EXPRESS PAPER

Open Access

CrossMark

In-line recognition of agglomerated pharmaceutical pellets with density-based clustering and convolutional neural network

Andraž Mehle[1*], Boštjan Likar[1,2] and Dejan Tomaževič[1,2]

Abstract

We present a method for recognition of agglomerates in images acquired during the coating process of pharmaceutical pellets. The pellets in the images are not perfectly dispersed, and it is often hard to differentiate between a random group of primary particles and a real agglomerate. The method utilizes a clustering-based image segmentation for candidate region detection and a convolutional neural network for classification of detected pellets to primary particles or agglomerates. We validated the performance of the method on real images of pharmaceutical pellets acquired during the coating process and achieved 93% classification accuracy.

Keywords: Particle recognition, Agglomeration detection, Neural network classification

1 Introduction

Pharmaceutical, chemical, cosmetic, and food industries utilize coating processes to modify particle properties, e.g., to improve esthetic appearance, mask odor and taste, enhance chemical and physical stability, flow ability, compressibility, etc. In the pharmaceutical industry, the coating process is often applied on pellets, i.e., small spherical particles with a narrow size distribution, which are then enclosed in capsules or compressed into tablets and contain the active pharmaceutical ingredient.

Pharmaceutical pellets are most commonly coated using a fluidized-bed coating method [1], where the carrier gas acts as a fluidizing medium for the particles and forces them into a circulating movement. During the coating, a spray nozzle inside the coater continually applies the coating dispersion onto the pellets.

One of the most undesirable phenomena that may significantly affect the coating process yield and the coating quality/uniformity is the agglomeration of particles being coated. Agglomeration occurs when a liquid bridge, due to an excess of the applied coating dispersion, is formed between primary particles and remains until complete

drying. The process parameters (e.g., temperature, quantity, and humidity of the fluidizing air, and spray rate) should be set such that the applied coating dispersion dries fast enough to prevent the agglomeration. However, since short production time is generally desired, coating processes are often driven close to the edge of the process design space, meaning only small variations in the process can lead to agglomeration.

It is common practice to assess the mass fraction of agglomerates at the end of the coating process by mechanically separating agglomerates and primary particles, e.g., by sifting. The final process yield is then determined based on the weight of dry materials entering the coating process (the mass of the particles and the dry coating dispersion mass) and on the weight of dry materials exiting the coating process (the mass of the coated particles without agglomerates). However, the sifting method has certain drawbacks: it is invasive (some agglomerates may break), it is time consuming, and it can be done only at the end of the process, when it is already too late for any intervention.

The pharmaceutical industry recently tends towards the development of novel analytical tools for monitoring of manufacturing processes, with the goal of ensuring the final product quality. The U.S. Food and Drug Administration (FDA) issued process analytical technology (PAT) guidance [2] to encourage the development of such tools.

*Correspondence: andraz.mehle@sensum.eu
[1]Sensum, Computer Vision Systems, Tehnološki park 21, 1000 Ljubljana, Slovenia
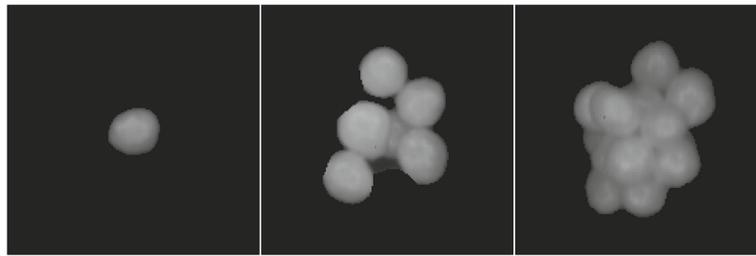Full list of author information is available at the end of the article

Mehle *et al. IPSJ Transactions on Computer Vision and Applications*   (2017) 9:7

Page 2 of 6

Fig. 1 A primary particle (*left*), a group of primary particles (*middle*), and an agglomerate of primary particles (*right*)

In scope of PAT guidance, it would be advantageous to estimate and to monitor the amount of agglomerates during a coating process in real time.

