Exploring X-ray Velocimetry (XV) imaging and Machine Learning for the detection and monitoring of cystic fibrosis disease in preclinical data

Chan Ka Hing
The University of Adelaide
kahing.chan@student.adelaide.edu.au

Viet Bach Tran
The University of Adelaide
vietbach.tran@student.adelaide.edu.au

Abstract

The purpose of this project is to explore the X-ray Velocimetry (XV) imaging in cystic fibrosis (CF) lung disease detection. X-ray velocimetry (XV) is a non-invasive lung function measurement tool based on the dynamics of airflow throughout the lung [1]. However, the utilisation of XV technology for CF detection is under exploration. This study investigated the CF disease through feature selection techniques and simple machine learning classifiers on mouse and rat experiments. These methods show robust and accurate classification results between healthy and CF animals in limited amount of data. These results show the potential of CF detection in human XV data.

1. Introduction

Cystic fibrosis (CF) is a life-limiting genetic condition where abnormal thick and sticky mucus is produced, leading to airways clogging, lung function impairment, and respiratory infections. Assessments of overall lung health in humans and animals are typically made using lung function tests that detect abnormalities by measuring the airflow. X-ray Velocimetry (XV) imaging offers a novel approach by delivering 4D (3D + Time) functional information across the entire lung and throughout the breathing cycle. This approach provides beneficial for monitoring and treating cystic fibrosis disease. However, the quantification of XV data are under exploration. Our project aims to provide machine learning methods to analyse mouse and rats XV data for CF classification.

2. Feature Selection

The X-ray Velocimetry (XV) technology records the specific ventilation in different regions of lungs. Specific Ventilation is defined as the change in volume of that region of lung between peak inspiration and exhalation (breathing all the way out to all the way in), divided by the volume of the region.

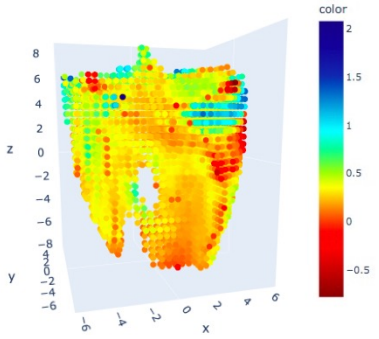

Figure 1. The 3D representation of specific ventilation of lungs. Each point represents the specific ventilation as a number.

2.1. Medical Report Features

The medical features are parameters extracted from specific ventilation for quantifying cystic fibrosis (CF). They include Ventilation Defect Percentage (VDP), Mean Specific Ventilation (MSV), Tidal Volume (TV), Ventilation Heterogeneity (VH), Ventilation Heterogeneity - Small Scale (VHSS) and Ventilation Heterogeneity - Large Scale (VHLS) and Heterogeneous Disease value (HD).

2.2. Histogram Features

The specific ventilation can be also represented by histogram. The histogram analysis generates features such as median, standard deviation, variance, skewness, kurtosis, signal-to-noise ratio, and entropy.

2.3. 3D Texture Features

We utilize two 3D texture features extraction methods namely, Gray-Level Co-occurrence Matrix (GLCM) and Local Binary Pattern (LBP).

2.3.1 Grey Level Co-occurrence Matrix (GLCM)

The Grey Level Co-occurrence Matrix (GLCM) is a statistical technique for analyzing the texture of an image. It in-

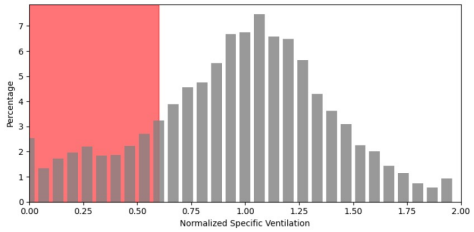

Figure 2. The histogram representation of normalised specific ventilation of lungs.

volves calculating a matrix that represents the frequency of pairs of pixel intensities occurring together at specified spatial relationships in a gray-scale image[2]. Once the GLCM is generated, various texture features, such as dissimilarity, contrast, correlation, energy, homogeneity, are derived from it to describe the texture of the image[2]. We apply GLCM in 3D lung data and extract the above mentioned texture features for analysis.

2.3.2 Local Binary Pattern (LBP)

The Local Binary Pattern (LBP) characterizes the local patterns of pixel intensities in an image by comparing the intensity values of its neighbors against the central pixel's intensity[3]. If the center pixel's value is greater than the neighbor's value, assign "0". Otherwise, assign "1". This gives an 8-digit binary number which is converted to a decimal value, creating a unique representation of the pattern.

