






Classification of Diffuse Lung Diseases Using Heterogeneous Ensemble Classifiers

Shyla Raj¹ , D. S. Vinod¹ , and Nagaraj Murthy² 

¹ Department of Information Science and Engineering, Sri Jayachamarajendra College of Engineering, Mysuru, Karnataka, India

shylanataraj@gmail.com, dsvinod@daad-alumni.de

² Department of Radiology, JSS Medical College, Mysuru, Karnataka, India
nagnishu@gmail.com

Abstract. Accurate classification of Diffuse Lung Diseases (DLD) plays a significant role in the identification of the lung pathology. Efficient classifiers based on various learning strategies have been proposed for multi class DLD classification. Due to imbalance in DLD class distribution the mis-classification probability of minority class is higher when compared to the majority class. To overcome the affects of imbalance in class distribution, the sampling approach is employed in the work, to balance the training set. It is observed that recognition rate of each DLD class is distinct based on the learning method adopted. Thus the complementary information offered by each classifier can be fused effectively to boost the classification performance. A heterogeneous ensemble classifier method based on weighted majority voting scheme is presented in this work to classify five DLD patterns imaged in High Resolution Computed Tomography (HRCT). The efficiency of the base and ensemble classifier is assessed based on recall, precision, F-measure and G-mean measure. By comparison it is found the results by ensemble of classifiers is superior than compared to its base classifiers.

Keywords: Diffuse lung diseases · Voting · Ensemble · TALISMAN · Classification

1 Introduction

DLDs are the diverse group of irreversible pulmonary disorders which causes difficulty in breathing and if untreated results in death. Globally the death rate due to DLD has increased by about 6% in the last decade [29]. Since the symptoms exhibited by DLDs are similar, the HRCT scans are utilized for accurate diagnosis. The HRCT modality is chosen because it shows significant difference between healthy and affected lung tissues. But manually investigating HRCT is strenuous due to inter and intra class variations among the lung patterns and subjective errors that may arise due to inexperience of the radiologist. Thus research has been ongoing to build Computer-Aided Diagnosis (CAD) to support the radiologist in the interpretation of HRCT scans. Identification/classification of lung patterns is one of the most step in CAD system.

During the conduction of applied research Hansen and Salamon [21] discovered that the prediction made by combination of classifiers achieve better classification results than compared to its individual counterparts. This observation has motivated this work to use ensemble of classifiers to classify DLD patterns. Ensemble learning is a machine learning framework in which individual decision of set of classifiers are fused in a particular way to achieve better classification. The decisions of three heterogeneous classifiers viz. Gaussian Support Vector Machine (GSVM), Weighted k Nearest Neighbour (Wk-NN) and Decision Tree (DT) are combined using weighted majority voting scheme in the work. The work focuses on classification of five DLD patterns namely: Emphysema (E), Fibrosis (F), Ground Glass Opacity (GGO), Healthy (H) and Micro-Nodules (MN).

Alike the other medical classification problems, the DLDs also suffers from the imbalance in the class distribution. The rate of occurrence of each pathology is distinct and hence some classes has fewer instances than the other. Applying classification on such dataset results in false assessment of overall accuracy. Since most of the machine learning algorithm aims to achieve higher overall accuracy, it tries to over-fit the majority class and neglect the minority class. This adversely affects the recognition rate of minority class. To overcome this problem, the minority class in the training set is oversampled in feature space to match the majority class. The Synthetic Minority Oversampling Technique (SMOTE) algorithm [7] is used for generating the synthetic samples.

To our best knowledge, this is the first work to address the imbalance in class distribution of 'TALISMAN' dataset by using the oversampling technique and adopt voting technique in ensemble learning for DLD pattern classification.

2 Related Work

Classification of DLDs is an important step in the CAD system for developed for differential diagnosis of DLD. The commonly used classifier include the Bayesian classifier, k-NN classifier, random forest and the widely utilized classifier in the literature is SVM. The works [36] and [6] employ the bayesian classifier, the k-NN classifier is used in [12, 26, 32]. The SVM is applied in [1, 3, 24, 25] and [14]. The feed forward neural network is adopted in [15, 16] and back propagation neural network in [35]. By analysing the above classification results it is observed that wrongly classified samples by each of the distinct classifiers is usually different. The DLD class wise recognition varies based on the learning technique employed in the classifier. Thus effective fusion of the complementary information from each classifiers can be used to boost the efficiency of DLD classification system. The multi-classifier approach or the ensemble learning is widely used in literature for various pattern recognition problems.

