Classification of scanning electron microscope images of pharmaceutical excipients using deep convolutional neural networks with transfer learning

Hiroaki Iwata, Yoshihiro Hayashi, Aki Hasegawa, Kei Terayama, Yasushi Okuno

PII: S2590-1567(22)00024-X

DOI: https://doi.org/10.1016/j.ijpx.2022.100135

Reference: IJPX 100135

To appear in:

Received date: 25 August 2022

Revised date: 12 October 2022

Accepted date: 13 October 2022

Please cite this article as: H. Iwata, Y. Hayashi, A. Hasegawa, et al., Classification of scanning electron microscope images of pharmaceutical excipients using deep convolutional neural networks with transfer learning, (2022), https://doi.org/10.1016/j.ijpx.2022.100135

Original Research

Classification of scanning electron microscope images of pharmaceutical excipients using deep convolutional neural networks with transfer learning

Hiroaki Iwata[1,#], Yoshihiro Hayashi[*,1,2,#], Aki Hasegawa[1], Kei Terayama[3], Yasushi Okuno[**,1]

[1]Graduate School of Medicine, Kyoto University, 53 Shogoin-kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan

[2]Pharmaceutical Technology Division, Nichi-Iko Pharmaceutical Co., Ltd.; 205-1, Shimoumezawa Namerikawa-shi, Toyama, 936-0857, Japan,

[3]Graduate School of Medical Life Science, Yokohama City University, Tsurumi-ku, Yokohama 230-0045, Japan

[#]These authors contributed equally to this work.

[*]Corresponding author. Pharmaceutical Technology Division, Nichi-Iko Pharmaceutical Co., Ltd.; 205-1, Shimoumezawa Namerikawa-shi, Toyama, 936-0857, Japan

[**]Corresponding author. Graduate School of Medicine, Kyoto University, 53 Shogoin-kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan.

E-mail address: yoshihiro-hayashi@nichiiko.co.jp (Y. Hayashi) and okuno.yasushi.4c@kyoto-u.ac.jp (Y. Okuno).

ABSTRACT

Convolutional Neural Networks (CNNs) are image analysis techniques that have been applied to image classification in various fields. In this study, we applied a CNN to classify scanning electron microscopy (SEM) images of pharmaceutical raw material powders to determine if a CNN can evaluate particle morphology. We tested 10 pharmaceutical excipients with widely different particle morphologies. SEM images for each excipient were acquired and divided into training, validation, and test sets. Classification models were constructed by applying transfer learning to pretrained CNN models such as VGG16 and ResNet50. The results of a 5-fold cross-validation showed that the classification accuracy of the CNN model was sufficiently high using either pretrained model and that the type of excipient could be classified with high accuracy. The results suggest that the CNN model can detect differences in particle morphology, such as particle size, shape, and surface condition. By applying Grad-CAM to the constructed CNN model, we succeeded in finding particularly important regions in the particle image of the excipients. CNNs have been found to have the potential to be applied to the identification and characterization of raw material powders for pharmaceutical development.

Keywords: Convolutional neural networks, Machine learning, Scanning electron microscope, Excipients, Powder, artificial intelligence

Abbreviations: API, active pharmaceutical ingredient; CNNs, convolutional neural networks; FN, false negative; FP, false positive; Grad-CAM, gradient-weighted class activation mapping; SEM, scanning electron microscopy; TN, true negative; TP, true positive

1. Introduction

In the pharmaceutical design of solid dosage forms, it is important to observe particles of raw materials on the nano- and microscales. This is because particle morphology such as particle size, particle shape, and surface condition has a strong influence on process performance and quality attributes of pharmaceutical products (Horio et al., 2014; Kudo et al., 2020; Paul et al., 2018). For instance, Horio et al. evaluated the effect of particle shape on powder flowability using eight types of microcrystalline cellulose (Horio et al., 2014). They showed that elongated particles tend to increase the strength of interactions between particles and resulted in poor flowability. After tablets are produced, elongated particles tend to have a greater hardness than spherical particles. Alyami et al. evaluated particle roughness and content uniformity of five excipients (Alyami et al., 2017). They showed that surface roughness affects homogeneity and content uniformity of fine and low-dose active pharmaceutical ingredients (APIs).

A scanning electron microscope (SEM) is a device that uses an electron beam to observe the surface of a sample. SEMs have high resolution and can produce magnified images of tens to millions of times magnification. SEMs are widely used to evaluate particle morphology of raw material powders of solid dosage forms because they can observe details of particles. Image analysis techniques for SEM images of particles have been developed to evaluate particle morphology quantitatively (Hamishehkar et al., 2010; Holgado et al., 1996; Myshkin et al., 1992; Passerini et al., 2002), but when observing particles, only qualitative evaluation of SEM images is often performed. This is mainly due to the need to measure individual particles manually, which requires an enormous amount of time for image analysis. Because the particle shape and distribution of many APIs and excipients are nonuniform, it is necessary to evaluate many particles to obtain accurate characterization

values (Silva et al., 2013). However, the shape of API and excipient particles is rarely simple, such as a perfect sphere, and they often have cohesive properties. Thus, it is difficult automatically to obtain the contours of individual particles. Shekunov et al. reported that SEM is the most essential technique for particle size analysis but that it is not very reliable for sizing by itself and should always be used in combination with other techniques (Shekunov et al., 2007). There are large statistical errors and biases associated with preferential particle orientation and particle agglomeration that are difficult to control and minimize. Therefore, image analysis techniques for SEM images of particles are not considered practical at present.

Convolutional neural networks (CNNs) are a type of deep learning structure designed for the classification and analysis of digital images and have recently attracted much attention in the field of computer vision and image processing (Abbas et al., 2019). CNNs have been applied in a wide range of fields including disease diagnosis (Lu et al., 2019), cell cycle phase classification (Nagao et al., 2020), plant classification (Onishi and Ise, 2021), and egg quality prediction (Ienaga et al., 2021). In the fields of pharmaceutical sciences, powder technology, and similar fields, several applications of CNNs have been reported (Calderon et al., 2018; Ficzere et al., 2022; Gambe-Gilbuena et al., 2020; Hirschberg et al., 2020; Lu et al., 2021; Ma et al., 2020; Probst et al., 2020; Yu et al., 2021). Ma et al. have developed an analysis program that incorporates a CNN to fully automate X-ray computed tomography image analysis of oral tablets for internal crack detection (Ma et al., 2020). Ficzere et al. showed that CNN can be used to detect defects in tablet film coatings in real time and further classify the defects into five categories (Ficzere et al., 2022). Lu et al. employed a CNN to identify, classify, and quantify three different physical mechanisms of emulsion in microscopy images automatically (Lu et al., 2021). They showed that the CNN achieved good performance with high classification accuracy. CNNs have also been applied to SEM

images to classify morphological features (Azimi et al., 2018; Ge et al., 2020; Modarres et al., 2017; Yu et al., 2021). For instance, Modarres et al. applied a CNN to multiple SEM images and successfully classified 10 categories including particles, nanowires, fibers, coated surfaces, and pillars (Modarres et al., 2017). Yu et al. applied CNN to SEM images of spray-dried microcapsules and successfully classified two morphological characteristics (aggregated or nonaggregated, and intact or broken) (Yu et al., 2021).