3. Machine Learning Classifiers

3.1. Dataset

The datasets contain 4 experiments, including Mouse B-ENaC Study, Mouse MPS Study, Rat PA Study, and Rat Sterile Bead Study. Since the amount of data are limited, we combined the mouse studies and rat studies and created the following datasets and labelled them accordingly for analysis:

- Mouse All (WT vs non-WT) [61 samples]
- Mouse All (WT vs B-ENaC vs MPS) [61 samples]
- Rat All (WT vs non-WT) [73 samples]
- Rat All (WT vs CF vs KO) [73 samples]

3.2. Classifiers

An ensemble machine learning model containing six constituent algorithms [Support Vector Machine (SVM), Decision Tree, K-nearest neighbors (KNN), Random Forest, Gradient Boosting, and Stochastic Gradient Descent (SGD)] was tested for its ability to classify CF in all four studies, denoted by 'ensemble all'. Since some of the classifiers showed over-fitting in most of the studies, we created another ensemble model consisting of SVM, KNN and SGD only, denoted by 'ensemble 3'. The datasets were applied to each individual classifiers, as well as the 'ensemble all' and 'ensemble 3' models.

3.3. Principle Component Analysis (PCA)

In order to identify the important features and remove the redundant features, we performed the Principle Component Analysis (PCA). PCA transforms high-dimensional features into a lower-dimensional features. It outputs the principal components that capture the most variance. We extracted 3 principal components (PC1 to PC3) and 5 principal components (PC1 to PC5), and input them to the classifiers.

3.4. Combination of features

As we have 3 types of features (report, histogram and 3D texture features), we examine different combinations of features to obtain the best result. The combinations of features are as followed:

- Report features
- Histogram features
- GLCM features
- 3D-LBP features
- 2D-LBP (x direction) features
- 2D-LBP (y direction) features
- 2D-LBP (z direction) features
- Report + Histogram features
- Report + GLCM features
- Report + 3D-LBP features
- Histogram + GLCM features
- Histogram + 3D-LBP features
- GLCM + 3D-LBP features
- Report + Histogram + GLCM + 3D-LBP features

4. Experimental result

4.1. Best Model result

The best result is obtained by using the 3D-LBP with PC1 to PC5 as features. Table 1,2,3, and 4 show the training, validation, and testing accuracy for all 4 studies. Figure 3 shows the comparison of testing accuracy among 4 studies by the best model. Comparing the mouse studies with rat studies in figure 3, we observed that the testing accuracy of mouse studies (over 80%) is higher than rat studies (from 50% to 60%).

Table 1. The training, validation, and testing accuracy of various models in Mouse all (WT vs non-WT) study

Mouse all (WT vs non-WT) study 3D LBP with PC1 to PC5			
Models	Training	Validation	Testing
SVM	0.976	0.810	0.789
Decision Tree	1.000	0.786	0.789
KNN	0.940	0.810	0.895
Random Forest	1.000	0.810	0.842
Gradient Boost	1.000	0.786	0.789
SGD	0.940	0.857	0.842
Ensemble all	1.000	0.810	0.895
Ensemble 3	0.976	0.857	0.842

Table 2. The training, validation, and testing accuracy of various models in Mouse all (WT vs B_ENaC vs MPS) study

Mouse all (WT vs B_ENaC vs MPS) study 3D LBP with PC1 to PC5			
Models	Training	Validation	Testing
SVM	0.940	0.857	0.789
Decision Tree	1.000	0.690	0.842
KNN	0.940	0.881	0.789
Random Forest	1.000	0.810	0.895
Gradient Boost	1.000	0.833	0.842
SGD	0.869	0.810	0.579
Ensemble all	0.976	0.857	0.895
Ensemble 3	0.940	0.905	0.842

Table 3. The training, validation, and testing accuracy of various models in Rat all (WT vs non-WT) study

Rat all (WT vs non-WT) study 3D LBP with PC1 to PC5			
Models	Training	Validation	Testing
SVM	0.706	0.549	0.636
Decision Tree	0.931	0.549	0.500
KNN	0.647	0.588	0.455
Random Forest	1.000	0.588	0.455
Gradient Boost	1.000	0.667	0.500
SGD	0.627	0.490	0.591
Ensemble all	0.931	0.608	0.500
Ensemble 3	0.725	0.529	0.591

4.2. Other Evaluation Metrics

Apart from computing the accuracy, we also adopt other evaluation metrics including sensitivity, specificity, precision, and F1-score. Table 5 shows various evaluation metrics for all studies. For mouse studies, both (WT vs non-WT) and (WT vs B_ENaC vs MPS), show outstanding performance in all evaluation metrics, which indicates high true positive and true negative rate, and a high proportion of correct predicted positive cases. While for rat studies, both (WT vs non-WT) and (WT vs KO vs CF), show bad performance in most of the metrics.