Dash et al. [10] presented a multi-classifier approach on the basis of winning neuron strategy for lung tissue classification. The work used the results of Neural Network (NN) and Naive Bayesian (NB) classifiers. Onan et al. [31] has employed ensemble classifier for text sentiment analysis. The decisions of NB, SVM, Logistic Regression, Discriminant Analysis and Bayesian Logistic Regression are fused using weighted voting scheme. Ye et al. [38] presented a decision machine based on weighted majority voting. The machine combined the benefits of SVM and Artificial Neural Network (ANN) for fault diagnosis. The network traffic was classified using multi classification approach

in [9]. The method explored combination techniques such as majority voting, weighted majority voting, Naive Bayes, Dempster-Shafer combiner, Behavior-Knowledge Space (BKS), Wernecke's (WER) method and oracle. Bashir et al. [4] proposed a novel ensemble learning based on enhanced bagging technique for heart diseases prediction. The framework was built using NB, quadratic discriminant analysis, SVM, linear regression and instance based learner.

3 Dataset

The work uses benchmark 'TALISMAN' dataset [13]. A total of 11, 053 patches of size 32×32 are extracted from the provided 1946 2-D Annotated Region Of Interest (AROI). The pattern wise distribution of DLD patches is represented in the Table 1. It can be inferred from Table 1 that the dataset is skewed in distribution.

Table 1. The DLD pattern wise distribution selected for work

DLD type	E	F	GGO	H	MN
Total samples	422	2989	2226	3032	2384

4 Methodology

The Algorithm 1 explains the steps of classifier ensemble employed in the work. In the first step, texture features are extracted for the entire dataset using Fuzzy Local Binary Pattern (FLBP), Grey Level Occurrence Matrix (GLCM) and Grey level Run length Matrix (GLRLM) and Intensity feature from Intensity Histogram (IH). In the second step, the dataset is divided into training and testing set by using stratified partitioning technique. The third step involves generating the synthetic samples for training set in feature space by oversampling using SMOTE algorithm. The fourth step involves classification. The GSVM, Wk-NN and DT are used as base classifier. In the final step, individual decisions from the base classifiers are fused by using weighted majority voting scheme to get the final class labels.

4.1 Feature Extraction

The DLD patterns are manifested as textural alternation in lung parenchyma hence texture based features are extracted in the work.

Algorithm 1: Ensemble classification

- Data:** The DLD image patches $p=\{p_1, p_2, \dots, p_n\}$
- Result:** Final class labels of DLD patches $X' = \{X'_1, X'_2, \dots, X'_N\}$
- Step 1:** Compute Texture and Intensity features for all DLD patches
 $Feature_Vector = \{FLBP; GLRLM; GLCM; IH\}$
- Step 2:** Generate training and testing sets by using stratified partitioning technique.
- Step 3:** Balance the training set using SMOTE in feature space
- Step 4:** Train GSVM, WK-NN and DT classifiers in parallel
- Step 5:** Obtain class labels for test data individually from trained classifiers $X^1 = \{X^1_1, X^1_2, \dots, X^1_N\}$; $X^2 = \{X^2_1, X^2_2, \dots, X^2_N\}$; $X^3 = \{X^3_1, X^3_2, \dots, X^3_N\}$
- Step 6:** Determine final class label of test data X' by performing $weighted_majority_voting = (X^1; X^2; X^3)$ (calls Algorithm 2)
-

FLBP proposed by Naresh and Nagendraswamy [30] is used in the work. FLBP is a powerful textural descriptor which overcomes the disadvantage of hard thresholding of the traditional LBP approach. The intensity of the image is transformed according to fuzzy triangular membership. Further the difference between center pixel and the neighbourhood is calculated and histogram is created. The main advantage of FLBP is that it gives both spatial and statistical information. For the work, FLBP patterns with 10 bins is extracted.

