Neural Computing and Applications
https://doi.org/10.1007/s00521-022-07566-x

ORIGINAL ARTICLE

Automated classification of emphysema using data augmentation and effective pixel location estimation with multi-scale residual network

T. Manikandan[1] · S. Maheswari[2]

Received: 7 April 2021 / Accepted: 28 June 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

Abstract
In modern medical diagnosis also, the emphysema is still recognized by the computed tomography (CT) scans with a set of defined patterns as a classification problem in computer vision. There were as many algorithms developed in the past that attempt to classify the underlying patterns and their relevant associated clusters by modeling an automated system. And this classification modeling approach is responsible for the benchmarking classification and quantification of various emphysematous tissues from lung CT images on different scales in the literature. Hence, with the same motivation and intents, this article put forth a multiscale residual network with data augmentation model (MS-ResNet-DA). First, a generative adversarial network (GAN) is employed to augment the training samples and avoid the overfitting problem. These images are again augmented based on different image processing methods. Then, the obtained images are learned by MS-ResNet to categorize the emphysema. Still, the accuracies of categorizing the centrilobular emphysema (CLE) and panlobular emphysema (PLE) are not satisfactory because they do not have spatial dependence. So, an enhanced MS-ResNet-DA (EMS-ResNet-DA) model is proposed, which applies an effective position estimation algorithm to measure relative and absolute location data of emphysema pixels in the images. The relative location data give the current location of the emphysema pixel by extracting the relative dislocation measures from CT images. Also, the absolute location estimation model is based on the position encoding network to match the diseased image with the reference emphysema images and validate whether location data are implicitly learned when trained on categorical labels. Moreover, these location data of all pixels in the images are learned by the MS-ResNet for emphysema classification. Finally, the experimental results demonstrated that the EMS-ResNet-DA achieves an overall classification accuracy of 94.6% that outclasses the conventional models.

Keywords Lung diseases · Emphysema classification · Deep learning · MS-ResNet · Data augmentation · GAN · Relative location estimation · Absolute location estimation · Position encoding network

1 Introduction

A substantial number of disorders that threaten the global population are lung-related. Cardiovascular complications like chronic obstructive pulmonary disease (COPD), asthma, fibrosis and pneumonia are impacting the foremost portion of society globally. Emphysema, which has a major risk factor for mortality worldwide, is one crucial constituent of COPD [1]. The disease can cause breathing difficulties due to the massive growth of the respiratory tract. In specific, emphysema is subclassified into CLE, paraseptic emphysema (PSE) and PLE [2]. They include various pathophysiological

✉ T. Manikandan
manikandant@vcew.ac.in; mani.siva88@gmail.com

S. Maheswari
s.maheswari@kongu.ac.in; s.maheswari.eee@kongu.edu

[1] Department of Electronics and Communication Engineering, Vivekanandha College of Engineering for Women, Thiruchengode, Tamilnadu, India

[2] Department of Electrical and Electronics Engineering, Kongu Engineering College, Erode, Tamilnadu, India

Published online: 27 July 2022

impacts [3, 4]. For instance, CLE is generally related to smoking and PSE is not always related to the major signs or neurological disabilities. As a result, it is essential to classify and quantify emphysema.

The CT scan has been deemed to be the foremost specific scanning technology to identify emphysema, classify its category and assess its seriousness [5]. In CT, emphysema's categories contain different radiology appearances [6]. Over the past few decades, several algorithms are suggested for classifying the emphysema in lung CT images. Such algorithms are categorized into supervised and unsupervised models [7, 8]. The goal of unsupervised algorithms is to find new emphysema categories, which go beyond the common categories recognized on autopsy. A generative framework [7] has been constructed to find the disease categories within emphysema and find patient groups, which are represented via different distributions of those categories. On the other hand, supervised algorithms concentrate on categorizing the typical emphysema categories detected on autopsy, which include various pathophysiologic significances.

The standard method of representing the emphysema patterns is depending on the local intensity distribution including dynamic intensity histogram [9] and kernel density measure [10]. The other algorithms define the structure of emphysema by the surface anatomy [11]. By using the adaptive multiple feature schemes [12], gradient magnitude [3] and the gray-level variance scheme [13], several techniques have been developed for emphysema patterns classification. Recently, few studies exposed the possible deep learning (DL) models for pulmonary disorder categorization. One of the most extensively utilized DL structures is the convolutional neural network (CNN) using different features from lung tissues or nodules [14]. But, it is not familiar to classify the emphysema categories since they cannot extract the high-level characteristics. Also, the majority of the algorithms to classify the emphysema was restricted to extracting low- or mid-level characteristics that possess restricted capabilities for differentiating various motifs [15].

Also, many algorithms depend on surface and intensity features; thus, intra- and inter-class variance problems remain challenging. The variances in CT scans can impact the efficiency of classifying the emphysema; therefore, it is required to develop a model that is highly strong to those variabilities. From this perspective, Peng et al. suggested a novel MS-ResNet [16] with 2 channels of unrefined CT images and its discrepancy execution part. In this MS-ResNet, the high-level features are discovered for efficiently characterizing the emphysema lesions. Initially, the raw image data are transformed into the discrepancy excitation field of individual view (i.e., patches with multiple scales) depending on Weber's principle to solve the inter-class variances. Then, both raw and transformed CT images are fed to the inputs of different channels in MS-ResNet for categorizing the emphysema into CLE, PSE, and PLE.

1.1 Problem description and major contributions

MS-ResNet leads to an overfitting problem, while the labeled training dataset is limited. As a result, it should ensure that the training dataset is adequate to avoid the overfitting problem. Therefore, in this paper, the MS-ResNet-DA model is proposed to augment the training samples and avoid the overfitting problem. The following are the key contributions of this study:

In this model, GAN is introduced as a data augmentation technique that uses a random normal vector as input to the network along with the dataset of lung CT images in 4 emphysema categories. The new images generated by GAN inherited the features of CLE, PSE, PLE, and non-emphysema (NE). Besides, those training image samples are again augmented through cropping, duplication, scaling, rotation, etc., for capturing the emphysema features, which are invariant to image-level affine transformations. Then, these samples are fed to the MS-ResNet for emphysema classification. But, the accuracies of classifying the CLE and PLE are quite low since they do not have spatial dependence. As a result, the EMS-ResNet-DA model is proposed to accurately combine pixel position information into the networks and increase the accuracy of classifying the categories of emphysema.

In this model, an effective position estimation algorithm is proposed to determine the emphysema pixel location in the images depending on the region, edge and vertex. It determines two different positions such as relative location and absolute location. The relative location provides the current location of emphysema via adding relative dislocation measures obtained from two lung disease scans. Easy addition of variables may reduce the dependability of the obtained variable measures as an emphysema pixel difference, resulting in a high inaccuracy of location information. So, absolute location estimation is essential for compensating for the location error created in relative location estimation. An absolute location estimation model based on a position encoding network is proposed to match the diseased image with the reference emphysema images and validate whether location information is implicitly learned when trained on categorical labels. This network model comprises a feed-forward convolutional encoder and simple position encoding units. First, the emphysema features at various levels of abstraction are extracted by an encoder unit. Then, these features are fed to the position encoding unit as input for predicting the absolute position

information at the end. After extracting location information of all pixels in the images obtained from the augmented training dataset, MS-ResNet is performed to classify the emphysema categories with maximum accuracy.

The remaining article is prepared as follows: Sect. 2 presents the classic models associated with the lung disease classification. Section 3 describes the MS-ResNet-DA and EMS-ResNet-DA models, and Section 4 displays their efficiency. Section 5 concludes the research work and recommends the future directions for this research work.

2 Literature survey

An enhanced CNN-based model [17] has been developed to categorize the Interstitial Lung Disease (ILD) from CT scans. Initially, the scan patches were acquired from the annotated Region-Of-Interests (ROIs), and they were spatially related or invariant to their overall intra-slice CT regions. Afterward, features with their classes were fed to the weakly supervised training-based CNN to categorize the ILD tags. But, if similar patches have many disease classes, then it was complex to classify the ILD.

