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Jan Verwaeren:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Thomas De Beer:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization.

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.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijpharm.2024.124629.

References

Aldrich, D.S., Cherris, R.T., Shabushnig, J.G., 2016. Visual Inspection and Particulate Control. Davis Healthcare International Publishing, Bethesda, MD.

Arden, N.S., Fisher, A.C., Tyner, K., Lawrence, X.Y., Lee, S.L., Kopcha, M., 2021. Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future. Int. J. Pharm. 602, 120554.

Brooks, J., 2019. COCO annotator. https://github.com/jsbroks/coco-annotator/.

Cai, Z., Vasconcelos, N., 2018. Cascade r-cnn: Delving into high quality object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 6154–6162.

Committee for Medicinal Products for Human Use, 2021. Committee for Medicinal Products for Human Use ICH Guideline Q13 on Continuous Manufacturing of Drug Substances and Drug Products. Vol. 31, EMEA European Medicines Agency, URL: www.ema.europa.eu/contact.

Corver, J., 2013. Method and system for freeze-drying injectable compositions , in particular pharmaceutical,WO2013036107.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. Ieee, pp. 248–255.

Ficzere, M., Péterfi, O., Farkas, A., Nagy, Z.K., Galata, D.L., 2023. Image-based simultaneous particle size distribution and concentration measurement of powder blend components with deep learning and machine vision. Eur. J. Pharm. Sci. 191, 106611, URL: https://www.sciencedirect.com/science/article/pii/S0928098723002415.

Langille, S.E., 2013. Particulate matter in injectable drug products. PDA J. Pharm. Sci. Technol. 67 (3), 186–200.

Lin, T., Maire, M., Belongie, S.J., Bourdev, L.D., Girshick, R.B., Hays, J., Perona, P., Ramanan, D., Doll'a r, P., Zitnick, C.L., 2014. Microsoft COCO: common objects in context. CoRR abs/1405.0312. URL: http://arxiv.org/abs/1405.0312.

Mazaheri, M., Saggu, M., Wuchner, K., Koulov, A.V., Nikels, F., Chalus, P., Das, T.K., Cash, P.W., Finkler, C., Levitskaya-Seaman, S.V., et al., 2023. Monitoring of visible particles in parenteral products by manual visual inspection-reassessing size threshold and other particle characteristics that define particle visibility. J. Pharm. Sci..

Morita, D., Raytchev, B., Elhanashi, A., Kawaguchi, M., Ogata, Y., Higaki, T., Kaneda, K., Nakashima, A., Saponara, S., 2023. Fast detection of bag-breakups in pulsating and steady airflow using video analysis and deep learning. J. Real-Time Image Process. 20 (6), 114.

Padilla, R., Passos, W.L., Dias, T.L., Netto, S.L., Da Silva, E.A., 2021. A comparative analysis of object detection metrics with a companion open-source toolkit. Electronics 10 (3), 279.

Péterfi, O., Madarász, L., Ficzere, M., Lestyán-Goda, K., Záhonyi, P., Erdei, G., Sipos, E., Nagy, Z.K., Galata, D.L., 2023. In-line particle size measurement during granule fluidization using convolutional neural network-aided process imaging. Eur. J. Pharm. Sci. 189, 106563.

Pharmacopeia, U., 2021. USP< 790> visible particulate in injections.

Pokhriyal, P., Chavda, V.P., Pathak, M., 2023. Future prospects and challenges in the implementation of AI and ML in pharma sector. In: Bioinformatics Tools for Pharmaceutical Drug Product Development. Wiley Online Library, pp. 401–416.

Poms, J., Sacher, S., Nixdorf, M., Dekner, M., Wallner-Mang, S., Janssen, I., Khinast, J.G., Schennach, R., 2019. The need for new control strategies for particulate matter in parenterals. Pharm. Dev. Technol. 24 (6), 739–750.

Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. Adv. Neural Inf. Process. Syst. 28.

Snyder, J., 1995. New machine vision technology focuses on performance, ease of use, lower costs. CHILTON's I CS 68 (12), 25–28.

Sun, J., Wu, Y., Qiu, Y., Miao, C., 2024. Position and morphology detection of mixed particles based on IPI and YOLOv7. Opt. Commun. 554, 130158.

Van Bockstal, P.-J., 2018. New approach suggests continuous lyophilization is possible. URL: https://www.pda.org/pda-letter-portal/home/full-article/new-approach-suggests-continuous-lyophilization-is-possible.

Veillon, R., Shabushnig, J., 2020. 40 Years of visual inspection: Where do we go from here? PDALetters.

Wang, C.-Y., Bochkovskiy, A., Liao, H.-Y.M., 2023. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7464–7475.

Wang, C.-Y., Liao, H.-Y.M., Yeh, I.-H., 2022. Designing network design strategies through gradient path analysis. arXiv preprint arXiv:2211.04800.

Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., Girshick, R., 2019. Detectron2. https://github.com/facebookresearch/detectron2.

Zhang, L., Xiong, N., Pan, X., Yue, X., Wu, P., Guo, C., 2023. Improved object detection method utilizing YOLOv7-tiny for unmanned aerial vehicle photographic imagery. Algorithms 16 (11), 520.