GLCM is a 2^{nd} order statistical method proposed by Haralick et al. [22]. GLCM can be used to analyse the spatial distribution of pixel intensities. GLCM yields a 2-D matrix which gives information about how frequently pixel with intensity 'a' appear in particular spatial relationship with pixel with intensity 'b'. The spatial relationship between pixels in analysed by considering the neighbouring properties of the 26-connected neighbours. From each orientation nine features are calculated. Finally corresponding values in all orientations are averaged to obtain final resultant vector. Nine GLCM features are used in the work [37].

GLRLM is also a 2^{nd} order statistical method proposed by Galloway [18] which yields a 2-D matrix RLM(a, b) which defines the relation between number of runs of gray level 'a' of length 'b' in a particular orientation. This matrix represents information regarding the connected length of a particular pixel in a particular direction. Run lengths are acquired in thirteen directions. In each direction the thirteen texture features [37] are acquired from the RLM matrix. The final GLRLM vector is calculated by averaging each of GLRLM feature in all directions.

4.2 Addressing Class Imbalance

In the medical recognition problems the recognition of each class is crucial, applying classification on dataset with imbalanced class distribution results in false assessment of overall accuracy because classifiers in general are biased towards the majority class

and tend to neglect the minority class. Hence the imbalance in class distribution need to be addressed. In the work, the oversampling approach is applied to balance only the training dataset and imbalance in testing set prevails.

The SMOTE algorithm, is an oversampling technique which creates synthetic samples in feature space for minority class from the existing minority samples instead of simply creating their copies. The SMOTE finds the n -nearest neighbours for each sample in minority class and takes the difference between the neighbours. The difference is multiplied with a random number between 0 and 1. This value generates a new point i.e. sample between the existing neighbours. The class re-distribution of DLD patches after employing SMOTE is represented in Table 2 respectively.

Table 2. The DLD pattern wise distribution after sampling

DLD type	E	F	GGO	H	MN
Training set	2124	2092	2118	2122	2119
Testing set	127	897	668	910	715

4.3 Classification

According to Krogh and Vedelsby [27] to built a good ensemble classifiers, the base classifiers must be accurate and as diverse as possible. The diversity of classifier refers to the learning approach adopted or in sub sampling the training examples. The SVM belongs to the family of generalized linear classifiers while k-NN belongs to the family of instance-based learning and the DT constructs tree like structure to classify the data. Thus three diverse classifiers GSVM, Wk-NN and DT are chosen for the work.

SVM. The SVM with the Gaussian kernel [2] is used for categorization of the DLD patterns. The one-versus-one (OVO) strategy is used for multi-class categorization. In the OVO approach a separate classifier is trained for each pair of class label. Thus it involves $M(M - 1)/2$ classifier for M class problem. All $M(M - 1)/2$ classifiers are involved in predicting the class label of unlabelled sample and sample is given the class label for which it gets majority number of votes. The best values of regularization parameter ‘C’ and the kernel width σ is found using trial and error approach on the training set.

Wk-NN. The Wk-NN a variation of traditional K-nn is used in the work [5]. The Wk-nn overcomes the shortcoming of simple majority voting employed in k-nn. In the Wk-nn each neighbour $n_i \in ne_K(x)$ is associated with a weight w_n , where $ne_K(x) = \{n_1, n_2, \dots, n_K\}$ are K points selected from training data. For each test data point to be classified different set of weights are assigned to the neighbour based on its inverse distance from new data point i.e. the neighbours closer to the test data point will have a greater influence than neighbours that are further away. This method gives more importance or greater weightage to the neighbours that are close to the test data point and the decision is less affected by the neighbours that are far from the new data point.

Decision Tree. Decision tree is a supervised, non-parametric learning algorithm [17]. The decision tree learns decision rules from labelled training data by constructing tree structure in the form of flowchart. The internal node of the tree presents the test on the attribute, every branch depicts the result of the test and every leaf node represents the class label. The decision tree can be easily transformed into if-then-else classification rules. Given a testing data, for which class is unknown, the target class is estimated by testing the data against the constructed decision tree.