In the past few years, various optical PAT methods were investigated for detection of agglomeration in fluidized-bed pellet coating processes. Wiegel et al. [3] explored spatial filter velocimetry (SFV) for in-line detection of agglomeration. They were able to detect the occurrence of agglomeration from changes in size distributions of particles. However, the SFV method has certain drawbacks: very narrow field of view, indirect size measurement based on one-dimensional particle chord lengths, and inability to detect overlapping particles. Malvern Instruments Ltd released an application note [4] describing the image analysis method for recognition of agglomerates based on multiple morphological properties of particles, such as size, convexity, and/or circularity. However, their method requires well-dispersed particle samples, making it unsuitable for in-line recognition of agglomerates. Možina et al. [5] proposed that visual imaging could be used as a PAT tool for pellet coating processes and demonstrated the possibility of agglomeration detection. However, they tested their method offline under controlled imaging conditions, where pellets and agglomerates were spatially well separated on an image plane.

There are many visual features that can generally be used to differentiate between primary particles and agglomerates. In case of pharmaceutical pellets, which are of approximately equal size, the particle size or shape would generally be sufficient. However, the pellets in images acquired in-line are not located in a single plane and are not perfectly dispersed; thus, many particles appear in groups in the acquired images. The crucial step of image analysis for recognition of agglomerates is to differentiate, in the groups of particles, those particles that are only visually in contact (i.e., they are occluded or overlapped from the point of view of a camera) from actual physical agglomerates (Fig. 1).
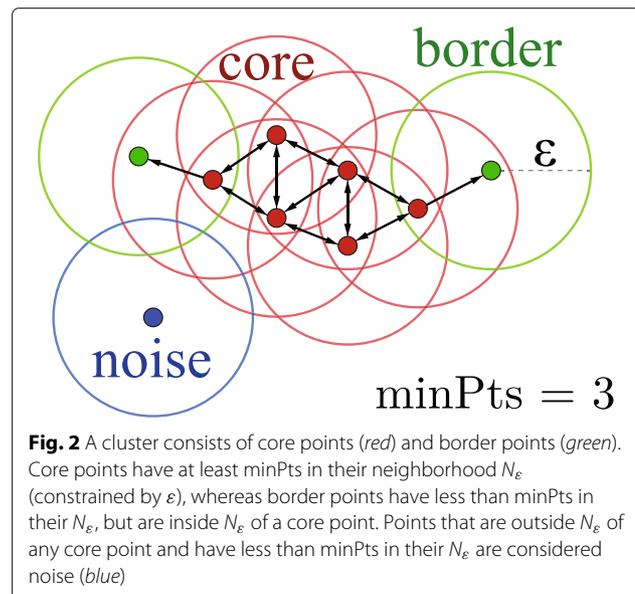
To address this problem, we propose a method for recognition of agglomerates based on density-based clustering and convolutional neural network. We validated the performance of the method on real images of pharmaceutical pellets acquired during the coating process.
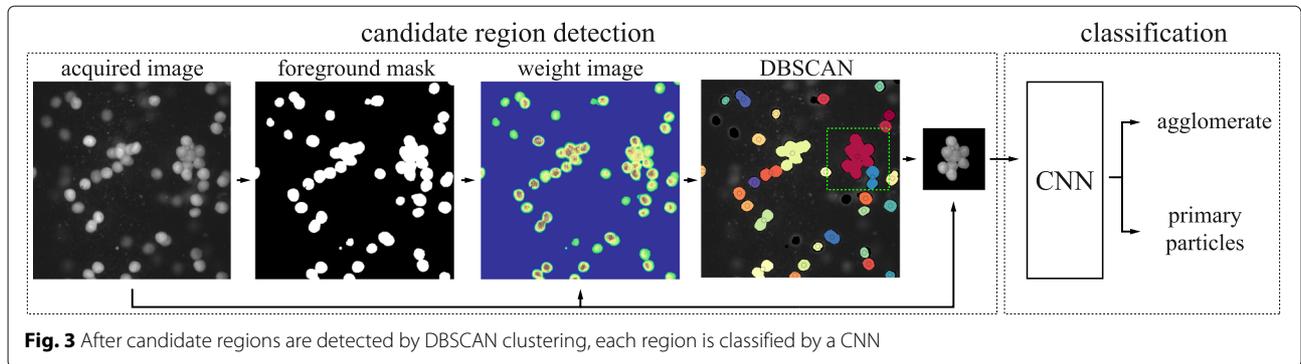
2 Method