Because CNNs have been widely applied as an image analysis technique, it could be applied to SEM images of pharmaceutical raw material powders. In other words, there is a possibility that particle classification and feature extraction can be performed by applying a CNN to SEM images. However, to our knowledge, there are still few examples of the application of CNNs to SEM images of raw materials or solid dosage forms, and it is unclear whether CNNs can classify types of particles and identify the characteristics of particles. If we can construct a classification model of SEM images of several particles, we can show that a CNN can recognize differences in particle size, particle shape, and surface roughness from SEM images, which may be applied to the development of a quality evaluation system for raw material powder. That is, an identification system for the type of sample and an evaluation system for lot-to-lot variation. For example, if the CNN model can classify lot differences in excipients for the same product, it may be possible to determine the particle morphology changes from lot to lot. Conversely, if the CNN model cannot classify lot differences, it may provide some assurance that there are no significant differences in particle morphology. Furthermore, if features related to differences in particles could be extracted, CNNs may be helpful in elucidating the cause. In pharmaceutical manufacturing, lot-to-lot variation in excipients is one of the factors that significantly affect quality attributes, and identifying the factors contributing to the variation is important for stable manufacturing (Thoorens et al., 2015; Wang et al., 2013).

The purpose of this study was to assess if a CNN can classify SEM images of pharmaceutical raw material powder used in oral solid dosage forms. Fig. 1 shows the experimental flow in this study. First, we selected 10 excipients with widely different particle morphologies as model raw material powders. Next, powder classification models were constructed by applying transfer learning to pretrained CNN models. Thus, a model trained on a large number of images was used as the initial structure, fully connected layers were replaced, and only the weights of the deeper layers were trained on our data set. The classification accuracy of the constructed model was then evaluated using a confusion matrix and statistical indices. Finally, we visualized the image regions that are important for the classification of the constructed CNN model.

Fig. 1

2. Materials and methods

2.1. Materials

Ten types of excipients, namely, microcrystalline cellulose (CEOLUS KG-802, KG-1000, PH-101, and PH-302, Asahi Kasei Chemicals Co. Ltd, Tokyo, Japan), lactose monohydrate (Pharmatose 100M, 125M, and 200M, DFE Pharma, Goch, Germany), mannitol (Mannit Q, Mitsubishi Life Science Institute, Tokyo, Japan; PEARLITOL 200SD, Roquette, Lestrem, France), and cornstarch (dried corn starch, Matsutani Chemical Industry Co., Ltd, Hyogo, Japan), were purchased commercially.

These excipients are widely used in the pharmaceutical development of oral solid dosage forms. Microcrystalline cellulose is one of the most important tableting excipients thanks to its outstanding dry binding properties, enabling the manufacture of tablets by direct compression (Thoorens et al., 2014). Mannitol and lactose are brittle excipients frequently

used as a filler (Janssen et al., 2021; Paul et al., 2018). Corn starch is used as a disintegrant (Desai et al., 2016).

2.2. Image acquisition using a scanning electron microscope

The 10 types of excipients were observed using an SEM on a Hitachi Miniscope TM1000 (Hitachi High-Technologies Co., Tokyo, Japan). The samples were prepared on carbon tape without evaporation coating. The magnification of the microscope was set to 150× for PEARLITOL 200SD, 150× for Pharmatose 200M, 250× for CEOLUS KG-802, KG-1000, PH-101, PH-302, and Mannit Q, and 500× for dried corn starch. Ten SEM images were captured for each excipient and a total of 100 original SEM images were obtained. The original SEM image properties were 1,280 × 960 pixels, horizontal and vertical resolution was 146 dpi, and 24 bits depth.

2.3. Image segmentation

The flow chart of image segmentation is shown in Fig. S1 (Supplementary Material). Ten original SEM images for each excipient were divided into training, validation, and test sets. The number of original SEM images in each dataset was set to 6, 2, and 2, respectively. Each original SEM image of 1,280 × 960 pixels was sequentially divided, resulting in a total of 20 subimages of 224 × 224 pixels. As a result, the number of subimages in the training, validation, and test set was set to 120, 40, and 40, respectively. Image augmentation methods such as rotation, height and width shift, or zoom were not applied because the original images needed to be preserved to ensure the authenticity of the features extracted from the training dataset.

2.4. Construction of the CNN model using the pretrained model and transfer learning

A CNN model for classifying particle images was constructed by applying transfer learning to a pretrained model. Transfer learning is a method that uses CNN models that have already been trained on a large set of image data called ImageNet (Lu et al., 2019; Yu et al., 2021). We focused on two types of pretrained models, namely, VGG-16 (Simonyan and Zisserman, 2015) and ResNet50 (He et al., 2016). Both models have received excellent results in the ImageNet Large Scale Visual Recognition Challenge, a contest for image classification accuracy (Russakovsky et al., 2015), and are among the commonly used CNN models for transfer learning (Morid et al., 2021). VGG16 has 13 convolutional layers and three fully connected layers for a total of 16 layers, whereas ResNet50 has a deeper network structure with a total of 50 layers. Both pretrained models can classify images into 1,000 object categories with high accuracy. To use the pretrained model for our task, the fully connected layers of the pretrained model were replaced with the softmax classifier as the final layer, the ReLu activation as the hidden layer, and a dropout layer to suppress overfitting. Only the parameters of the fully connected and last convolution layer of the CNN model were trained on our data, while the other layers were frozen, i.e., the parameters of the pretrained model were used. To optimize training parameters, we examined various values (learning rate: 1×10^{-1}, 1×10^{-2}, 1×10^{-3}, 1×10^{-4}, and 1×10^{-5} and batch size: 16, 32, 48, and 64). The values that yielded the highest accuracy were selected. The epoch and momentum were fixed at 100 and 0.9, respectively.

2.5. Performance evaluation of the CNN model

A confusion matrix was used to describe the achieved model classification showing true positive (TP), false negative (FN), false positive (FP), and true negative (TN). Then, the

values from the confusion matrix were used to measure performance by calculating accuracy, precision, recall, and F-measure defined by the following equations (Fawcett, 2006).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TN+FP}} \quad (3)$$

$$\text{F} - \text{measure} = 2 \times \frac{\text{Pecision} \times \text{Recall}}{\text{Precision+Recall}} \quad (4)$$

The performance of the CNN model was evaluated by 5-fold cross-validation as shown in Fig. S2 (Supplementary Material). The 10 SEM images were divided so that the training set, validation set, and test set contained 6, 2, and 2 SEM images, respectively. Note that the 10 SEM images were acquired so as not to overlap the area captured when the SEM images were taken. Therefore, the test set was used as external validation data to evaluate the prediction accuracy for unknown images that were not used for training and validation. In addition, five datasets were created with different combinations of SEM images included in each dataset. The CNN model was iteratively constructed using these datasets, and the classification accuracy with respect to the test set was evaluated five times repeatedly. Finally, the average value of each statistical index was calculated.

2.6. Feature visualization

Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al., 2020) was used to verify which parts of the SEM image the trained CNNs focused on for classification. Grad-CAM is a method that uses the gradient information input to the last all-junction layer of a CNN to identify the regions most involved in class prediction. In other words, it is possible to visualize the basis for the decision.

2.7. Computer software

All of the basic algorithmic operations concerning the CNN approaches and the Grad-CAM technique were implemented using the TensorFlow and Keras deep learning libraries for Python (Abadi et al., 2016).

3. Results and discussion

3.1. SEM images and particle properties of model excipients

SEM images and characteristics of the ten excipients are summarized in Fig. 2 and Table 1. Because the size of the particles varies greatly from excipient to excipient, the magnification was taken at different settings for each excipient. In our opinion, for the CNN model to learn SEM images of particles, the particles must be present in the image at an appropriate size so that their characteristics can be identified. Therefore, the magnification was set low enough so that at least one or more particles were present in the divided image and the approximate outline of the particles could be determined. Moreover, for excipients with smaller particles, the magnification was set high enough to allow humans to grasp the characteristics of the particles. For example, PEARLITOL 200SD was a very large particle, so the magnification was low, while dried corn starch was a small particle, so the magnification was high. However, for the four types of CEOLUS and three types of Pharmatose, the magnification was standardized within each brand so that comparisons could be made between grades. Since the background of the SEM image could affect the learning and prediction of the CNN model, plain carbon tape, which is commonly used in SEM photography, was used in this study.