Automated DL-based categorization of centrilobular emphysema from CT scans [18] was developed. The CNN was integrated into the long short-term memory (LSTM) network to predict the topic-level values from CT images. Also, the Fleischner framework was used to handle the training data. But, it requires more memory to train the data and causes an overfitting problem.

Automatic categorization of emphysema from weakly labeled high-resolution CT lung images [19] has been designed. First, the COPD patient's CT images were collected and the lung parenchyma in these images was represented via the textual features. Also, these images were trained by miSVM (multiple instances SVM) and multiple instance learning by embedded instance selection (MILES). On the other hand, the training samples were not adequate and the data imbalance problem has occurred.

An improved pneumonia prediction model [20] was suggested based on machine and DL algorithms. First, the dataset comprising a group of stroke people with and without pneumonia was obtained. After, the time-aware characteristics have been obtained and provided to the linear regression, SVM, XGBoost, multi-linear perceptron and recurrent neural network to categorize pneumonia. However, many kinds of lung disorders were needed for categorization.

The lung diseases were recognized and categorized from the CT images [21] by extracting the Haralick texture characteristics, Zernike's moments, Gabor and Tamura surface characteristics. After that, these features were fed to the improvised crow search grey wolf and cuttlefish algorithms to decide the best attribute subset. Also, K-nearest neighbor, SVM, random forest (RF) and decision tree classifiers have been used to categorize COPD. But, it has high computation time complexity.

A new framework [22] has been suggested to split tree-like structures via adjacent data and introduced a graph-cut scheme for guaranteeing the spatial and directional dependability of the constructed subtrees. Also, an RF classification was conducted to categorize the pulmonary artery vein by predicting the correlation between lung syndromes and lung tissue patterns.

An enhanced method [23] has been developed to categorize various classes of lung diseases from CT images using the ensemble of Riesz and deep characteristics. Primarily, the discriminative parametric pulmonary surface patterns were trained by the Riesz descriptions. After, the attributes from deep CNNs were determined by adjusting the inception V3 model. Such learned descriptions were fused in a softmax layer to analyze the lung tissues.

Segmentation and classification of lung parenchyma [24] were designed based on the CNN structure. Initially, a clustering scheme was applied to automatically create a dataset for training the CNN. After, a k-means clustering scheme was performed to segment the CT slices into image patches. Moreover, a cross-shaped proof, a size-junction, simultaneous testing and a patch development were utilized for generating the resultant dataset. At last, this dataset was fed to the CNN to categorize COPD. But, the dimension of patches was constant, and its effect on the efficiency of CNN was not analyzed.

A level of accuracy in the medical systems using an open dataset was analyzed [25] by the DL. First, the spectrogram attributes and labels of the annotated lung images were extracted. Then, such attributes were applied to the 2D-CNN model to categorize lung diseases. But, the training data were not sufficient. A two-phase method was recommended [26] to visualize the COPD growth at the extent of unorganized medical files. First, a 4-layer DL structure was created, which exploits a certain RNN to discover the inaccurate period-lapse parts. After, a temporal image was created by the irregular period-lapse parts to categorize the conditions of COPD. But, it needs more memory for training.

An enhanced scheme called segmentation-based deep fusion network (SDFN) [27] was developed that formulates the domain data and the high-resolution data of neighboring pulmonary areas. Initially, the neighboring pulmonary areas have been detected and clipped via the pulmonary area creator. After, 2 CNN-based classifiers have been applied as attribute miners to acquire the discriminative characteristics of the whole lung X-ray scans and the clipped pulmonary area images. Moreover, the acquired

characteristics have been combined through the attribute merging scheme to categorize the disorders. Conversely, the amount of inaccurately labeled images was not limited.

A new dual asymmetric attribute training model, namely DualCheXNet [28], has been developed to classify the multi-label thoracic diseases in lung X-rays. In this model, two asymmetric subnetworks depending on ResNet and DenseNet have been fused to obtain more discriminative characteristics of multiple irregularities from the original lung X-rays in an adaptive manner. Besides, attribute- and decision-level merging schemes were executed to create the paired attribute training embedded in DualCheXNet. Further, an iterative training mechanism was applied for combining the loss function of the categorizers into a single error and optimizing the task of paired attribute training alternatively.

A technique called dp-BREATH [29] has been suggested for identifying irregularities in lung areas and guiding the physician's concentration to the irregular pulmonary area. At first, the lung abnormalities were identified depending on the normal lung behavior by the superpixel-based method and heat map image. Also, the areas of healthy tissue were modeled by the statistical method. Each irregular area was developed and categorized based on its likelihood of having every considered irregularity. Moreover, a higher detection of radiographic structures with the probability of a specific pulmonary area to comprise irregularities was achieved by the dp-BREATH. But, it has high complexity due to the mixture of interstitial thickening and honeycombing patterns.

A neuro-heuristic method [30] has been designed for recognizing pulmonary disorders from X-ray scans. In this method, image descriptors were used depending on the spatial distribution of hue, saturation and brightness ranges on X-rays. Also, a neural network co-execution with heuristic schemes was applied for detecting deteriorated pulmonary tissues in X-rays. The neural network was used to assess the image and whether the likelihood of lung diseases was recognized and the heuristic algorithm was used to recognize the deteriorated tissues in the X-ray depending on the utilization of the fitness factor. But, the accuracy was less and the computation time was high.

A DL model [31] has been developed, which uses an enhanced multi-resolution CNN for automatically segmenting the pulmonary in many injured people. Transfer training has been adopted for accommodating the limited amount of training samples. First, it was pre-learned on person CT scans and then adjusted on canine, porcine and ovine CT scans with pulmonary damages. But, it does not analyze the segmentation efficiency in the target domain.

A novel hybrid method [32] was developed depending on modular artificial neural networks (ANN) with fuzzy logic to diagnose lung disorders. An optimal collection of characteristics was obtained by the genetic algorithm, and these were classified to identify the lung diseases. But, it has high complexity and needs to balance the trade-off between sensitivity and specificity.

An unsupervised DL model [33] was designed to investigate the annotation-less and class-imbalanced cases that normally represent pulmonary tumors. The feasibility of using medical domain knowledge like minimum intensity projection (minIP) was analyzed to emphasize the minimal variances of emphysema areas in low-dose CT scans. Also, detection maps were created to represent the model predictions and to act as an efficiency analysis. But, it did not categorize emphysema depending on severity levels, i.e., various emphysema classes.

The scalar and spatial distribution of the pulmonary damages [34] was identified by partitioning the low attenuation clusters. Also, the low attenuation clusters of efficient tiny airway disorder have been defined via the parametric response map (PRM) on the co-enrolled CT scan. Moreover, the voxel-wise criticality was estimated by the local low attenuation clusters on the co-enrolled CT to define the Bullous PRM (BPRM). But, the number of patients involved in this analysis was limited.

3 Proposed methodology

In this section, the ideology and mechanism of the proposed MS-ResNet-DA and EMS-ResNet-DA models are briefed. First, the lung CT images are collected as a training dataset, which is given to the MS-ResNet-DA model. In this model, GAN and different methods are performed for augmentation. Such samples are classified into four categories of emphysema by MS-ResNet. To further increase the classification accuracies of CLE and PLE, the EMS-ResNet-DA model is proposed, which applies the Effective Pixel Location Map Estimation (EPLME) algorithm to estimate the relative and absolute locations of emphysema pixels in the lung disease images. Moreover, these multi-scale features/pixels of emphysema are learned by the MS-ResNet for classifying the CLE, PSE and PLE effectively.