The proposed method comprises two major stages: first, candidate regions are detected by clustering-based image segmentation, such that groups of primary particles that are possibly agglomerated represent a single region. Next, the candidate regions are classified as primary particles or agglomerates by a convolutional neural network trained on images of candidate regions and a manually obtained ground truth.

2.1 Candidate region detection

First, a global threshold value is determined by Otsu's method [6] and image thresholding is performed to extract the binary mask of the foreground, i.e., the particle region on an image. The foreground is further segmented to candidate regions by DBSCAN clustering algorithm [7]. DBSCAN is a density-based algorithm for discovering arbitrarily shaped clusters in large spatial data sets with noise, where the number of clusters is determined automatically. Furthermore, DBSCAN clusters only points in dense areas, whereas points in sparse areas are considered outliers or noise (Fig. 2).

Fig. 2 A cluster consists of core points (*red*) and border points (*green*). Core points have at least minPts in their neighborhood N_ε (constrained by ε), whereas border points have less than minPts in their N_ε, but are inside N_ε of a core point. Points that are outside N_ε of any core point and have less than minPts in their N_ε are considered noise (*blue*)

Mehle *et al. IPSJ Transactions on Computer Vision and Applications* (2017) 9:7

Page 3 of 6

Fig. 3 After candidate regions are detected by DBSCAN clustering, each region is classified by a CNN

For segmentation of agglomerates, each image pixel inside the foreground region represents an input point p for clustering. Furthermore, each point p is given a weight $W(p)$ based on intensity gradient magnitude $G(p)$ normalized by the image intensity $I(p)$ in that point:

$$W(p) = 1 - G(p)/I(p), \tag{1}$$

where $I(p)$ and $G(p)$ are normalized to values between 0 and 1. With this notation, points with low intensity and high gradient magnitude, e.g., points at the border of pellets are given lower weights than points in the central area of pellets. Furthermore, high gradient magnitudes that may appear in the central area are suppressed. A density in the N_ε of a given point is then calculated as the sum of weights $W(p)$ in the $N_\varepsilon(p)$:

$$d(p) = \sum_{p_i} W(p_i), \; p_i \in N_\varepsilon(p). \tag{2}$$

The described candidate region detection with DBSCAN (Fig. 3) separates the particles based on image gradients and intensities, which allows for easy classification of the primary particles and the agglomerates based on their size. However, due to particle overlapping or occlusion there are many groups of primary particles where the separation with DBSCAN fails. Thus, we propose a classification based not only on particle size but on a set of learned visual features that can better differentiate between primary particles (single or grouped) and real agglomerates.

2.2 Classification

After the candidate region detection, each particle inside the candidate region is classified as a primary particle or as an agglomerate (Fig. 3). For this purpose, the center of gravity of the candidate region is calculated, and the intensity values of each candidate region are mapped into the center of an empty candidate image of a fixed size, i.e., an input for classification. The size of the candidate images
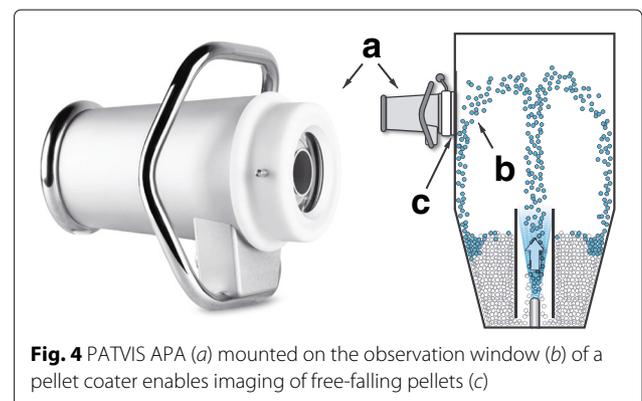
is selected based on the expected maximal size of imaged particles. The classification is based on machine learning method that utilizes a deep convolutional neural network.