4.4 Voting Scheme

Voting techniques are simple, efficient and widely used method for decision fusion. The voting technique can be weighted or un-weighted. In un-weighted/simple majority voting scheme the final decision X' is the output which atleast more than half number of base classifiers C agree on.

$$X' = mode\{C_1(y), C_2(y) \dots C_n(y)\} \quad (1)$$

In weighted majority scheme [8] each base classifier is assigned with the weights. The weights to the classifier are assigned automatically based on the misclassification cost (ϵ_i) [38]. The weight (w_i) is inversely proportional to the misclassification cost, mathematically it is writtenas:

$$w_i = \log_2((1 - \epsilon_i)/\epsilon_i) \quad (2)$$

The individual weights are normalized such that sum of all weights equals to 1. The final decision X' is the one for which summation of individual classifier decision with its corresponding weights, is the highest. The final decision X' of weighted majority is given as follows:

$$X' = argmax_j \sum_{i=1}^N w_i \chi L(C_i(y) = j) \quad (3)$$

where L is the set of unique class labels and χL is the characteristic function:

$$[C_i(y) = j \in L] \quad (4)$$

Algorithm 2 explains the weighted majority voting scheme.

Algorithm 2: Weighted majority voting

Data: Class label of individual classifier $X^1 = \{X_1^1, X_2^1, \dots, X_N^1\}$;
 $X^2 = \{X_1^2, X_2^2, \dots, X_N^2\}$; $X^3 = \{X_1^3, X_2^3, \dots, X_N^3\}$
 Weights for individual classifier $w = \{w_1, w_2, w_3\}$

Result: Final class label $X' = \{X_1', X_2', \dots, X_N'\}$

Step 1: Gather the class label from individual classifier.

Step 2: Calculate the weights of individual classifier using equation (2).

Step 3: Final class label is deduced by considering the majority of the individual weighted decision as shown in equation (3).

4.5 Experimental Set-up and Evaluation Metric

A stratified ten cross fold validation technique is used to evaluate the efficiency of the ensemble classifier. The stratification approach is employed on both balanced and imbalanced dataset, to ensure that representative samples from each class should appear in both training and testing set. The simple majority voting scheme assigns equal weights to all the classifier whereas in the weighted majority voting scheme the weights to the classifiers are assigned based on the performance of the classifier.

The classification performance is evaluated on the basis of Recall, Precision, Fscore and G-mean measures. In the medical research domain, the main goal is to reduce the false negatives. Hence to gauge the efficiency of ensemble classification, recall measure is used. The precision measure is equally important as recall as it defines the classifier's exactness. A high precision value indicates lower number of False Positives. F-score is a weighted average measured based precision and recall value. G-mean is the harmonic mean of sensitivity and specificity which provides better assessment of classifier than the accuracy measure in case of imbalanced dataset.

5 Result and Discussion

Table 3 shows the performance comparison between the individual base classifiers and the results obtained from the voting scheme for imbalanced dataset. Amongst the voting schemes it can be noted that the performance of the simple majority performs better in terms of recall and G-mean than the weighted majority but it fails in achieving higher Precision. Hence weighted majority voting is considered better than simple voting.

Table 3. Performance comparison of base classifiers and ensemble classifier for imbalanced dataset

Method	GSVM	WKNN	DT	Simple majority	Weighted majority
Recall	69.12	74.12	73.16	78.16	76.54
Precision	84.14	76.50	77.24	82.35	85.16
F-score	71.26	75.25	74.59	79.78	79.24
G-mean	78.39	83.49	82.58	86.03	84.84

Table 4 gives more insights about how different classes of DLDs are recognised by each learning algorithm. As mentioned earlier each individual classifier identifies DLD type distinctly based on the learning technique adopted. The Emphysema (E) and Micro-nodules (MN) is well recognized by WKNN while Fibrosis (F), Healthy (H) and Ground Glass Opacity (GGO) by GSVM. It can be noted that, no individual classifier is able to recognize all lung patterns effectively. Thus fusing the individual decisions helps in the overall performance improvement of multi-class classification problem.