Fig. 2

Table 1

3.2. Construction of a CNN model for classifying the 10 excipients

To optimize the CNN model parameters, the classification accuracy on the test set of 20 types of CNN models built with different batch sizes and learning rates was evaluated in each pretrained model (Fig. 3). The classification accuracy showed widely varying values depending on batch size and learning rate. The classification accuracy was extremely poor when the learning rate was high, suggesting that the relationship between batch size and classification accuracy is nonlinear. Batch size had little effect on classification accuracy, but when the learning rate was high on ResNet50, classification accuracy increased with lower batch size. The highest classification accuracy was found at a batch size of 48 and a learning rate of 1×10^{-3} when the pretrained model was VGG16. On the other hand, for a ResNet50, a model with a batch size of 16 and a learning rate of 1×10^{-2} had the highest classification accuracy. Therefore, subsequent studies focused on the best models for this condition.

Fig. 3

A typical example of the training progress of the CNN models is shown in Fig. S3 (Supplementary Material), including classification accuracy and loss for each epoch of training and validation. The results show that the accuracy of a training set improves with increasing epochs and converges without significant change after about the 20th epoch, at most. Similarly, the model loss of training set decreased with increasing epochs and did not change significantly after the 20th epoch. By contrast, the validation set required more epochs to converge than the training set due to greater variability in model accuracy and model loss. Model accuracy and model loss were seen to vary widely in some cases up to about the 40th epoch. In some cases, the model accuracy improved slightly after about the 70th epoch (Fig. S3b). Finally, the CNN model converged with high classification accuracy and low loss on all datasets, indicating that it can learn features to distinguish between excipient types. Furthermore, the high accuracy of classification not only on the training set

but also on the validation set suggests that the CNN model is well trained, with little overfitting.

3.3. Evaluation of the constructed CNN model using the test set

To assess the accuracy of the model in more detail quantitatively, we used a confusion matrix and four statistical indices: classification accuracy, precision, reproducibility, and F-measure (Figs. 4 and 5). As for both pretraining models, all four statistical indices were higher than 0.860, indicating that the 10 excipients could be classified with high accuracy. These results also suggest that our CNN model is capable of automatically extracting particle features from SEM images. In other words, we found that the CNN can automatically extract the size, shape, and surface condition of particles. In fact, because Pharmatose 100M, 125M, and 200M are used as grades with different particle sizes, the fact that they could be classified suggests that CNN can evaluate differences in particle size. Moreover, the nominal mass median diameter of CEOLUS PH-101, KG-802, and KG-1000 are all equally 50 μm (Horio et al., 2014), suggesting that CNNs can also recognize information about particle shape and surface conditions.

Fig. 4

Fig. 5

Focusing on the classification accuracy of individual excipients, the CEOLUS series had slightly lower classification accuracy than the other excipients (Figs. 4 and 5). The results of the confusion matrix show that CEOLUS grades are often mistaken for the other CEOLUS grades although they are not mistaken for other excipients (Fig. 4). In particular, the classification accuracy of CEOLUS PH-101 and KG-802 was lower than that for CEOLUS PH-302 and KG-1000 for both models. KG-1000 has the most elongated filamentous particles, whereas PH-302 features the most highly rounded particles of the CEOLUS grades.

In contrast, CEOLUS PH-101 and KG-802 are intermediate grades, so it is assumed that they were more easily mistaken. This is partly because CEOLUS is a mixture of elongated and round particles, and the particle morphology is less uniform. In other words, if the particles are biased in the split image, the likelihood of misclassification will increase. For example, in the case of CEOLUS KG-802, if there happen to be many elongated string-like particles in the split image, the probability of misclassification to KG-1000 will increase. To prevent this, it is necessary to adjust the measurement magnification and the number of segmented pixels so that more particles are present in the segmented image.

For both pretraining models, the mean value of the F-measure was 0.970, indicating that the classification accuracy was comparable. ResNet50 has deeper layers than VGG16, and ResNet50 has higher classification accuracy for the ImageNet classification dataset (He et al., 2016). In both cases, training set converged in relatively few epochs on our dataset, suggesting that the task was too simple for either model to show any difference in classification accuracy. In the future, when the task is made more difficult by increasing the number of SEM images of excipients, the classification accuracy will likely differ depending on the pretrained model.

3.4. Visualization of particle features used in the CNN classification

Grad-CAM is one of the algorithms that visualize features of CNN models and has been developed for interpreting the deep learning classification (Selvaraju et al., 2020). It can highlight particular image regions that provide meaningful information for model prediction. This algorithm helps us to discover the features that deep learning used, which means we can know whether deep learning used detailed features, and it can provide understandable visual information about model performance.

Fig. 6 shows the attention map visualized by Grad-CAM. The Grad-CAM analysis showed that in most cases, the CNN model appropriately focused on the regions in the image where particles were present, rather than on the regions where no particles were present. The exception to this was Pharmatose 100M, where the model using ResNet50 gave greater importance to regions without particles. The details are unknown, although they appear to indicate the contours of the particles. The results of the Grad-CAM analysis also showed that there were differences in the degree of importance among the particles. For instance, Mannit Q appeared to focus on smaller spherical particles rather than larger spherical particles. In addition, the critical areas of the image differed slightly depending on the pretraining model. For example, focusing on Pharmatose 200M, particles near the center were more important in VGG16, whereas those in the lower right were more important in ResNet50. The difference in Grad-CAM can be attributed to differences in the model structure. For example, VGG16 has a structure consisting of 16 layers, with successive convolutional layers with small filters, and the filters are repeatedly reduced to half their size by pooling layers (Simonyan and Zisserman, 2015). By contrast, ResNet50 consists of 50 layers deeper than VGG16 and has a structure called skip connection that adds a given signal to the output of a layer slightly higher than it (He et al., 2016). Therefore, because the internal structure of the CNN models of VGG16 and ResNet50 is very different, it is likely that the important image regions will also be observed to be different in some cases.

Fig. 6

Our study is the first to show that Grad-CAM can visualize important parts of the image, even in SEM images of excipient particles. How the CNN model classifies images is a black box that is difficult for humans to understand. Even so, Grad-CAM is a useful tool because it can show the criteria the CNN model uses to classify images in a way that humans can comprehend, at least partially. However, even with Grad-CAM it was difficult for us to fully

understand which particles were of particular importance. To clarify this, it is necessary to quantify the size, shape, and surface properties of individual particles and evaluate their correlation with importance, which is considered a future task.

4. Conclusions

In this study, a CNN was introduced to classify SEM images of pharmaceutical raw material powders. Specifically, a method of VGG16 or ResNet50 with transfer learning was proposed to classify 10 different excipients with different particle morphology automatically. We demonstrate that CNN achieved good performance with high classification accuracy, indicating that CNN can detect differences in particle morphology such as particle size, shape, and surface properties. Moreover, Grad-CAM analysis enabled us to visualize particle features as good classifiers for the excipients. In summary, the CNN has been found to have the potential to be applied to the identification and characterization of raw material powders for pharmaceutical development.

Acknowledgment

This paper is based on partial results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: Large-Scale

Machine Learning on Heterogeneous Distributed Systems. arXiv e-prints. https://doi.org/arxiv.org/abs/1603.04467.