A convolutional neural networks (CNN) is a type of neural network with its architecture primarily designed for object recognition tasks [8]. Generally, CNN architecture comprises three basic structures: convolutional layers, pooling layers, and fully connected layers.

The convolutional layer aims to learn the parameters of filters that activate on some type of visual feature of the inputs. It comprises a bank of filters that perform a 2D filtering (convolution) on the input image data and produce a 2D feature map.

The pooling layer is used for sub-sampling (dimensionality reduction) and feature selection by merging local information. Therefore, it compresses or generalizes feature representations and generally reduce the overfitting of the model to the training data.

The fully connected layers have neurons with connections to all activations in the previous layer (as seen in regular neural networks) and are used at the end of the CNN. The last fully connected layer in CNNs used for classification has a non-linear activation function or a softmax activation in order to output probabilities of class predictions.
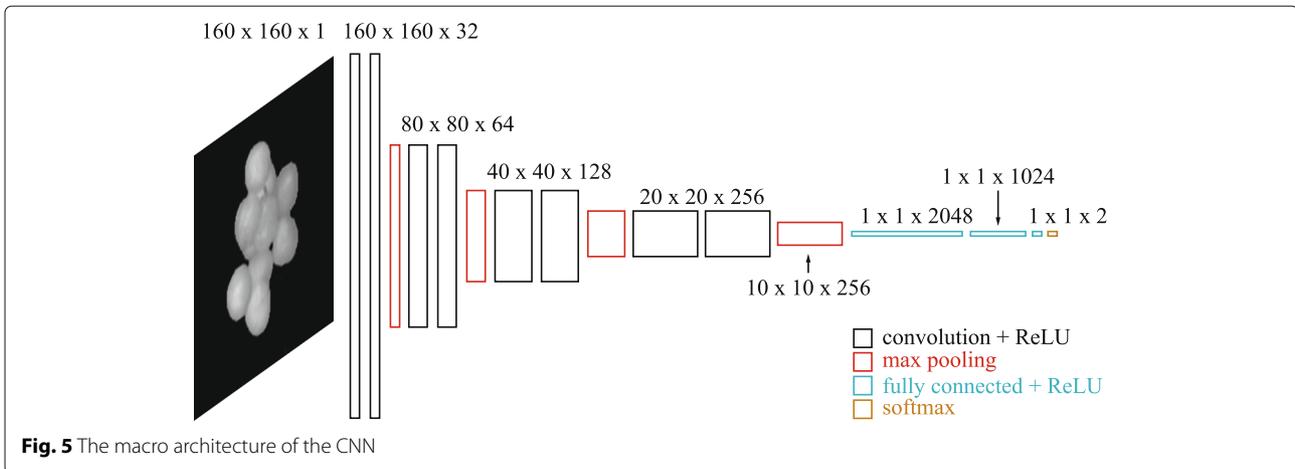

Fig. 4 PATVIS APA (*a*) mounted on the observation window (*b*) of a pellet coater enables imaging of free-falling pellets (*c*)

Mehle *et al. IPSJ Transactions on Computer Vision and Applications* (2017) 9:7

Page 4 of 6

Fig. 5 The macro architecture of the CNN

3 Experiment

3.1 Coating process
We executed a coating process of pellets in a pilot-scale fluidized-bed coater. The process parameters were deliberately set to induce substantial particle agglomeration.

3.2 Image acquisition
Pellet images were acquired by an in-line visual inspection system PATVIS APA (Sensum, Slovenia [9]) through the observation window of the fluidized-bed coater. The mounting position enabled imaging of the pellets in free fall (Fig. 4). The image resolution was 512×512 pixels with pixel size of 31.2 μm. Due to telecentricity of the lens, the pixel size was independent of the distance from the camera.

3.3 Implementation details
After the image segmentation by DBSCAN, the regions that were substantially smaller than the pellets too dark or contacting the image border were discarded from further analysis. Candidate images were normalized by subtracting the mean and dividing by the standard deviation.