3.1 Multi-scale residual network model (MS-ResNet) classification with data augmentation model

Figure 1 depicts the overall block diagram of this MS-ResNet-DA model to classify the emphysema types. In this model, GAN is employed for augmenting the training examples and avoiding the overfitting issue by creating the

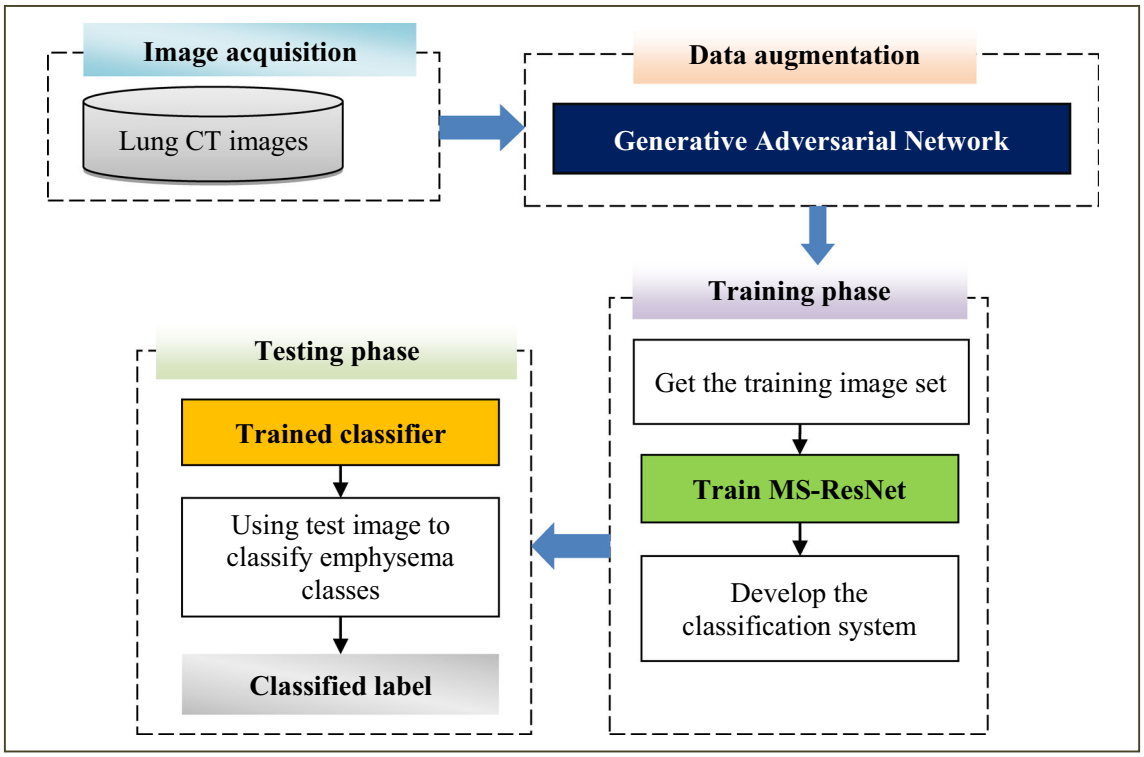


Fig. 1 Block diagram of MS-ResNet-DA model for emphysema classification

synthetic image samples while learning the MS-ResNet model. GANs are constructed by the adversarial nets to increase the CT scans.

The architecture of GAN involves two main modules such as a generator G and a discriminator D. G is utilized to translate an example from a regular distribution to the image distribution and D is learned to differentiate whether the example fits in the actual image distribution. Both G and D are trained simultaneously depending on the player concept. Figure 2 portrays a typical architecture of GAN for image augmentation.

In all iterations, G transforms an arbitrary vector Z sampled from a regular distribution into a sample scan $X_{sample} = G(Z)$ using a series of deconvolution and activation units where G is a differentiable function represented by a multilayer perceptron. The network D is trained for increasing the likelihood of deciding the accurate label for both actual and generated image samples from G.

Also, G is trained simultaneously for reducing $\log(1 - D(G(Z)))$. These two networks are learned in parallel by a two-layer minimax competition with a value factor $V(D, G)$:

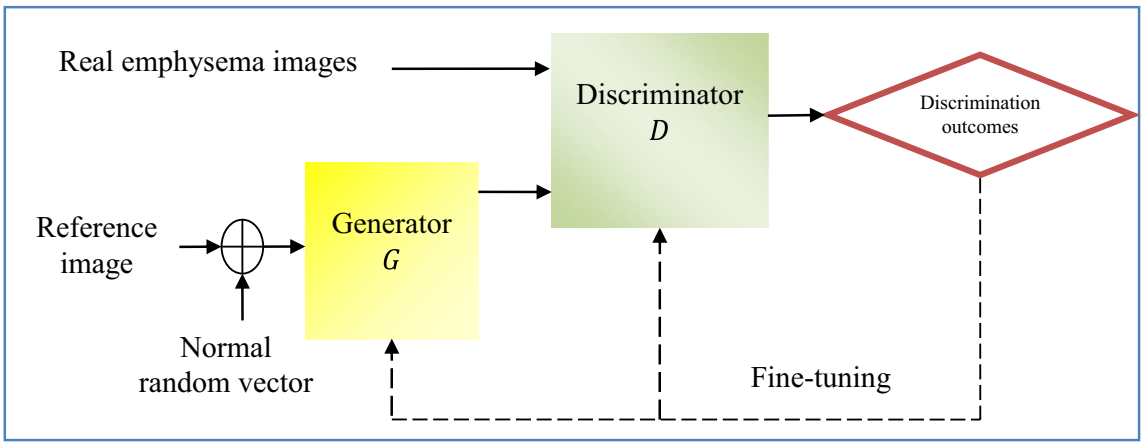


Fig. 2 Data augmentation based on GAN architecture

$$\min_{G}\max_{D} V(D,G) = \mathbb{E}_{X\sim P_{data}}[\log D(X)] + \mathbb{E}_{Z\sim noise}[\log(1-D(G(Z)))] \quad (1)$$

In Eq. (1), $\mathbb{E}$ indicates the expectation factor, $D(X)$ defines the likelihood of X fitting in the real image and $G(Z)$ or X_{sample} represents the image produced via G from Z.

Initially, the dataset of emphysema CT images is split into 4 major types such as NE, CLE, PSE and PLE. After, a GAN is trained by considering 4 categories of emphysema images with a random normal distribution vector as its input to create counterfeit images. For all classes of counterfeit images, their mean counterfeit images are found. The average of counterfeit CLE and PLE images is added, and the obtained output is deducted from the average of counterfeit PSE images to get the reference image. This reference image is applied as input to the GAN at every interval to create counterfeit images. The created counterfeit images have inherited the characteristics of 4 types of emphysema images. D differentiates whether the image is a real or counterfeit image from G.

Simultaneously, G creates images, which are highly hard to identify by D as real or counterfeit images. If the optimal of (1) is attained, then G creates an image identical to the real images, which cannot be differentiated by D. Utilizing back-propagation, G trains the counterfeit images, which extremely be similar to the real training images; thus, D cannot differentiate them from the original images.

These images are again augmented by different schemes such as cropping, duplicating and random translation within the voxels range [1, 0 or −1 pixels in every dimension (3D)], flipping, scaling, swapping in 3D axis and rotating on different angles [0°, 90°, 180°, 270°] in the dataset. In particular, random translations and rotation among input batches are carried out for upsampling and downsampling. These schemes are more useful to capture the emphysema features invariant to image-level affine transformation.

3.2 Effective MS-ResNet-DA model for emphysema classification

In this EMS-ResNet-DA model, the EPLME algorithm is used to get the relative and absolute location maps for a certain lung CT image for classifying emphysema classes. First, a bootstrap model of processing is applied for initiating the 3D representation of emphysema feature points. It utilizes the matched feature points from the two lung disease images for estimating the depth of these points and so every image is provided with a primary 3D instantiation. The feature points from two lung disease images are extracted and matched on grounds of image-plane proximity and feature similarity. Such matches are used for estimating the relative dislocation between pixel locations. Figure 3 depicts the overall block diagram of this EMS-ResNet-DA model to classify the emphysema types. This dislocation measure is used for estimating the relative location as follows:

4 Relative location estimation

It successively determines the current location map of emphysema pixels by adding the relative dislocation measures captured from any 2 lung disease scans. Consider D is a single-frame occasion latency and f'_n is a characteristic spot. Then, this scheme recursively determines the relative dislocation B_n between 2 sf'_n, i.e. (l_{n-1}, l_n), and adjusts the location of an emphysema pixel via gatheringB_n. Here, the dislocation estimation is performed depending on stereo matching, wherein the relative location prediction scheme captures B_n of lung disease images using f_n of the present given scan I_n and f'_n of the earlier scanI_{n-1}. B_n is denoted as $TR_nE_nV_n$ where T is the sampling period, R_n, E_n and V_n are the region, edge and vertex information in a lung disease images. In this model, T is assigned to 1 s for ensuring that the overlapping region between 2 consecutive frames is roughly ½ the entire scan dimension. Also, (x_n, y_n, z_n) is the 3D coordinates in which x_n and y_n are defined in terms of the height and width of an image, and accordingly z_n is the image depth.

