T1030-12-68

Applications of Deep Learning to Predict Aerosol Performance of Dry Powders for Inhalation

Junhuang Jiang [a], Han-Hsuan Peng [a], Zhenpei Yang [b], Xiangyu Ma [c], Sawittree Sahakijpijarn [d], Chaeho Moon [a], Defang Ouyang [e] and Robert O. Williams III [a,*]

[a] Department of Molecular Pharmaceutics and Drug Delivery, College of Pharmacy, The University of Texas at Austin, TX, US
[b] Department of Computer Science, The University of Texas at Austin, TX, USA
[c] Global Investment Research, Goldman Sachs, NY, USA
[d] TFF Pharmaceuticals, Inc., Austin, TX, USA
[e] State Key Laboratory of Quality Research in Chinese Medicine, Institute of Chinese Medical Sciences (ICMS), University of Macau, Macau, China

CONTACT INFORMATION: bill.williams@austin.utexas.edu

aaps PharmSci 360

PURPOSE

Background and motivation of this study

1. Dry powder inhalers (DPI) are widely used devices to treat respiratory diseases.
2. Aerosol performance of dry powders for inhalation is important for the product development process and is typically characterized by fine particle fraction (FPF) and mass median aerodynamic diameter (MMAD).
3. However, a conventional trial-and-error method for formulation screening and optimization is time-consuming and laborious.

OBJECTIVE(S)

We propose to develop a deep learning-based method to estimate the aerosol performance based on the properties of the powder made by thin-film freezing, specifically using image analysis.

METHOD(S)

- **Data collection and labeling**
  1. 132 raw SEM images obtained from published literature and additional experiments were used for training.
  2. Image processing techniques (i.e., image augmentation and CLAHE (**Fig. 1**))
  3. SEM images were manually classified into "acceptable" formulations with recovered FPF ≥ 60% and "unacceptable" formulations with recovered FPF < 60%.
- **Convolutional neural networks (CNN)** (**Fig. 2**)
- **Model evaluation**
  1. 316 resized SEM were used for testing
  2. Evaluation metrics (accuracy and receiver operating characteristic (ROC) aera under curve (AUC))

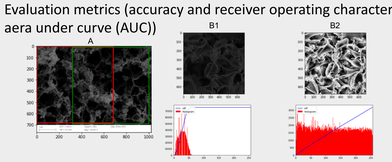

Figure 1. Image processing techniques: Resize (left) and CLAHE (right)

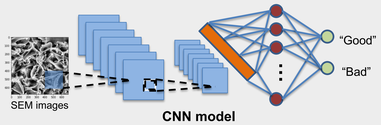

Figure 2. CNN model for image classification

RESULT(S)

Image processing results

- After image augmentation, a balanced dataset containing 4266 SEM images was obtained.
- CLAHE has been demonstrated to greatly improve the image contrast.
- The CNN model applying CLAHE showed higher accuracies and lower loss values than the CNN model not applying CLAHE in both training and validation subsets (**Fig. 3**).

CNN model prediction results

- Overall, the CNN model performed well with relatively high accuracy and ROC AUC values of 0.839 and 0.927, respectively (**Fig. 4**).
- In addition, to further understand the CNN model performance, we split the dataset by different drug formulations and evaluated their accuracies (**Table 1**). The tacrolimus, remdesivir, and voriconazole TFF formulations have accuracies of 0.948, 0.730, and 0.821, respectively (**Table 1**).

CNN model visualization

- Feature map, which contains the images captured by applying the filters to the input images was applied to visualize the CNN model (**Fig. 5**).
- The first and second convolutional layer feature maps were displayed similarly as compared to the original SEM images. In contrast, the last feature maps learned distinct activation for the two classes (**Fig. 5**).
- the CNN model visualization by feature maps showed differences in the internal structure of "acceptable" and "unacceptable" dry powder.