The convolutional neural network model used in our experiments was based on the network used by the VGG team in the ILSVRC-2014 competition [10]. The architecture of our network is presented in Fig. 5. The size of all convolutional filters and max-pooling filters was 3×3 and 2×2, respectively.

The training of the CNN was performed in batches of 50 images using stochastic gradient descent with Nesterov momentum for optimization. The learning rate was slightly decreased after each batch.

3.4 Training and validation
The experimental image database used for training and evaluation of the CNN consisted of 2000 candidate images with various region sizes. We estimated the average area

of a single pellet A and assort candidate images into size classes based on A (Table 1, Fig. 6).

The ground truth was obtained by manually classifying candidate images. With regard to the ground truth, 1100 images included single particles and 900 images included agglomerates.

The image database was randomly divided into the training set (60%, 1200 images), the validation set (20%, 400 images), and the test set (20%, 400 images).

The training images were augmented by randomly applying noise with standard deviation in the range from 0 to 0.01 and rotation from 0 to 360° (other spatial transformations, such as scale and perspective, were not considered, since the camera was still and telecentric lens was used). Consequently, the model was trained with altered training images in each training cycle. This substantially reduces the overfitting of the model to the training data, hence, providing the model that generalizes well.

Once the model was trained, its performance in terms of classification accuracy was validated on the test image set. Furthermore, the receiver operating curve (ROC) was calculated to assess the performance of the classifier over its entire operating range. The performance was compared to the classification based on the area of candidate regions (i.e., an estimate of particle size).

4 Results and discussion
The trained CNN model achieved 93% classification accuracy on the test set. Furthermore, the obtained ROC curve (Fig. 7) shows that the CNN classifier substantially outperformed the area-based classification. At false positive rate

Table 1 The size classes and approximate percentage of corresponding candidate images in the database

Size class	A	$2A$	$3A$	$4A$	$5A$	$6A$	$\geq 7A$
Approx. %	30	20	10	10	10	10	10

Mehle *et al. IPSJ Transactions on Computer Vision and Applications* (2017) 9:7
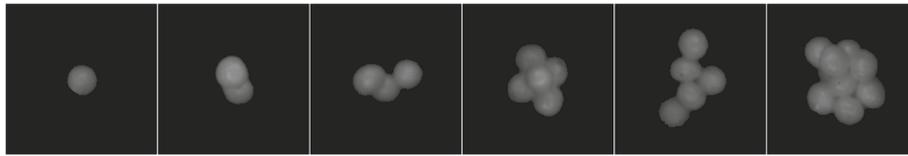
Page 5 of 6

Fig. 6 Examples of candidate images of various size classes

0.05, the CNN classifier achieved 0.92 true positive rate (threshold 0.46), while the area-based classifier achieved 0.70 true positive rate (area threshold $4.0A$, A – area of a single pellet). Note that high area threshold of $4.0A$ does not allow for the detection of agglomerates with smaller areas that usually form at the beginning of agglomeration occurrence, which substantially reduces the sensitivity of the method and the possibility of timely intervention. On the contrary, the CNN classifier shows good performance independent of the candidate region area; it was able to correctly recognize 92% of agglomerates while falsely recognizing 5% of primary particles as agglomerates.

We chose the machine learning approach because it is hard to manually design features for classification that are general enough for reliable differentiation between primary particles and agglomerates in various particle constellations.

The performance of the candidate region detection was not assessed separately because in real application it is not crucial to recognize all particles in each image due to high image acquisition rate that easily provides statistically significant sample.

The main drawback of the proposed classification method is that a candidate image may generally include both agglomerates and primary particles. This introduces ambiguities to both the training and the prediction of the

CNN classifier. However, the acquisition of the ground truth used in our experiment implied that those cases are rare. Furthermore, while obtaining the ground truth, it is sometimes very hard to decide whether a group of particles is a real agglomerate or not; thus, it would be advantageous for the training stage to separately acquire images of only primary particles and only real agglomerates.

5 Conclusions

The results show that the use of a trained convolutional neural network classifier substantially improves the recognition of agglomerates compared to classification based on particle size.