Table 4. The training, validation, and testing accuracy of various models in Rat all (WT vs KO vs CF) study

Rat all (WT vs KO vs CF) study 3D LBP with PC1 to PC5			
Models	Training	Validation	Testing
SVM	0.843	0.490	0.545
Decision Tree	0.951	0.333	0.455
KNN	0.627	0.431	0.364
Random Forest	1.000	0.294	0.545
Gradient Boost	1.000	0.373	0.409
SGD	0.598	0.451	0.409
Ensemble all	0.980	0.431	0.500
Ensemble 3	0.745	0.451	0.500

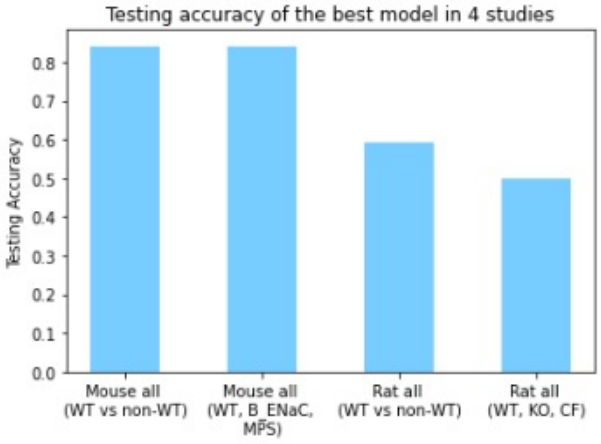

Figure 3. The comparison of testing accuracy between 4 studies by the best model.

Table 5. Evaluation metrics for the testing data in all studies using 3D-LBP (PC1 – PC5) features

Evaluation metrics for the testing data in all studies using 3D-LBP (PC1 – PC5) features					
Studies	Sensitivity	Specificity	Precision	F1-score	Accuracy
Mouse all (WT vs non-WT)	0.778	0.900	0.875	0.824	0.842
Mouse all (WT vs B_ENaC vs MPS)	0.833	0.917	0.873	0.853	0.842
Rat all (WT vs non-WT)	0.857	0.500	0.632	0.727	0.591
Rat all (WT vs KO vs CF)	0.506	0.747	0.514	0.510	0.500

4.3. Principle Component Analysis result

The PCA (PC1 to PC3) visualisation of the 3D-LBP features for all studies are shown in figure 4,5,6 and 7. For the mouse studies, both (WT vs non-WT) and (WT vs B_ENaC vs MPS), the PCA 3D plots show clear separation between clusters of different classes, which corresponds to the classifiers results. While for the rat studies, both (WT vs non-WT) and (WT vs KO vs CF), the clusters of different classes are overlapping with each other, indicating an ambiguous separation between classes based on the 3D-LBP features.

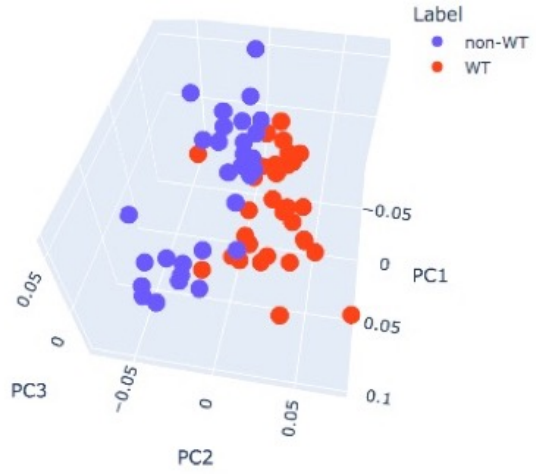

Figure 4. The 3D visualisation of PC1 to PC3 of Mouse all (WT vs non-WT) study using 3D-LBP features.