Consider R'_n and R_n are the vectors from l_{n-1} and l_n to the 3D location f_n. Similarly, r'_n and r_n are the normalized vectors of R'_n and R_n acquired through transforming the characteristic spot in the image space into the actual image coordinates. Observe that the vector R'_n and R_n passes via f'_n in the earlier frame.

Initially, R'_n and R_n should be determined to get B_n. Here, R'_n is provided by $R'_n = \left(f_{n,x} - l_{n-1,x}, f_{n,y} - l_{n-1,y}, f_{n,z} - l_{n-1,z}\right)$ where l_{n-1} is provided and f_n is found by ray tracing with l_{n-1} and r'_n. After that, $R_n = \left(R_{n,x}, R_{n,y}, R_{n,z}\right)$ is calculated by,

$$\frac{R_{n,x}}{r_{n,x}} = \frac{R_{n,y}}{r_{n,y}} = \frac{R_{n,z}}{r_{n,z}} = \frac{f_{n,z} - l_{n,z}}{r_{n,z}} \quad (2)$$

Here, $r_n = \left(r_{n,x}, r_{n,y}, r_{n,z}\right)$. Moreover, the dislocation measure B_n is represented as:

$$B_n\left(f'_n, f_n, x_{n-1}, y_{n-1}\right) = R'_n - R_n \quad (3)$$

After the current location f_n of an emphysema pixel is estimated by adding the estimated dislocation measure B_n to the previous location f_{n-1}. In relative location estimation, the location error is approximately proportional to the

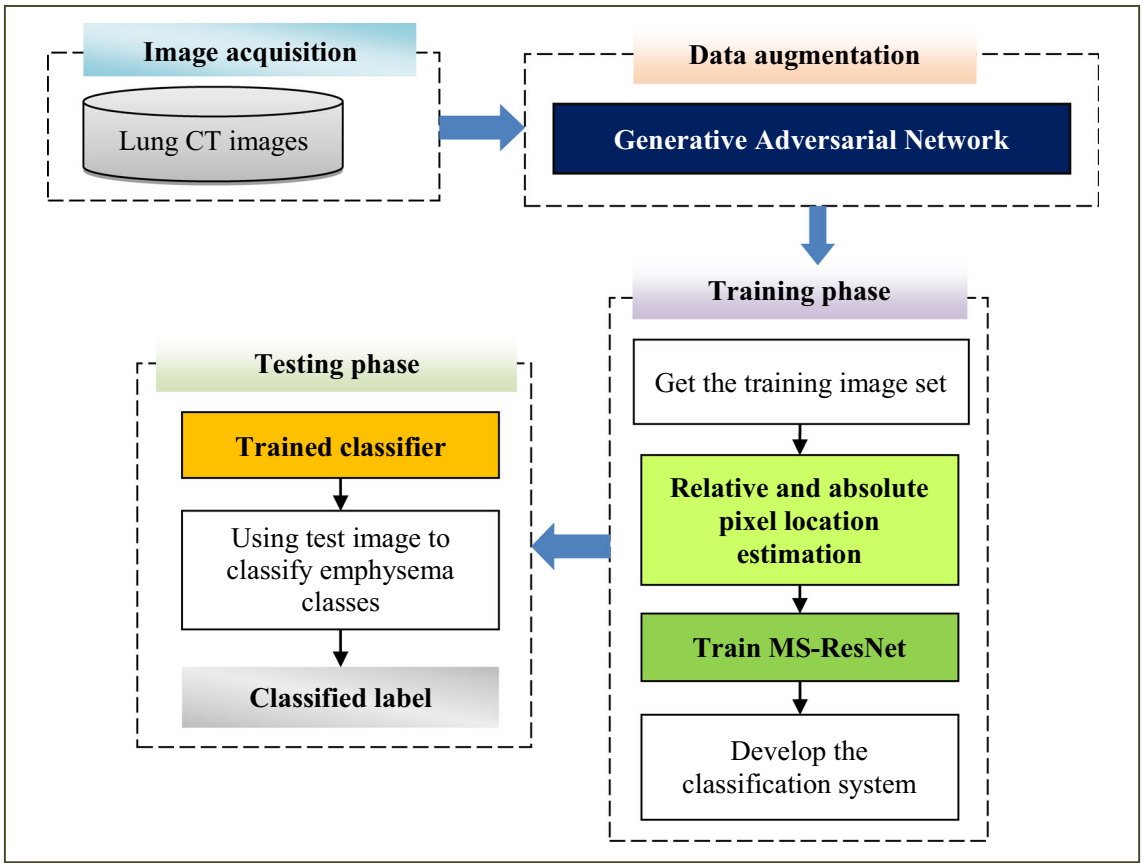


Fig. 3 Block diagram of MS-ResNet-DA model for emphysema classification

depth error, so a small depth error is needed to achieve a better classification. As a result, a relative location map for a specific lung disease image is estimated using the effective location encoding model.

5 Absolute location estimation

For an input image $I_n \in \mathbb{R}^{h\times w\times 3}$, the main aim is to estimate the gradient-like location information mask $\widehat{f}_l \in \mathbb{R}^{h\times w}$ where every pixel range indicates the absolute coordinates of a pixel from left → right or top → bottom. To solve this problem, absolute location map estimation is performed by using a position encoding network, which comprises two major units such as a feed-forward convolutional encoder network (f_{ecn}) and a simple location encoding unit (f_{leu}). Here, the encoder network extracts features at various abstraction levels from shallow layers to deep layers. The location encoding unit considers multiscale features from the encoder network as input and estimates the absolute location information at the end. Figure 4 portrays the position encoding network for absolute location estimation.

Encoder Unit: The ResNet is used for constructing the encoder networks (f_{ecn}) via eliminating the average pooling layer and the layer that allocates types. Usually, the encoder unit has 5 feature extractors represented by $(f_{f1}, f_{f2}, f_{f3}, f_{f4}, f_{f5})$. The extracted multi-scale features from bottom-to-top layers of the canonical network are represented by $(f_{l1}, f_{l2}, f_{l3}, f_{l4}, f_{l5})$. The main function is defined as:

$$f_{li} = f_{fi}(W_\alpha * I_n) \tag{4}$$

In Eq. (4), W_α are weights that are fixed and $*$ is the convolution process. Observe that in probing the encoding network, only the location encoding unit (f_{leu}) is trained for extracting the emphysema pixel location when f_{ecn} unit is applied for preserving their prior weights.