Abbas, Q., Ibrahim, M.E.A., Jaffar, M.A., 2019. A comprehensive review of recent advances on deep vision systems. Artif. Intell. Rev. 52, 39–76. https://doi.org/10.1007/s10462-018-9633-3

Alyami, H., Dahmash, E., Bowen, J., Mohammed, A.R., 2017. An investigation into the effects of excipient particle size, blending techniques & processing parameters on the homogeneity & content uniformity of a blend containing low-dose model drug. PLoS One 12, 1–19. https://doi.org/10.1371/journal.pone.0178772

Azimi, S.M., Britz, D., Engstler, M., Fritz, M., Mücklich, F., 2018. Advanced steel microstructural classification by deep learning methods. Sci. Rep. 8, 1–14. https://doi.org/10.1038/s41598-018-20037-5

Calderon, C.P., Daniels, A.L., Randolph, T.W., 2018. Deep Convolutional Neural Network Analysis of Flow Imaging Microscopy Data to Classify Subvisible Particles in Protein Formulations. J. Pharm. Sci. 107, 999–1008. https://doi.org/10.1016/j.xphs.2017.12.008

Desai, P.M., Liew, C.V., Heng, P.W.S., 2016. Review of disintegrants and the disintegration phenomena. J. Pharm. Sci. 105, 2545–2555. https://doi.org/10.1016/j.xphs.2015.12.019

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. 27, 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Ficzere, M., Mészáros, L.A., Kállai-Szabó, N., Kovács, A., Antal, I., Nagy, Z.K., Galata, D.L., 2022. Real-time coating thickness measurement and defect recognition of film coated tablets with machine vision and deep learning. Int. J. Pharm. 623, 121957. https://doi.org/10.1016/j.ijpharm.2022.121957

Gambe-Gilbuena, A., Shibano, Y., Krayukhina, E., Torisu, T., Uchiyama, S., 2020. Automatic Identification of the Stress Sources of Protein Aggregates Using Flow

Imaging Microscopy Images. J. Pharm. Sci. 109, 614–623. https://doi.org/10.1016/j.xphs.2019.10.034

Ge, M., Su, F., Zhao, Z., Su, D., 2020. Deep learning analysis on microscopic imaging in materials science. Mater. Today Nano 11, 100087. https://doi.org/10.1016/j.mtnano.2020.100087

Hamishehkar, H., Emami, J., Najafabadi, A.R., Gilani, K., Minaiyan, M., Mahdavi, H., Nokhodchi, A., 2010. Effect of carrier morphology and surface characteristics on the development of respirable PLGA microcapsules for sustained-release pulmonary delivery of insulin. Int. J. Pharm. 389, 74–85. https://doi.org/10.1016/j.ijpharm.2010.01.021

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-Decem, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hirschberg, C., Edinger, M., Holmfred, E., Rantanen, J., Boetker, J., 2020. Image-based artificial intelligence methods for product control of tablet coating quality. Pharmaceutics 12, 1–9. https://doi.org/10.3390/pharmaceutics12090877

Holgado, M.A., Fernández-Hervás, M.J., Alvarez-Fuentes, J., Vela, M.T., Rabasco, A.M., Fini, A., 1996. Characterization of modified paracetamol by means of SEM and fractal analysis. Int. J. Pharm. 142, 143–151. https://doi.org/10.1016/0378-5173(96)04659-5

Horio, T., Yasuda, M., Matsusaka, S., 2014. Effect of particle shape on powder flowability of microcrystalline cellulose as determined using the vibration shear tube method. Int. J. Pharm. 473, 572–578. https://doi.org/10.1016/j.ijpharm.2014.07.040

Ienaga, N., Higuchi, K., Takashi, T., Gen, K., Tsuda, K., Terayama, K., 2021. Vision-based egg quality prediction in Pacific bluefin tuna (Thunnus orientalis) by deep neural network. Sci. Rep. 11, 1–10. https://doi.org/10.1038/s41598-020-80001-0

Janssen, P.H.M., Depaifve, S., Neveu, A., Francqui, F., Dickhoff, B.H.J., 2021. Impact of powder properties on the rheological behavior of excipients. Pharmaceutics 13, 1198. https://doi.org/10.3390/pharmaceutics13081198

Kudo, Y., Yasuda, M., Matsusaka, S., 2020. Effect of particle size distribution on flowability of granulated lactose. Adv. Powder Technol. 31, 121–127. https://doi.org/10.1016/j.apt.2019.10.004

Lu, S., Lu, Z., Zhang, Y.D., 2019. Pathological brain detection based on AlexNet and transfer learning. J. Comput. Sci. 30, 41–47. https://doi.org/10.1016/j.jocs.2018.11.008

Lu, T., Yu, F., Xue, C., Han, B., 2021. Identification, classification, and quantification of three physical mechanisms in oil-in-water emulsions using AlexNet with transfer learning. J. Food Eng. 288, 110220. https://doi.org/10.1016/j.jfoodeng.2020.110220

Ma, X., Kittikunakorn, N., Sorman, B., Xi, H., Chen, A., Marsh, M., Mongeau, A., Piché, N., Williams, R.O., Skomski, D., 2020. Application of Deep Learning Convolutional Neural Networks for Internal Tablet Defect Detection: High Accuracy, Throughput, and Adaptability. J. Pharm. Sci. 109, 1547–1557. https://doi.org/10.1016/j.xphs.2020.01.014

Modarres, M.H., Aversa, R., Cozzini, S., Ciancio, R., Leto, A., Brandino, G.P., 2017. Neural Network for Nanoscience Scanning Electron Microscope Image Recognition. Sci. Rep. 7, 1–12. https://doi.org/10.1038/s41598-017-13565-z

Morid, M.A., Borjali, A., Del Fiol, G., 2021. A scoping review of transfer learning research on medical image analysis using ImageNet. Comput. Biol. Med. 128, 104115. https://doi.org/10.1016/j.compbiomed.2020.104115

Myshkin, N.K., Grigoriev, A.Y., Kholodilov, O. V., 1992. Quantitative analysis of surface topography using scanning electron microscopy. Wear 153, 119–133. https://doi.org/10.1016/0043-1648(92)90265-A

Nagao, Y., Sakamoto, M., Chinen, T., Okada, Y., Takao, D., 2020. Robust classification of cell cycle phase and biological feature extraction by image-based deep learning. Mol. Biol. Cell 31, 1346–1354. https://doi.org/10.1091/MBC.E20-03-0187

Onishi, M., Ise, T., 2021. Explainable identification and mapping of trees using UAV RGB image and deep learning. Sci. Rep. 1–15. https://doi.org/10.1038/s41598-020-79653-9

Passerini, N., Albertini, B., González-Rodríguez, M.L., Cavallari, C., Rodriguez, L., 2002. Preparation and characterisation of ibuprofen-poloxamer 188 granules obtained by melt granulation. Eur. J. Pharm. Sci. 15, 71–78. https://doi.org/10.1016/S0928-0987(01)00210-X

Paul, S., Chang, S.Y., Dun, J., Sun, W.J., Wang, K., Tajarobi, P., Boissier, C., Sun, C.C., 2018. Comparative analyses of flow and compaction properties of diverse mannitol and lactose grades. Int. J. Pharm. 546, 39–49. https://doi.org/10.1016/j.ijpharm.2018.04.058

Probst, C., Zayats, A., Venkatachalam, V., Davidson, B., 2020. Advanced Characterization of Silicone Oil Droplets in Protein Therapeutics Using Artificial Intelligence Analysis of Imaging Flow Cytometry Data. J. Pharm. Sci. 109, 2996–3005. https://doi.org/10.1016/j.xphs.2020.07.008

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet Large Scale Visual Recognition Challenge. Int. J. Comput. Vis. 115, 211–252. https://doi.org/10.1007/s11263-015-0816-y

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2020. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Int. J. Comput. Vis. 128, 336–359. https://doi.org/10.1007/s11263-019-01228-7

Shekunov, B.Y., Chattopadhyay, P., Tong, H.H.Y., Chow, A.H.L., 2007. Particle size analysis in pharmaceutics: Principles, methods and applications. Pharm. Res. 24, 203–227. https://doi.org/10.1007/s11095-006-9146-7

Silva, A.F.T., Burggraeve, A., Denon, Q., Van Der Meeren, P., Sandler, N., Van Den Kerkhof, T., Hellings, M., Vervaet, C., Remon, J.P., Lopes, J.A., De Beer, T., 2013. Particle sizing measurements in pharmaceutical applications: Comparison of in-process methods versus off-line methods. Eur. J. Pharm. Biopharm. 85, 1006–1018. https://doi.org/10.1016/j.ejpb.2013.03.032

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. 1–14.