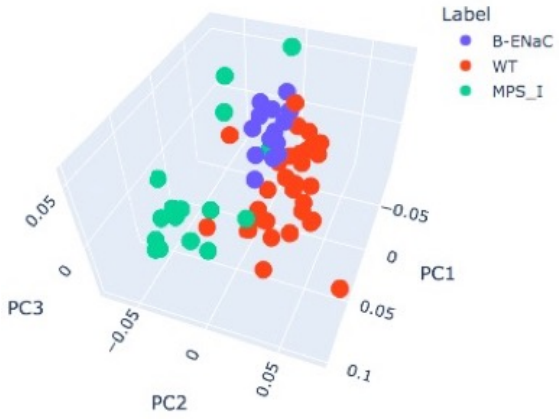

Figure 5. The 3D visualisation of PC1 to PC3 of Mouse all (WT vs B_ENaC vs MPS) study using 3D-LBP features.

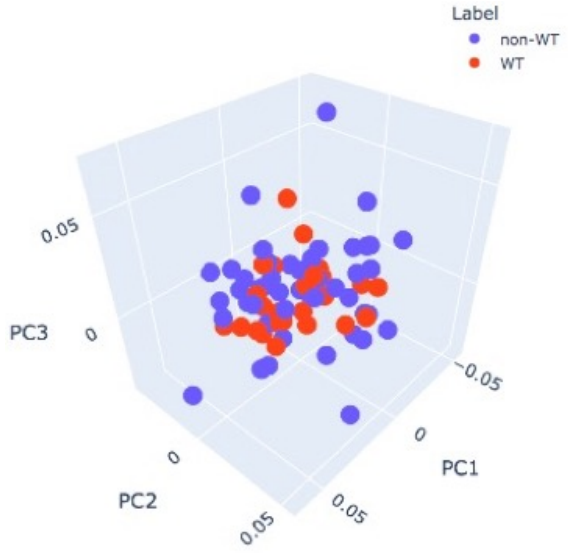

Figure 6. The 3D visualisation of PC1 to PC3 of Rat all (WT vs non-WT) study using 3D-LBP features.

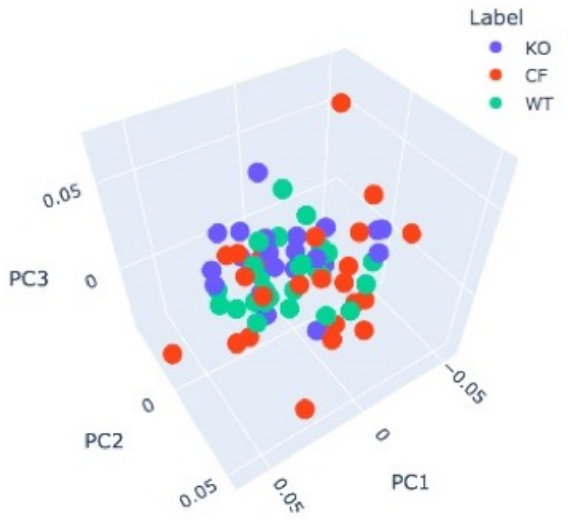

Figure 7. The 3D visualisation of PC1 to PC3 of Rat all (WT vs KO vs CF) study using 3D-LBP features.

4.4. Experimental results of all combinations of features

The experimental results for different combinations of features are shown in figure 9. Generally, the testing accuracy of mouses studies in most of the feature combinations are higher than rat studies. For rat studies, the report features and histogram features give higher accuracy. The report and histogram features captures the global changes, while 3D texture features captures the local pattern. Therefore, rat studies might show more differences in global fea-

tures. Although the (report + 3D-LBP features) and (histogram + 3D-LBP features) have higher testing accuracy than 3D-LBP features, the best model is using 3D-LBP features only because it gives the most robust and consistent result along with high accuracy.

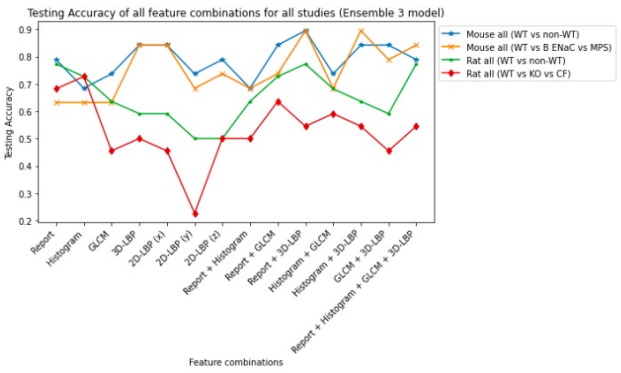

Figure 8. The testing accuracy of all feature combinations for all studies (Ensemble 3 model).