Location Encoding Unit: It considers multiscale features $(f_{l1}, f_{l2}, f_{l3}, f_{l4}, f_{l5})$ from f_{ecn} as input and creates a required location map $\widehat{f}_l$ by a transformation function $\mathcal{T}_l$. At first, $\mathcal{T}_l$ executes a bilinear interpolation process on the feature maps to obtain similar spatial dimensions resulting in a feature map f_{lc}. After obtaining similar spatial dimensions for multiscale features, those features are fused together after that a single convolutional unit with a kernel

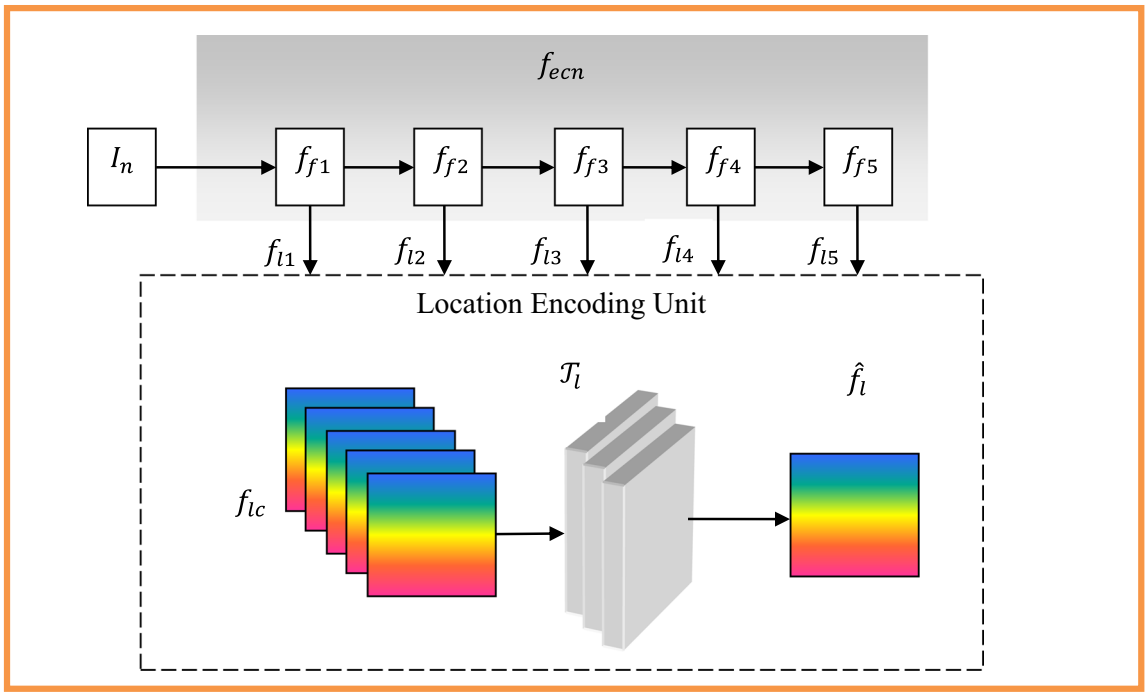


Fig. 4 Architecture of position encoding network model

dimension 3×3 without padding. The major process is defined as:

$$f_{lc} = (f_{l1} \oplus f_{l2} \oplus f_{l3} \oplus f_{l4} \oplus f_{l5}) \tag{5}$$

$$\widehat{f}_l = W_{lc} * f_{lc} \tag{6}$$

In Eq. (6),W_{lc} denotes the learnable weights involved with $\mathcal{T}_l$. The primary goal of f_{leu} is to validate whether pixel information is implicitly learned while trained on categorical labels. Also, it models the correlation between hidden location information and the gradient-like ground-truth mask.

The outcome is predicted to be random if there is no pixel location information encoded in the feature maps and vice versa. The presence of pixel location information in a network is validated by allocating a normalized gradient-like location map as a ground truth. Primarily, the gradient-like masks are created in horizontal and vertical directions. This is used to validate whether the network can learn absolute location on one or two axes.

Also, the location information in the multi-level features is probable to be modeled via a transformation by f_{leu}. The design of gradient ground truth is considered as a category of the random label since there is no correlation between the given image and the ground-truth regarding location.

This f_{leu} creates the pixel location map $\widehat{f}_l$ of interest. During training, for a given input image $I_n \in \mathbb{R}^{h \times w \times 3}$ and related ground-truth location map, the supervisory signal is applied on $\widehat{f}_l$ via upsampling it to the dimension of the ground-truth location map. After, a pixel-wise mean-squared error is defined to compute the discrepancy between the estimated and ground-truth location maps as:

$$\Delta_{\widehat{f}_l} = \frac{1}{2s} \sum_{i=1}^{s} (x_i - y_i)^2 \tag{7}$$

In Eq. (7), $x \in \mathbb{R}^s$ and $y \in \mathbb{R}^s$ are the vectorized predicted location and ground-truth map, accordingly where s is the spatial resolution, x_i and y_i are a pixel of $\widehat{f}_l$ and ground-truth location maps, accordingly. So, an absolute location map of emphysema pixels in a particular lung disease image is obtained by the position encoding network.

6 MS-ResNet

Once obtaining the relative and absolute location maps for a certain lung disease image, the MS-ResNet is trained by all patches of lung images and their corresponding location maps to create the trained model. Moreover, the trained classifier (EMS-ResNet) model is used to classify the emphysema categories of unknown lung disease CT image samples. Figure 5 portrays the presented EMS-ResNet model for emphysema classification. Thus, this model can increase the classification accuracy of CLE and PLE by augmenting the quantity of learning examples efficiently.

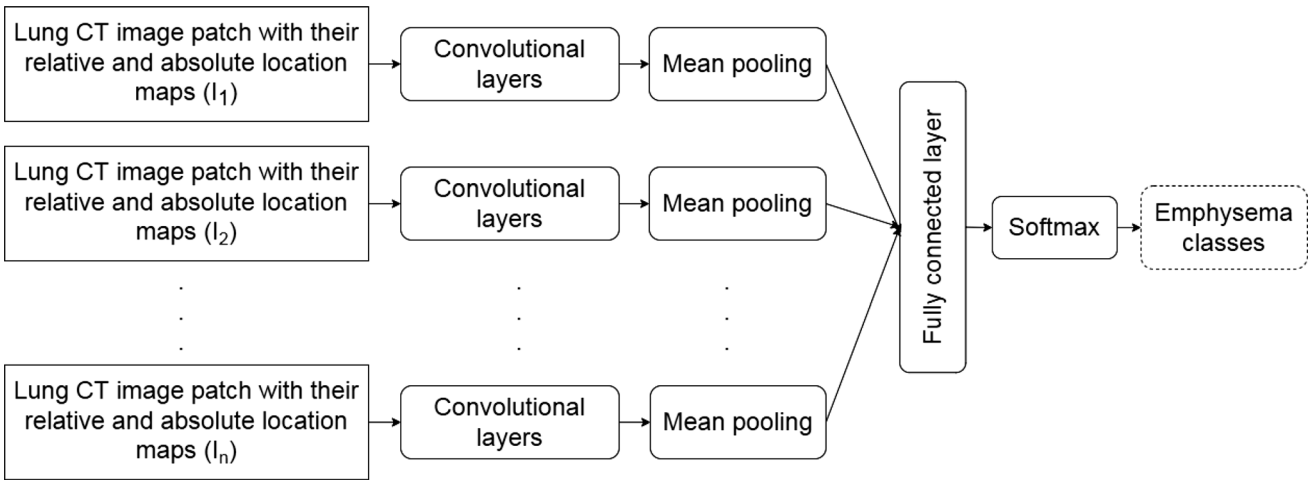


Fig. 5 EMS-ResNet model for emphysema classification

7 Experimental results

The MS-ResNet-DA and EMS-ResNet-DA models are implemented in MATLAB 2017b by using the number of lung CT images. Also, the effectiveness of these models is evaluated with the MS-ResNet based on precision, recall, F-measure and accuracy of each category of emphysema. In this experiment, a total of 115 CT images from 39 patients (9 non-smokers, 10 smokers excluding COPD and 20 smokers with COPD) are obtained from an open database available in [35]. It contains 115 high-resolution CT image slices of dimension 512 × 512 pixels, and these are labeled as normal tissues (NT-61 images), CLE (21 images), PLE (27 images) and PSE (6 images). For data augmentation, 65% of samples are considered for training and 35% are considered for testing of the GAN model. After augmentation, a total of 5520 CT images are created and labeled as the NT (1221 images), CLE (1485 images), PLE (1413 images) and PSE (1401 images). Among those images, 70% of images are taken for training and the remaining 30% are taken for testing of the MS-ResNet/EMS-ResNet model.