Thoorens, G., Krier, F., Leclercq, B., Carlin, B., Evrard, B., 2014. Microcrystalline cellulose, a direct compression binder in a quality by design environment - A review. Int. J. Pharm. 473, 64–72. https://doi.org/10.1016/j.ijpharm.2014.06.055

Thoorens, G., Krier, F., Rozet, E., Carlin, B., Evrard, B., 2015. Understanding the impact of microcrystalline cellulose physicochemical properties on tabletability. Int. J. Pharm. 490, 47–54. https://doi.org/10.1016/j.ijpharm.2015.05.026

Wang, T., Alston, K.M., Wassgren, C.R., Mockus, L., Catlin, A.C., Fernando, S.R., Fernando, S., Basu, P.K., Hoag, S.W., 2013. The creation of an excipient properties database to support quality by design (QbD) formulation development. Am. Pharm. Rev. 16.

Yu, F., Lu, T., Han, B., Xue, C., 2021. A quantitative study of aggregation behaviour and integrity of spray-dried microcapsules using three deep convolutional neural networks with transfer learning. J. Food Eng. 300, 110515. https://doi.org/10.1016/j.jfoodeng.2021.110515

Tables

Table 1

Characteristics of the ten model excipients.

Excipients	Component	Particle shapes
CEOLUS KG-802	Microcrystalline cellulose	Fibrous elongated
CEOLUS KG-1000		
CEOLUS PH-101		Fibrous elongated and round
CEOLUS PH-302		
Dried corn starch	Corn starch	Nearly spherical but with angles
Mannit Q	Mannitol	Spherical
PEARLITOL 200SD		Round but rugged shape
Pharmatose 100M	Lactose monohydrate	Angular and slightly elongated in shape
Pharmatose 125M		
Pharmatose 200M		

Figure Legends

Fig. 1. Flow chart of this research.

Fig. 2. Original SEM images of the 10 model excipients.

Fig. 3. Relationship between CNN model parameters and classification accuracy. Shaded conditions indicate optimal conditions.

Fig. 4. Confusion matrix of the test set showing the classification accuracy of each excipient: (a) VGG16 and (b) ResNet50.

Fig. 5. Statistical indices for each excipient in each pretrained model.

Fig. 6. Typical example of an image, visualizing how important it is to classify excipients using Grad-CAM. For each excipient, the original SEM image and the use of VGG16 and ResNet50 as pretrained models are shown.

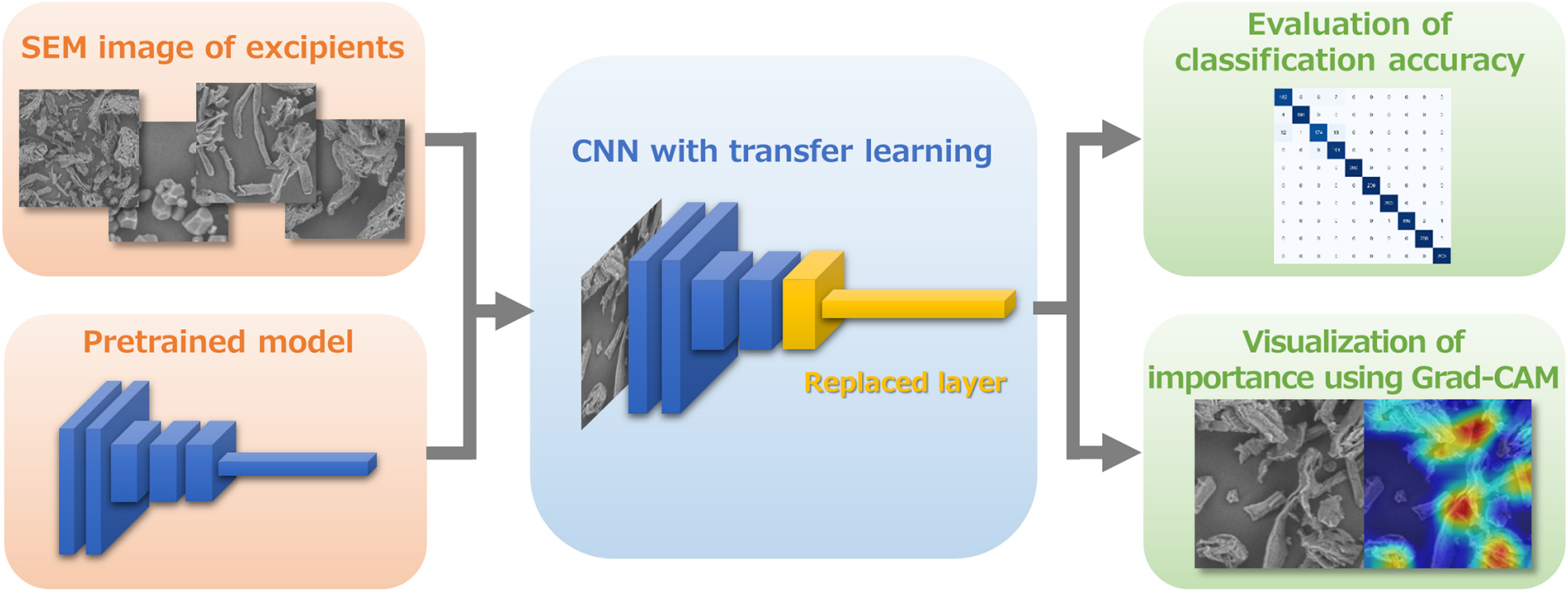
Graphical abstract

Conflict of interest statement

The authors declare no conflict of interest.

CRediT author statement

Hiroaki Iwata: Methodology, Software, Validation, Investigation, Formal analysis, Resources, Data Curation, Writing - Original Draft. **Yoshihiro Hayashi:** Conceptualization, Validation, Investigation, Formal analysis, Resources, Data Curation, Writing - Original Draft. **Aki Hasegawa:** Validation, Investigation, Formal analysis, Writing - Review & Editing. **Kei Terayama:** Writing - Review & Editing. **Yasushi Okuno:** Supervision, Project administration, Funding acquisition.