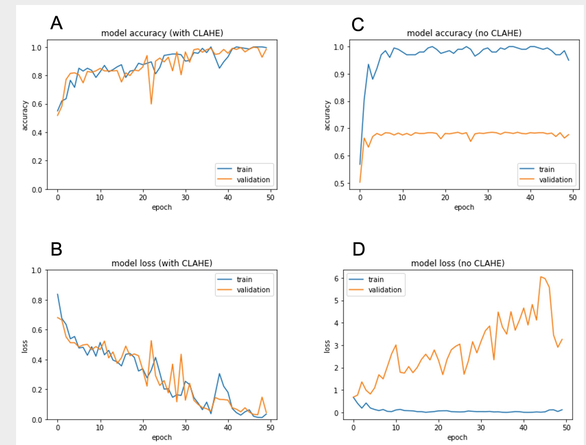

***Figure 3.** (A)The training and validation process of the CNN model. CNN accuracy results for images processed by CLAHE. (B) CNN loss results for images processed by CLAHE. (C) CNN accuracy results for images without CLAHE processing. (D) CNN loss results for images without CLAHE processing.*

TFF formulations	Total SEM images	True positive	True negative	False positive	False negative	Accuracy
Tacrolimus	134	102	25	7	0	0.948
Remdesivir	126	46	46	18	16	0.730
Voriconazole	56	0	46	10	0	0.821
Total	316	148	117	35	16	0.839

***Table 1.** True positive, true negative, false positive, false negative, and accuracies results of three TFF drug formulations (i.e., tacrolimus, remdesivir, and voriconazole) in the testing set.*

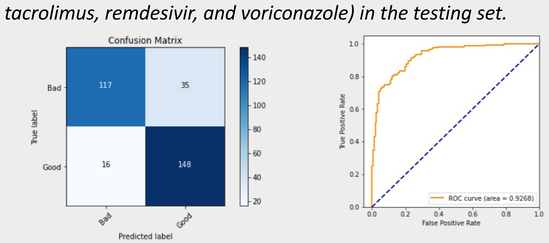

***Figure 4.** Predictive performance of the CNN model. Confusion matrix (Left) and ROC AUC (Right). The overall accuracy result (0.839) was calculated based on confusion matrix.*

CONCLUSION(S)

1. In this study, we have successfully developed a deep-learning-based aerosol performance analyzing method and achieved high accuracy and adaptability.
2. Most importantly, this study can be used to reduce the workload of conventional experimental analysis.

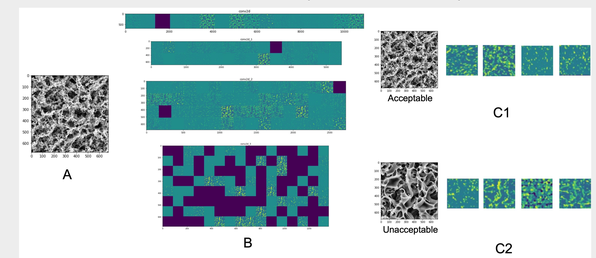

***Figure 5.** CNN model visualization by feature maps. Processed SEM image (A). CNN model feature maps in first (top) to fourth convolutional layer (bottom) (B). Feature maps in the last convolutional layers for "Acceptable" TFF formulation (C1) and "Unacceptable" formulation (C2).*

FUNDING

This study was partially supported by a sponsored research agreement with TFF Pharmaceuticals, Inc.

DECLARATION OF COMPETING INTEREST

The University of Texas System has licensed IP on thin-film freezing to TFF Pharmaceuticals, Inc. Williams reports financial support from TFF Pharmaceuticals. Williams reports a relationship with TFF Pharmaceuticals, Inc. that includes consulting or advisory, equity or stocks, and research funding. Moon reports a relationship with TFF Pharmaceuticals, Inc. that includes consulting or advisory.

REFERENCES

This poster was published as part of the paper [1].

1. Jiang, Junhuang, et al. "The Applications of Machine Learning (ML) in Designing Dry Powder for Inhalation by Using Thin-film-freezing Technology." International Journal of Pharmaceutics (2022): 122179.

TEXAS
The University of Texas at Austin
College of Pharmacy