It can be noticed from Table 4 that the recall measures of minority class E, GGO and MN are lower than the rest in imbalanced dataset. This is because, the classifier

Table 4. Illustration of recall measure of each DLD class for imbalanced dataset

Method	GSVM	WKNN	DT	Weighted majority
Emphysema (E)	19.61	57.87	49.08	48.03
Fibrosis (F)	93.86	81.85	85.16	91.39
Ground Glass Opacity (GGO)	74.51	74.03	73.83	78.10
Healthy (H)	85.35	81.12	83.12	87.08
Micro-nodules (MN)	76.28	76.85	74.96	78.08

tries to improve the recognition rate of the majority class while the mis-classification of minority class is neglected. To overcome this bias, oversampling the minority class is employed to match them to majority class.

Table 5 depicts the performance comparison between the individual base classifiers and the results obtained from ensemble classifier. These results are obtained after balancing the dataset. In here the weighted majority voting performs better than the simple majority voting in all measures. On comparing the results of ensemble with base classifiers, it can be noted that there is considerable improvement in results by the ensemble classifier.

Table 5. Performance comparison of base classifiers and ensemble classifier for balanced data

Method	GSVM	WKNN	DT	Simple majority	Weighted majority
Recall	76.38	76.36	75.49	79.95	80.24
Precision	81.79	74.13	71.10	80.54	81.23
F-score	78.28	75.03	72.18	80.14	80.62
G-mean	84.73	84.78	84.23	87.20	87.40

The class wise recall value obtained for balanced dataset is presented in Table 6. It can be well noted that the recognition of the minority class patterns E, GGO and MN has been enhanced after re-balancing the dataset. Contrasting the results between balanced Table 6 and imbalanced data Table 4, it can be perceived that balancing the dataset helps in the recognition of all DLD types, thus improving the overall recognition performance.

Table 7 illustrates the comparison result of proposed ensemble classifier with existing work in literature. By contrasting the results, it can be inferred that proposed method provides promising results than the most of existing work.

The results of class-wise recall comparison of proposed work with other state of art techniques in literature is tabulated in Table 8. The recognition rate for the Fibrosis (F) pattern by the proposed method outperforms all the existing methods including deep learning [19]. The recognition rate of Ground Glass (GGO) and Healthy (H) are equally promising when compared to other existing methods. While Emphysema (E) and Micronodules (Mn) patterns need further analysis to improve their recognition.

Table 6. Illustration of recall measure of each DLD class for balanced dataset

Method	GSVM	WKNN	DT	Weighted majority
Emphysema (E)	55.91	71.50	70.24	69.61
Fibrosis (F)	92.80	80.56	79.98	89.31
Ground Glass Opacity (GGO)	77.56	75.30	78.29	80.66
Healthy (H)	83.10	77.71	75.37	84.13
Micro-nodules (MN)	72.53	76.74	73.58	77.50

Table 7. Contrasting the proposed work with existing work in literature

Method	Recall (%)	Precision (%)	F-score (%)
Proposed	80.2	81.2	80.6
Joyseeree et al. (2018) [24]	74.5	80.7	77.1
Song et al. (2015) [33]	84.1	82.5	83.3
Song et al. (2013) [34]	82.6	80.7	81.5
Li et al. (2013) [28]	74.4	70.2	na
Depeursinge et al. (2012) [14]	75.8	76.3	76

Table 8. The DLD pattern wise comparison of proposed work with existing work in literature

Method	E (%)	F (%)	GGO (%)	H (%)	Mn (%)
Proposed	69.61	89.31	80.66	84.13	77.50
Joyseeree et al. (2018) [24]	57.25	82.41	72.73	72.6	87.54
Gao et al. (2016) [19]	82.70	89.10	81.51	91.42	87.99
Gupta et al. (2016) [20]	75.00	71.8	62.00	81.20	71.4
Shin et al. (2016) [23]	91.0	83.0	70.0	68.0	79.0
Song et al. (2015) [33]	79.6	85.4	80	88.5	87.2
Depeursinge et al. (2012) [11]	72.70	84.20	68.40	82.70	83.50

6 Conclusion

The oversampling approach has been adopted in the work to overcome the imbalance in the class distribution and the bias towards minority class. An ensemble classifier based on weighted majority voting is presented for multi-class categorization of DLD patterns. The individual decisions of the base classifiers are fused to achieve higher classification efficiency. The experimental results clearly exhibit the performance boost in Recall, Precision, F-score and G-mean values by ensemble classifier than its base

classifier. Thus it can be concluded that balancing the dataset and adopting ensemble approach for classification helps in improving the overall performance of multi-class DLD classification problem.