Graphics Abstract

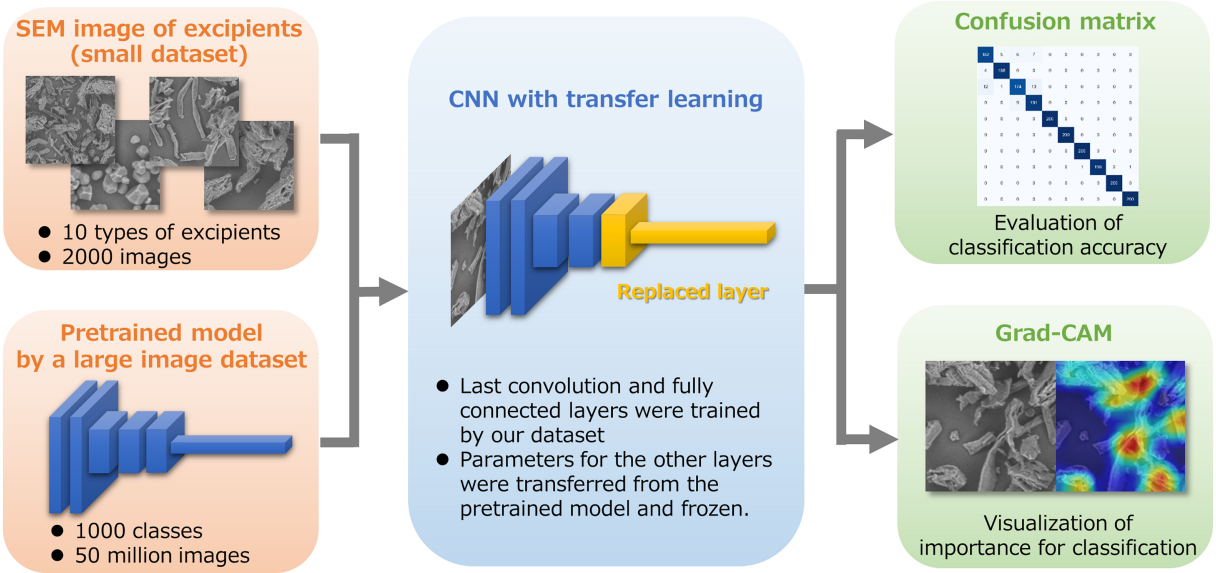

SEM image of excipients (small dataset)
10 types of excipients
2000 images
Pretrained model by a large image dataset
1000 classes
50 million images
CNN with transfer learning
Replaced layer
Last convolution and fully connected layers were trained by our dataset
Parameters for the other layers were transferred from the pretrained model and frozen.
Confusion matrix
Evaluation of classification accuracy
Grad-CAM
Visualization of importance for classification

Figure 1

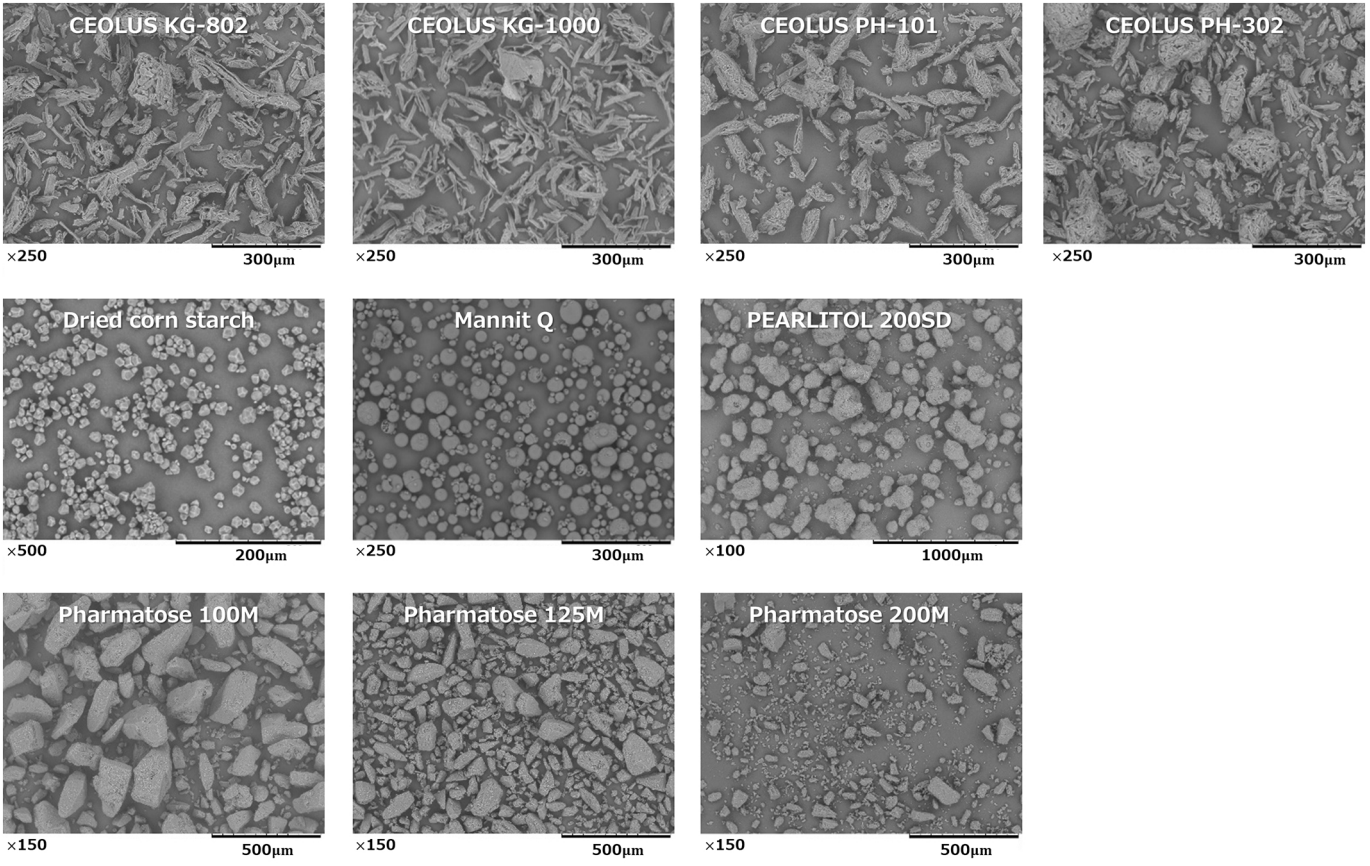

CEOLUS KG-802
×250
300μm
CEOLUS KG-1000
×250
300μm
CEOLUS PH-101
×250
300μm
CEOLUS PH-302
×250
300μm
Dried corn starch
×500
200μm
Mannit Q
×250
300μm
PEARLITOL 200SD
×100
1000μm
Pharmatose 100M
×150
500μm
Pharmatose 125M
×150
500μm
Pharmatose 200M
×150
500μm