COPD is defined by airflow restrictions. It consists of tiny airway disorder and emphysema that is defined by the progressive degradation of pulmonary tissue. Emphysema is the key part of COPD, and accurate categorization of emphysema versus NT facilitates deep investigation of the disorder. The focus of this work is the examination of emphysema on CT scans. Emphysematous foci or blisters are observable on CT scans as regions of unusually small attenuation ranges nearby the air attenuation values. On CT, emphysema is divided into three subcategories: CLE, defined as small areas of small attenuation; PSE, defined as a monolayered region of small attenuation along the pleura, surrounded by a lobular hepatic septum that appears to be a thin white wall; PLE is defined as a low-decay lung in which the pulmonary blood vessels become smaller and smaller. Figure 6 shows examples of CLE and PSE and NT.

7.1 Analysis of pixel estimation module

To verify the existence of location information on the network, we implement a randomization test by assigning a location map (e.g., normalized gradient) as the ground truth. First, create masks such as gradients in the horizontal (H) and vertical (V) directions. The main motivation for generating these two modes is to ensure that the model can learn absolute positions on one or both axes.

The presence of pixel location information encoded in the pre-trained classifiers is analyzed in terms of mean absolute error (MAE) and Spearmen correlation (SPC).

7.1.1 SPC

It is the Spearmen's relationship between the ground truth and the predicted pixel location map for a given lung CT image. The formula for calculating the Spearman rank correlation coefficient is:

$$r = \frac{cov(ry_i, \widehat{ry_i})}{std(ry_i)\widehat{std(ry_i)}} \tag{8}$$

where cov()—is covariation function, std()—is standard deviation, ry_i—rank of the ground-truth location map, $\widehat{ry_i}$—rank of the predicted location map, r—Spearman's correlation.

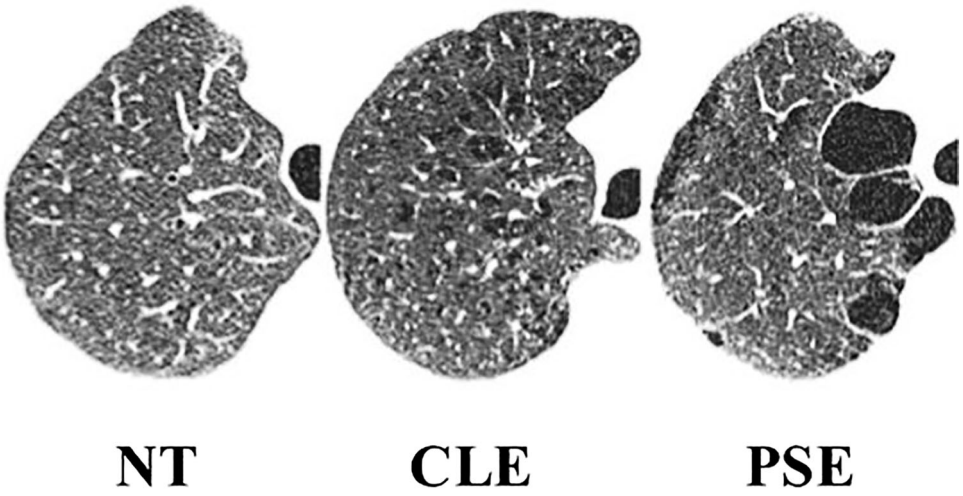


Fig. 6 Different lung tissue patterns in CT a NT. **b** CLE. **c** PSE

7.1.2 MAE

It is the mean pixel-wise variance between the estimated location map and the ground-truth gradient location map. The mean absolute error is provided by the formula:

$$MAE = \frac{1}{n}\sum_{n=1}^{n}(y_i - \widehat{y_i}) \tag{9}$$

where n—number test of samples, ry_i—actual ground-truth gradient location map, $\widehat{y_i}$—predicted location map.

Table 1 gives the quantitative analysis of pixel location estimation by implementing the different algorithms (e.g., dp-BREATH [29], minIP-DL [33], BPRM [34] and proposed EPLME) on lung CT images for emphysema classification.

As illustrated in Table 1, EMS-ResNet-DA can easily extract the emphysema pixel location information from the pre-trained classifiers. It indicates the efficiency of pixel location extraction using test datasets across H and V ground-truth patterns. This reveals that the EMS-ResNet-DA can extract the pixel's true location information. Figure 7 depicts a few samples of classified emphysema images in different classes using MS-ResNet, MS-ResNet-DA and EMS-ResNet-DA models.

In Fig. 7, based on the pixel estimation (emphysema pixel location maps), CLE is shown in blue, PLE is shown in green, and PSE is shown in red, in the simulation results. Typically, the pixel location map estimation uses pixel intensity distributions to indicate the local behavior or characteristics of an image, which can be well suited for detecting and identifying emphysema patterns at specific locations. It can be used to address irregularities present in medical images, facilitating pattern recognition methods.

7.2 Analysis of emphysema classification module

7.2.1 Accuracy

It is the fraction of accurately categorized examples in each category of emphysema over the sum amount of examples considered.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{10}$$

Table 1 Quantitative analysis of pixel location map estimation using different algorithms

Ground-truth patterns	Methods	NT		CLE		PLE		PSE	
		MAE	SPC	MAE	SPC	MAE	SPC	MAE	SPC
H	dp-BREATH	0.154	0.805	0.182	0.817	0.168	0.822	0.157	0.817
	minIP-DL	0.149	0.823	0.176	0.835	0.159	0.851	0.150	0.824
	BPRM	0.142	0.859	0.171	0.866	0.152	0.874	0.143	0.839
	EPLME	0.137	0.870	0.164	0.892	0.145	0.903	0.136	0.861
V	dp-BREATH	0.142	0.861	0.151	0.875	0.147	0.882	0.158	0.857
	minIP-DL	0.137	0.884	0.146	0.899	0.144	0.901	0.153	0.876
	BPRM	0.131	0.906	0.142	0.913	0.139	0.920	0.148	0.894
	EPLME	0.126	0.925	0.138	0.934	0.132	0.938	0.144	0.915