4.5. 2D-LBP results

We divide the 3D lung into 2D slices along x,y,z directions and analyse them separately using Local Binary Pattern (LBP). As shown in table 6, the features extracted from x-direction (sagittal) slices give the best performance in all studies, followed by z-direction (transverse) slices. The y-direction (coronal) slices show the lowest testing accuracy in all studies.

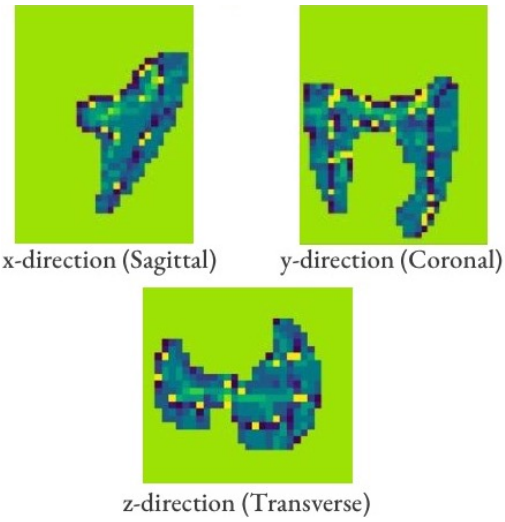

Figure 9. The LBP example of 2D slices in x-direction (Sagittal), y-direction (Coronal) and z-direction (Transverse).

Table 6. The classification testing accuracy of 2D-LBP slices in all 4 studies, considering PC1 to PC5 of the 2D-LBP features.

2D-LBP slices testing accuracy in all 4 studies			
Studies	x (sagittal)	y (coronal)	z (transverse)
Mouse all (WT vs non-WT)	0.842	0.737	0.789
Mouse all (WT vs B_ENaC vs MPS)	0.842	0.684	0.737
Rat all (WT vs non-WT)	0.591	0.500	0.500
Rat all (WT vs KO vs CF)	0.455	0.227	0.500

5. Discussion

From the above experiments, we obtained the following findings:

- The best features selection algorithm is 3D Local Binary Pattern (LBP). It captures the texture features in the whole 3D lungs by comparing each voxel values with its neighbors, regardless of the absolute value of the voxels.
- Our best classification model shows high accuracy and robustness for mouse studies (89.5% testing accuracy). However, it shows over-fitting and low accuracy for rat studies. It might be due to the different amount of information from the data. The mouse studies contain B_ENaC and MPS mouse, which show more obvious symptoms. While rat studies contain KO and CF rats, which show less symptoms in the data. Another reason might be the diversity of ages and sizes in rat datasets, which introduces complexity in the analysis. Lastly, poor experimental performance in measuring the features in rat's lung might also capture less distinct symptoms.
- Among all the classifiers, the 'ensemble 3' model shows the most robust result with high accuracy in all studies. The training, validation, and testing accuracy in 'ensemble 3' model are similar in all studies.
- Considering different combination of features, we observed that the report and histogram features give good performances in rat studies, indicating the importance of global features in analysing rat studies.
- In the 2D-LBP analysis, x-direction (sagittal) slices show the best performance. This might indicate more distinct featuers in sagittal plane. However, further medical analysis is required.

6. Code

https://github.com/aperperidis/summer_internship_cystic_fibrosis

7. Conclusion

In summary, our exploration of X-ray Velocimetry (XV) imaging and machine learning for cystic fibrosis (CF) detection showcased the effectiveness of the 3D Local Binary Pattern (LBP) feature selection. The 'ensemble 3' model, incorporating SVM, KNN, and SGD, demonstrated consistent high accuracy across mouse studies (89.5% testing accuracy). Challenges arose in rat studies, indicating the need for further refinement. Overall, the combination of XV imaging and machine learning holds promise for advancing CF detection, with potential applications in medical diagnostics.

References

[1] Werdiger, F, Donnelley, M, Dubsky, S, Murrie, RP, Carnibella, RP, Samarage, CR, ... Morgan, KS 2020, 'Quantification of muco-obstructive lung disease variability in mice via laboratory X-ray velocimetry', Scientific Reports, vol. 10, no. 1, pp. 10859–10859.

[2] Nandan Mohanty, S, Garcia Diaz, V & Satish Kumar, GAE 2023, 'GLCM Based Feature Extraction and Medical X-ray Image Classification Using Machine Learning Techniques', in Intelligent Systems and Machine Learning, vol. 470, Springer, Switzerland.

[3] Paulhac, L, Makris, P & Ramel, J-Y 2008, 'Comparison between 2D and 3D Local Binary Pattern Methods for Characterisation of Three-Dimensional Textures', in ICIAR, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 670–679.