References

1. Ajin, M., Mredhula, L.: Diagnosis of interstitial lung disease by pattern classification. *Procedia Comput. Sci.* **115**, 195–208 (2017)
2. Amari, S.I., Wu, S.: Improving support vector machine classifiers by modifying kernel functions. *Neural Netw.* **12**(6), 783–789 (1999)
3. Bagesteiro, L.D., Oliveira, L.F., Weingaertner, D.: Blockwise classification of lung patterns in unsegmented CT images. In: 2015 IEEE 28th International Symposium on Computer-Based Medical Systems (CBMS), pp. 177–182. IEEE (2015)
4. Bashir, S., Qamar, U., Khan, F.H.: BagMOOV: a novel ensemble for heart disease prediction bootstrap aggregation with multi-objective optimized voting. *Aust. Phys. Eng. Sci. Med.* **38**(2), 305–323 (2015)
5. Bicego, M., Loog, M.: Weighted k-nearest neighbor revisited. In: 2016 23rd International Conference on Pattern Recognition (ICPR), pp. 1642–1647. IEEE (2016)
6. Chabat, F., Yang, G.Z., Hansell, D.M.: Obstructive lung diseases: texture classification for differentiation at ct. *Radiology* **228**(3), 871–877 (2003)
7. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002)
8. Cordella, Luigi P., De Stefano, C., Fontanella, F., Scotto di Freca, A.: A weighted majority vote strategy using Bayesian Networks. In: Petrosino, A. (ed.) ICIAP 2013. LNCS, vol. 8157, pp. 219–228. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41184-7_23
9. Dainotti, A., Pescapé, A., Sansone, C.: Early classification of network traffic through multi-classification. In: Domingo-Pascual, J., Shavitt, Y., Uhlig, S. (eds.) TMA 2011. LNCS, vol. 6613, pp. 122–135. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-20305-3_11
10. Dash, J.K., Mukhopadhyay, S., Garg, M.K., Prabhakar, N., Khandelwal, N.: Multiclassifier framework for lung tissue classification. In: Students' Technology Symposium (TechSym), 2014 IEEE, pp. 264–269. IEEE (2014)
11. Depeursinge, A., Foncubierta-Rodríguez, A., Van de Ville, D., Müller, H.: Multiscale lung texture signature learning using the Riesz transform. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012. LNCS, vol. 7512, pp. 517–524. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33454-2_64
12. Depeursinge, A., Sage, D., Hidki, A., Platon, A., Poletti, P.A., Unser, M., Müller, H.: Lung tissue classification using wavelet frames. In: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2007, pp. 6259–6262. IEEE (2007)
13. Depeursinge, A., Vargas, A., Platon, A., Geissbuhler, A., Poletti, P.A., Müller, H.: Building a reference multimedia database for interstitial lung diseases. *Comput. Med. Imaging Graph.* **36**(3), 227–238 (2012)
14. Depeursinge, A., Van de Ville, D., Platon, A., Geissbuhler, A., Poletti, P.A., Müller, H.: Near-affine-invariant texture learning for lung tissue analysis using isotropic wavelet frames. *IEEE Trans. Inf Technol. Biomed.* **16**(4), 665–675 (2012)
15. Dudhane, A., Shingadkar, G., Sanghavi, P., Jankharia, B., Talbar, S.: Interstitial lung disease classification using feed forward neural networks. In: ICCASP, Advances in Intelligent Systems Research, vol. 137, pp. 515–521 (2017)