Figure 2

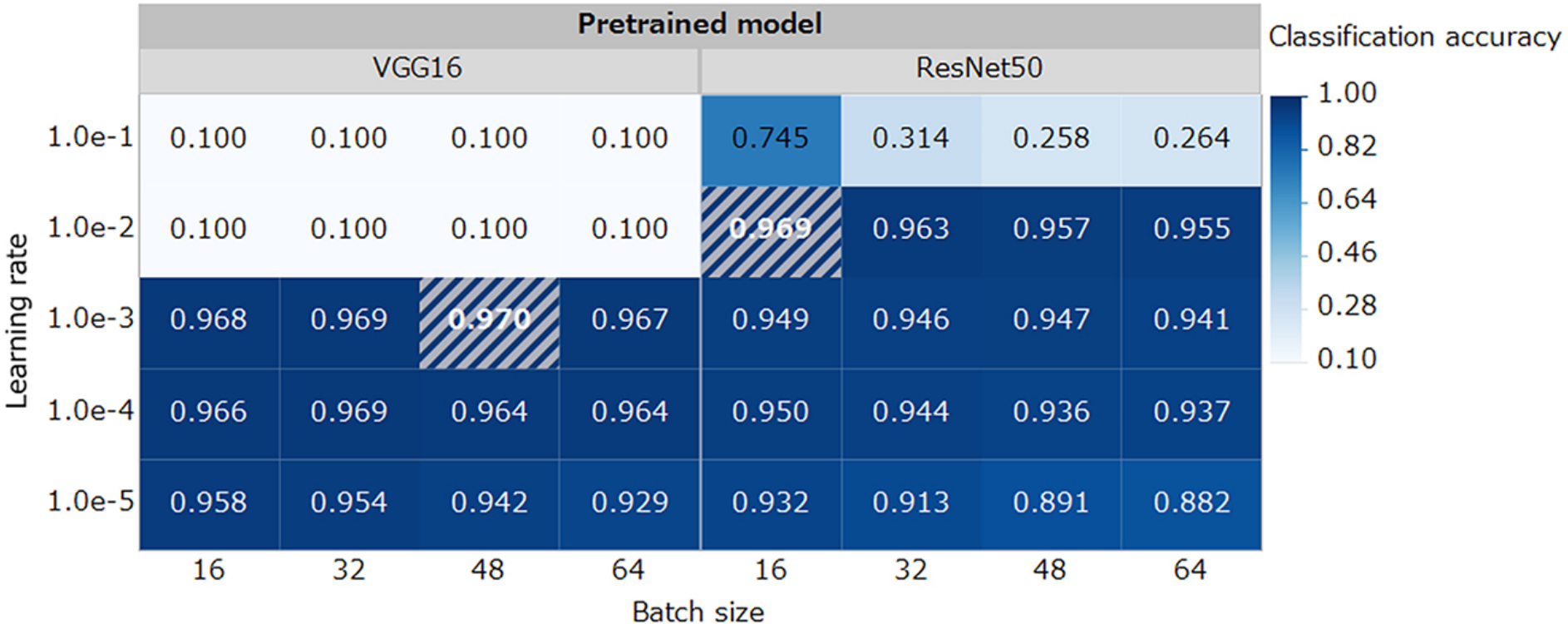

Pretrained model
VGG16
ResNet50
Classification accuracy
1.00
0.82
0.64
0.46
0.28
0.10
Learning rate
1.0e-1
1.0e-2
1.0e-3
1.0e-4
1.0e-5
0.100 0.100 0.100 0.100 0.745 0.314 0.258 0.264
0.100 0.100 0.100 0.100 0.969 0.963 0.957 0.955
0.968 0.969 0.970 0.967 0.949 0.946 0.947 0.941
0.966 0.969 0.964 0.964 0.950 0.944 0.936 0.937
0.958 0.954 0.942 0.929 0.932 0.913 0.891 0.882
16 32 48 64 16 32 48 64
Batch size

Figure 3

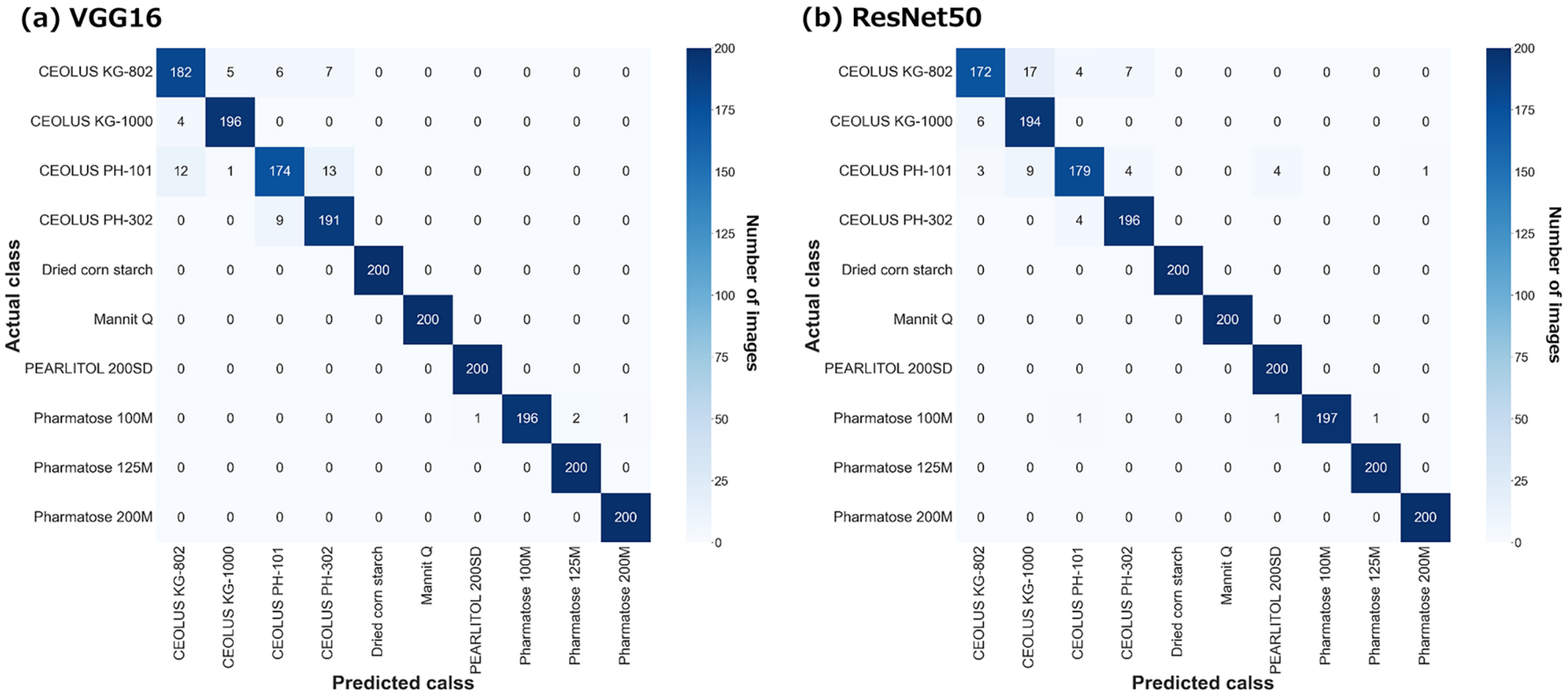

(a) VGG16
CEOLUS KG-802 182 5 6 7 0 0 0 0 0 0
CEOLUS KG-1000 4 196 0 0 0 0 0 0 0 0
CEOLUS PH-101 12 1 174 13 0 0 0 0 0 0
CEOLUS PH-302 0 0 9 191 0 0 0 0 0 0
Dried corn starch 0 0 0 0 200 0 0 0 0 0
Mannit Q 0 0 0 0 0 200 0 0 0 0
PEARLITOL 200SD 0 0 0 0 0 0 200 0 0 0
Pharmatose 100M 0 0 0 0 0 0 1 196 2 1
Pharmatose 125M 0 0 0 0 0 0 0 0 200 0
Pharmatose 200M 0 0 0 0 0 0 0 0 0 200
Actual class
CEOLUS KG-802
CEOLUS KG-1000
CEOLUS PH-101
CEOLUS PH-302
Dried corn starch
Mannit Q
PEARLITOL 200SD
Pharmatose 100M
Pharmatose 125M
Pharmatose 200M
Predicted calss
Number of Images
200
175
150
125
100
75
50
25
0
(b) ResNet50
CEOLUS KG-802 172 17 4 7 0 0 0 0 0 0
CEOLUS KG-1000 6 194 0 0 0 0 0 0 0 0
CEOLUS PH-101 3 9 179 4 0 0 4 0 0 1
CEOLUS PH-302 0 0 4 196 0 0 0 0 0 0
Dried corn starch 0 0 0 0 200 0 0 0 0 0
Mannit Q 0 0 0 0 0 200 0 0 0 0
PEARLITOL 200SD 0 0 0 0 0 0 200 0 0 0
Pharmatose 100M 0 0 1 0 0 0 1 197 1 0
Pharmatose 125M 0 0 0 0 0 0 0 0 200 0
Pharmatose 200M 0 0 0 0 0 0 0 0 0 200
Actual class
CEOLUS KG-802
CEOLUS KG-1000
CEOLUS PH-101
CEOLUS PH-302
Dried corn starch
Mannit Q
PEARLITOL 200SD
Pharmatose 100M
Pharmatose 125M
Pharmatose 200M
Predicted calss
Number of Images
200
175
150
125
100
75
50
25
0