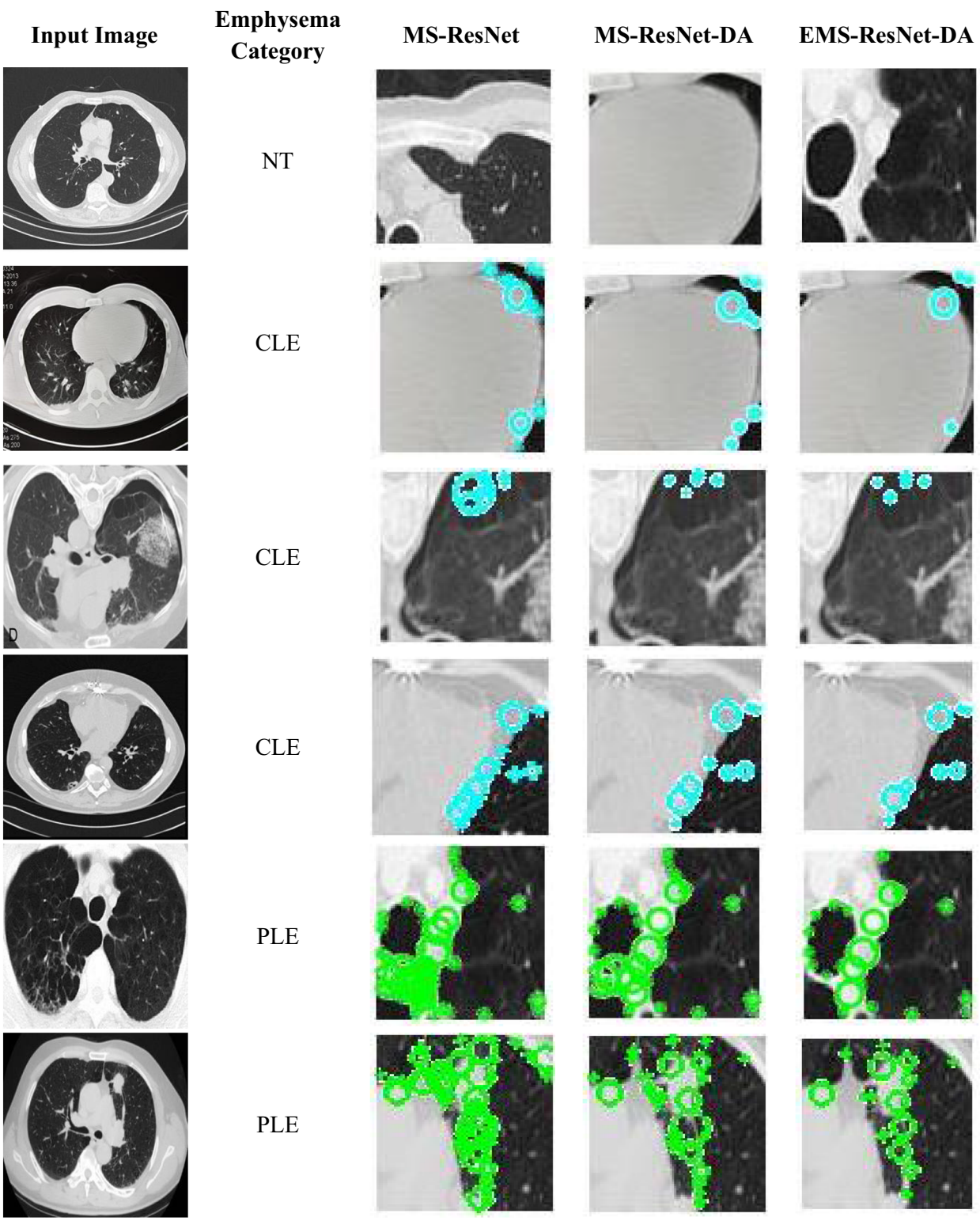


Fig. 7 Experimental results of classified emphysema categories for different input image samples

TP (True Positive) defines the outcome that the model accurately classifies the emphysema types as itself, e.g., CLE is classified as CLE. TN (True Negative) defines the outcome that the model the classifier accurately classifies the NE as NE. FP (False Positive) defines the outcome that the model inaccurately classifies the emphysema (CLE/PSE/PLE) as NE. FN (False Negative) defines the outcome

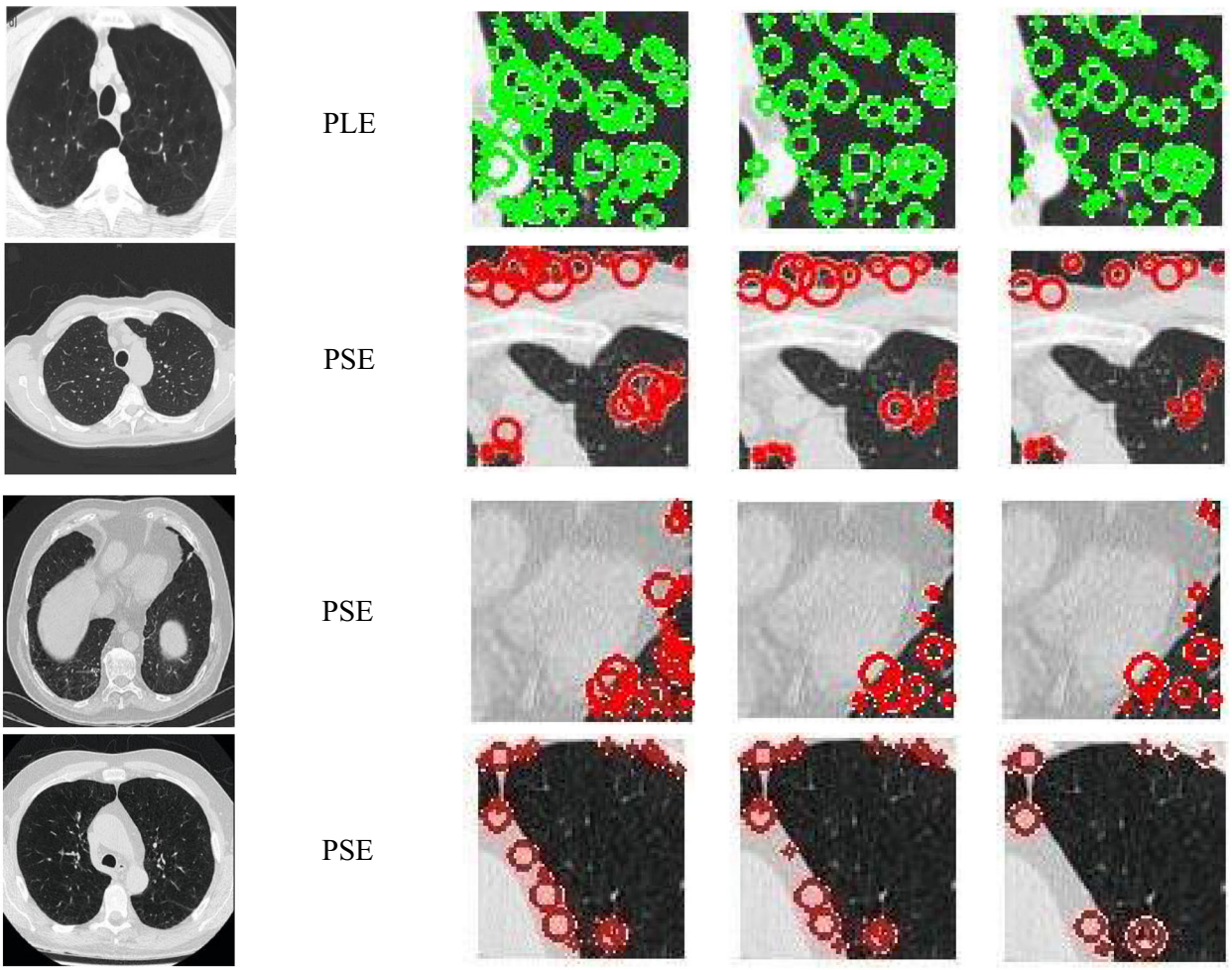


Fig. 7 continued

that the model inaccurately classifies the NE as emphysema.

Figure 8 portrays the accuracy (in %) attained by a Texton-based model [11], Joint Weber-based Rotation Invariant Uniform Local Ternary Pattern (JWRIULTP)