16. Dudhane, Akshay A., Talbar, Sanjay N.: Multi-scale directional mask pattern for medical image classification and retrieval. In: Chaudhuri, Bidyut B., Kankanhalli, Mohan S., Raman, B. (eds.) Proceedings of 2nd International Conference on Computer Vision & Image Processing. AISC, vol. 703, pp. 345–357. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-7895-8_27
17. Friedl, M.A., Brodley, C.E.: Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **61**(3), 399–409 (1997)
18. Galloway, M.M.: Texture analysis using grey level run lengths. NASA STI/Recon Technical Report N 75 (1974)
19. Gao, M., et al.: Holistic classification of ct attenuation patterns for interstitial lung diseases via deep convolutional neural networks. *Comput. Methods in Biomech. Biomed. Eng. Imaging Vis.* **6**(1), 1–6 (2018)
20. Gupta, R.D., Dash, J.K., Mukhopadhyay, S.: Content based retrieval of interstitial lung disease patterns using spatial distribution of intensity, gradient magnitude and gradient direction. In: 2016 International Conference on Systems in Medicine and Biology (ICSMB), pp. 58–61. IEEE (2016)
21. Hansen, L.K., Salamon, P.: Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(10), 993–1001 (1990)
22. Haralick, R.M., Shanmugam, K., et al.: Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **6**, 610–621 (1973)
23. Hoo-Chang, S., et al.: Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **35**(5), 1285 (2016)
24. Joyseeree, R., Müller, H., Depeursinge, A.: Rotation-covariant tissue analysis for interstitial lung diseases using learned steerable filters: performance evaluation and relevance for diagnostic aid. *Comput. Med. Imaging Graph.* **64**, 1–11 (2018)
25. Kale, M., Mukhopadhyay, S., Dash, J.K., Garg, M., Khandelwal, N.: Differentiation of several interstitial lung disease patterns in HRCT images using support vector machine: role of databases on performance. In: Medical Imaging 2016: Computer-Aided Diagnosis, vol. 9785, p. 97852Z. International Society for Optics and Photonics (2016)
26. Korfiatis, P.D., Karahaliou, A.N., Kazantzi, A.D., Kalogeropoulou, C., Costaridou, L.I.: Texture-based identification and characterization of interstitial pneumonia patterns in lung multidetector CT. *IEEE Trans. Inf Technol. Biomed.* **14**(3), 675–680 (2010)
27. Krogh, A., Vedelsby, J.: Neural network ensembles, cross validation, and active learning. In: Advances in Neural Information Processing Systems, pp. 231–238 (1995)
28. Li, Q., Cai, W., Feng, D.D.: Lung image patch classification with automatic feature learning. In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6079–6082. IEEE (2013)
29. Naghavi, M., et al.: Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the global burden of disease study 2016. *Lancet* **390**(10100), 1151–1210 (2017)
30. Naresh, Y., Nagendraswamy, H.: A novel fuzzy LBP based symbolic representation technique for classification of medicinal plants. In: 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 524–528. IEEE (2015)
31. Onan, A., Korukoğlu, S., Bulut, H.: A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification. *Expert Syst. Appl.* **62**, 1–16 (2016)
32. Sluimer, I.C., van Waes, P.F., Viergever, M.A., van Ginneken, B.: Computer-aided diagnosis in high resolution ct of the lungs. *Med. Phys.* **30**(12), 3081–3090 (2003)

33. Song, Y., Cai, W., Huang, H., Zhou, Y., Wang, Y., Feng, D.D.: Locality-constrained subcluster representation ensemble for lung image classification. *Med. Image Anal.* **22**(1), 102–113 (2015)
34. Song, Y., Cai, W., Zhou, Y., Feng, D.D.: Feature-based image patch approximation for lung tissue classification. *IEEE Trans. Med. Imaging* **32**(4), 797–808 (2013)
35. Uchiyama, Y., et al.: Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography. *Med. Phys.* **30**(9), 2440–2454 (2003)
36. Uppaluri, R., Hoffman, E.A., Sonka, M., Hunninghake, G.W., McLennan, G.: Interstitial lung disease: a quantitative study using the adaptive multiple feature method. *Am. J. Respir. Crit. Care Med.* **159**(2), 519–525 (1999)
37. Vallières, M., Freeman, C.R., Skamene, S.R., El Naqa, I.: A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys. Med. Biol.* **60**(14), 5471 (2015)
38. Ye, F., Zhang, Z., Chakrabarty, K., Gu, X.: Board-level functional fault diagnosis using artificial neural networks, support-vector machines, and weighted-majority voting. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **32**(5), 723–736 (2013)