Our future work includes the evaluation of classification performance of the trained model on other images from the same coating process and on images from different coaters. We want to investigate whether it is possible to obtain a general model for various coating processes or should the model be trained for each process separately.

The recognition of agglomerates of particles is not only beneficial to the fluidized-bed coating processes but also to other processes in which agglomeration of particles occurs either as an accurately controlled process or as an undesired phenomenon. In both cases, real-time estimation of the amount of agglomerates provides means for process monitoring and process control. Such processes
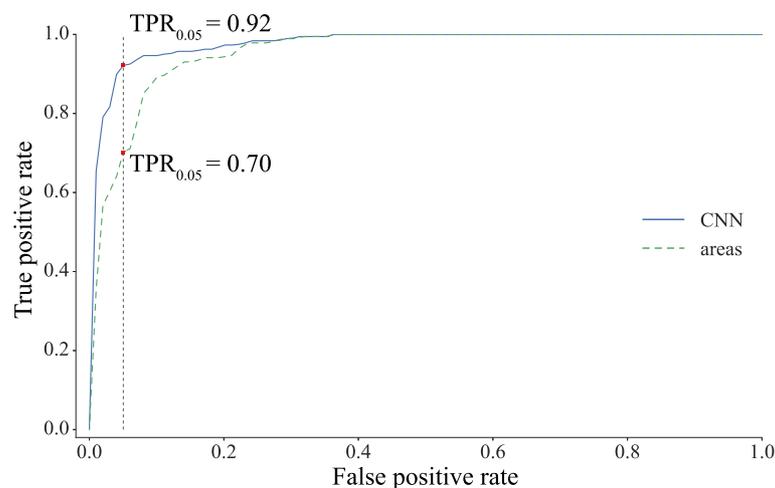

Fig. 7 Receiver operating curve for classification with CNN and classification based on areas

Mehle *et al. IPSJ Transactions on Computer Vision and Applications* (2017) 9:7

Page 6 of 6

include, for example, pan coating, fluid-bed gasification, combustion and polymerization, fluid-bed, high-shear or twin-screw granulation, crystallization, bioreactor fermentation, water purification, and flocculation.

Authors' contributions

AM designed the agglomerate recognition framework, performed the experiments, and drafted the manuscript. BL and DT supervised the work and revised the manuscript critically. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Teunou E, Poncelet D (2002) Batch and continuous fluid bed coating — review and state of the art 53(4):325–340. doi:10.1016/S0260-8774(01)00173-X
2. US Food and Drug Administration (2004) Guidance for industry: PAT – A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. [Internet]. US Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. Available from: https://www.fda.gov/. URL to PDF: https://www.fda.gov/downloads/drugs/guidances/ucm070305.pdf. Accessed 8 Dec 2016
3. Wiegel D, Eckardt G, Priese F, Wolf B (2016) In-line particle size measurement and agglomeration detection of pellet fluidized-bed coating by spatial filter velocimetry 301:261–267. doi:10.1016/j.powtec.2016.06.009
4. Malvern Instruments Ltd (2015) Identification of Agglomerates Using Automated Image Analysis (Application Note) [Internet]. Malvern Instruments Ltd, Grovewood Road, Malvern, Worcestershire, UK. WR14 1XZ. Available from: http://www.malvern.com/identifying-agglomerates. Accessed 8 Dec 2016
5. Možina M, Tomaževič D, Leben S, Pernuš F, Likar B (2010) Digital imaging as a process analytical technology tool for fluid-bed pellet coating process 41(1):156–162. doi:10.1016/j.ejps.2010.06.001
6. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9(1):62–66. doi:10.1109/TSMC.1979.4310076
7. Ester M, Kriegel HP, Sander J, Xu X (1996) A , density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. AAAI Press, Portland Vol. KDD'96. pp 226–231
8. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989) Backpropagation applied to handwritten zip code recognition 1(4):541–551. doi:10.1162/neco.1989.1.4.541
9. Sensum (2016). http://www.sensum.eu. Accessed 28 Nov 2016
10. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556. http://arxiv.org/abs/1409.1556