Figure 4

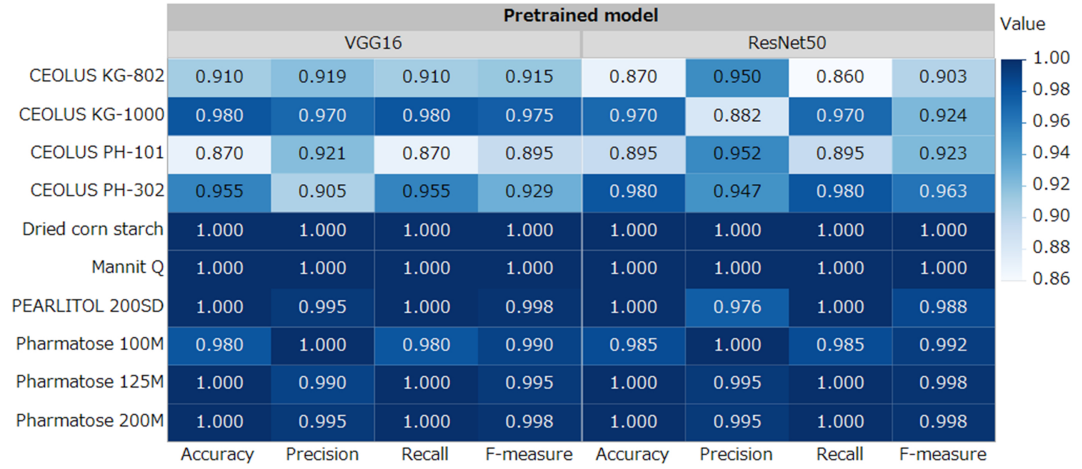

Pretrained model
VGG16
ResNet50
Value
1.00
0.98
0.96
0.94
0.92
0.90
0.88
0.86
CEOLUS KG-802 0.910 0.919 0.910 0.915 0.870 0.950 0.860 0.903
CEOLUS KG-1000 0.980 0.970 0.980 0.975 0.970 0.882 0.970 0.924
CEOLUS PH-101 0.870 0.921 0.870 0.895 0.895 0.952 0.895 0.923
CEOLUS PH-302 0.955 0.905 0.955 0.929 0.980 0.947 0.980 0.963
Dried corn starch 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000
Mannit Q 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000
PEARLITOL 200SD 1.000 0.995 1.000 0.998 1.000 0.976 1.000 0.988
Pharmatose 100M 0.980 1.000 0.980 0.990 0.985 1.000 0.985 0.992
Pharmatose 125M 1.000 0.990 1.000 0.995 1.000 0.995 1.000 0.998
Pharmatose 200M 1.000 0.995 1.000 0.998 1.000 0.995 1.000 0.998
Accuracy Precision Recall F-measure Accuracy Precision Recall F-measure

Figure 5

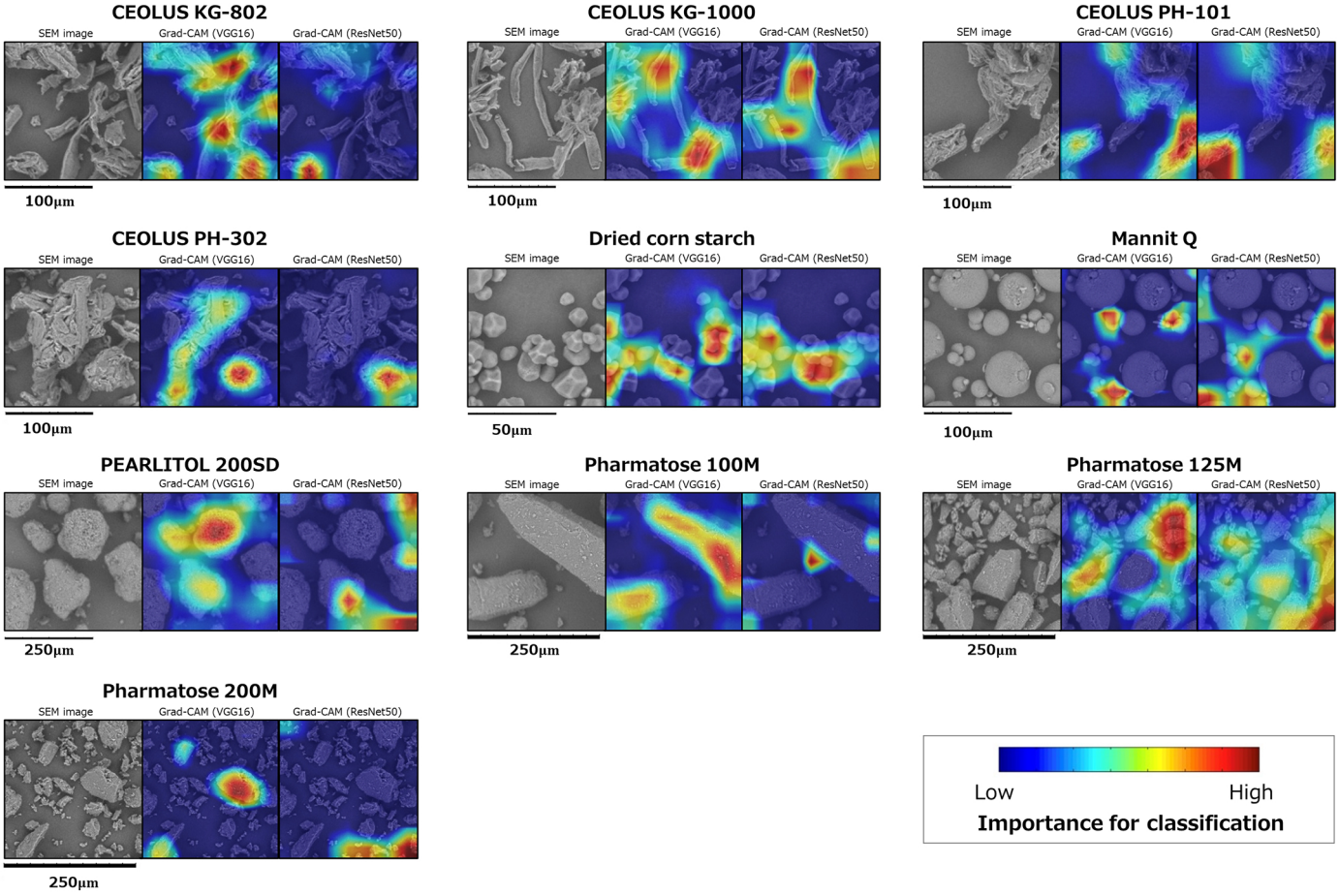

CEOLUS KG-802
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
100μm
CEOLUS KG-1000
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
100μm
CEOLUS PH-101
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
100μm
CEOLUS PH-302
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
100μm
Dried corn starch
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
50μm
Mannit Q
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
100μm
PEARLITOL 200SD
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
250μm
Pharmatose 100M
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
250μm
Pharmatose 125M
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
250μm
Pharmatose 200M
SEM image
Grad-CAM (VGG16)
Grad-CAM (ResNet50)
250μm
Low
High
Importance for classification

Figure 6