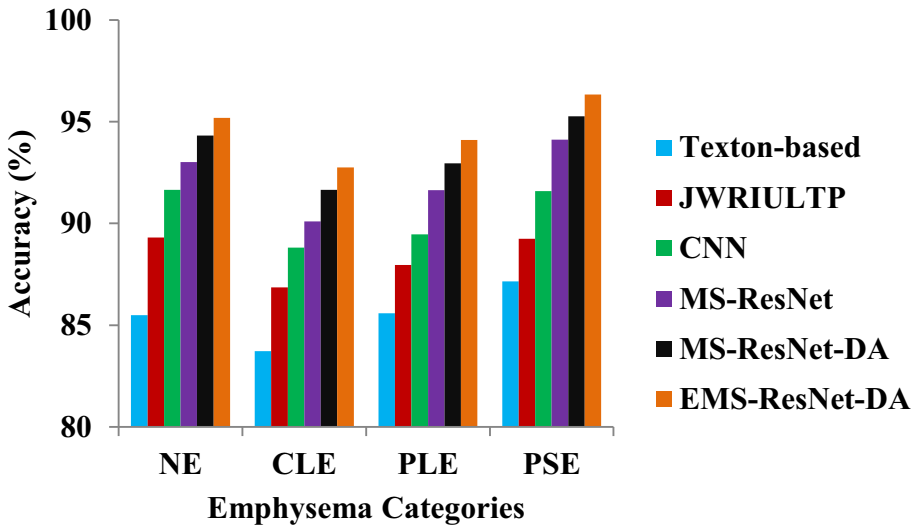


Fig. 8 Accuracy of each category of emphysema

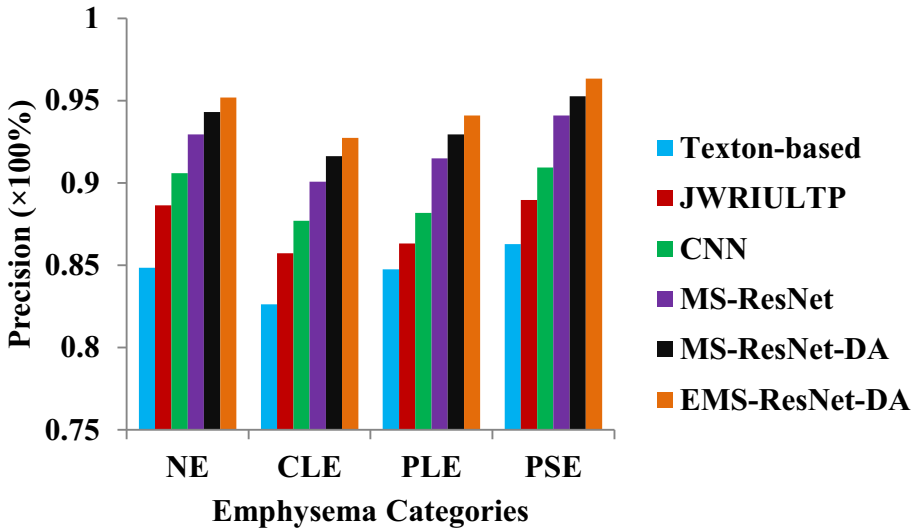


Fig. 9 Precision of each category of emphysema

[8], CNN [14], MS-ResNet [16], MS-ResNet-DA and EMS-ResNet-DA models for classifying different emphysema categories. It analyzes that the accuracy of EMS-ResNet-DA for each category of emphysema is higher than the other classification models since it increases the quantity of learning examples and estimates the accurate location of emphysema pixels for classification. For the case of PLE classification, the accuracy of EMS-ResNet-DA is 9.96% greater than the Texton-based model, 7% greater than the JWRIULTP, 5.18% greater than the CNN, 2.7% greater than the MS-ResNet and 1.24% greater than the MS-ResNet-DA. Thus, it proves that the proposed models can increase the accuracy of classifying the emphysema types compared to all other existing models.

7.2.2 Precision

It is the percentage of accurately categorized examples of emphysema at TP and FP rates.

Figure 9 shows the precision values obtained by the Texton-based model [11], JWRIULTP [8], CNN [14], MS-ResNet [16], MS-ResNet-DA and EMS-ResNet-DA models to classify the types of emphysema. It notices that the precision of EMS-ResNet-DA for every type of emphysema is greater than the other classification models. For example, the precision of EMS-ResNet-DA for classifying the PLE category is 11.03% higher than the Texton-based model, 9.01% higher than the JWRIULTP, 6.71% higher than the CNN, 2.85% higher than MS-ResNet, and 1.24% higher than MS-ResNet-DA model. Therefore, it is verified that the proposed models maximize the precision of emphysema classification compared to all other existing models.

7.2.3 Recall

It is the percentage of accurately categorized examples of emphysema at TP and FN rates.

$$\text{Precision} = \frac{\text{No. of exactly classified emphysema samples}}{\text{No. of exactly classified emphysema samples} + \text{No. of inexactly classified emphysema samples}} \tag{11}$$

$$\text{Recall} = \frac{\text{No. of exactly classified emphysema samples}}{\text{No. of exactly classified emphysema samples} + \text{No. of inexactly classified NE samples}} \tag{12}$$

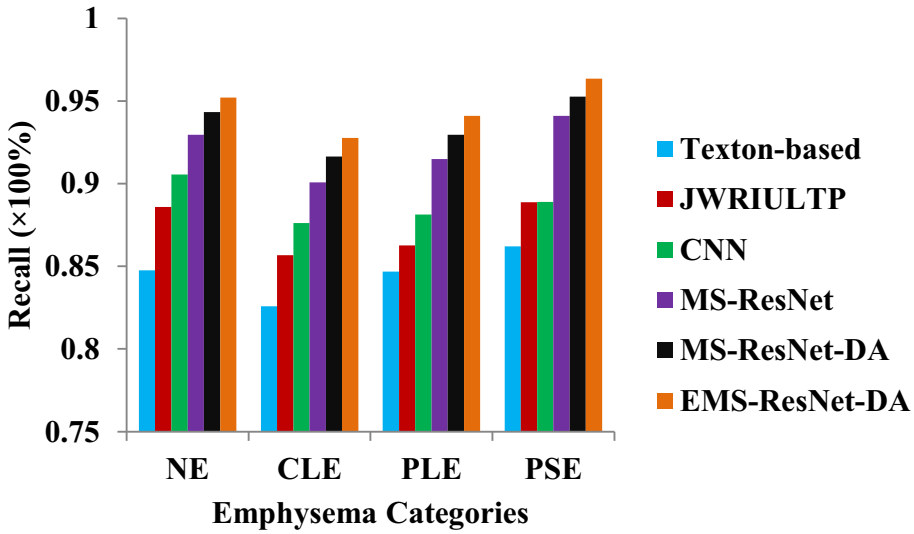


Fig. 10 Recall of each category of emphysema

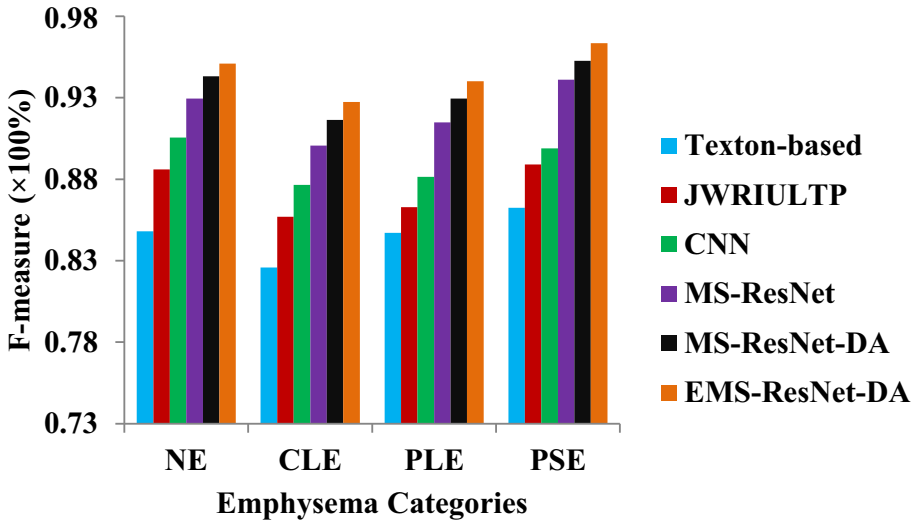


Fig. 11 F-measure of each category of emphysema

Figure 10 shows the recall values for Texton-based model [11], JWRIULTP [8], CNN [14], MS-ResNet [16], MS-ResNet-DA and EMS-ResNet-DA models to classify the types of emphysema. It addresses that the recall of EMS-ResNet-DA for every type of emphysema is increased compared to the other models. For example, the recall of EMS-ResNet-DA for classifying the PLE category is 11.11% higher than the Texton-based model, 9.09% higher than the JWRIULTP, 6.76% higher than the CNN, 2.85% higher than MS-ResNet and 1.23% higher than MS-ResNet-DA classification models. So, it is proved that the proposed models can achieve better recall compared to all other existing models.

7.2.4 F-measure

It is determined by

$$F-\text{measure} = 2 \times \frac{(\text{Precision} \bullet \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{13}$$

Figure 11 illustrates the F-measure of Texton-based model [11], JWRIULTP [8], CNN [14], MS-ResNet [16], MS-ResNet-DA and EMS-ResNet-DA models to classify the types of emphysema. It observes that the F-measure of EMS-ResNet-DA for each category of emphysema is increased compared to the other models. For example, the F-measure of EMS-ResNet-DA to classify the PLE category is 10.98% greater than the Texton-based model, 8.96% greater than the JWRIULTP, 6.65% greater than the CNN, 2.77% and 1.14% greater than MS-ResNet and MS-ResNet-DA, respectively. So, it is concluded that the EMS-ResNet-DA increases the f-measure of classifying the emphysema categories compared to all other existing models.

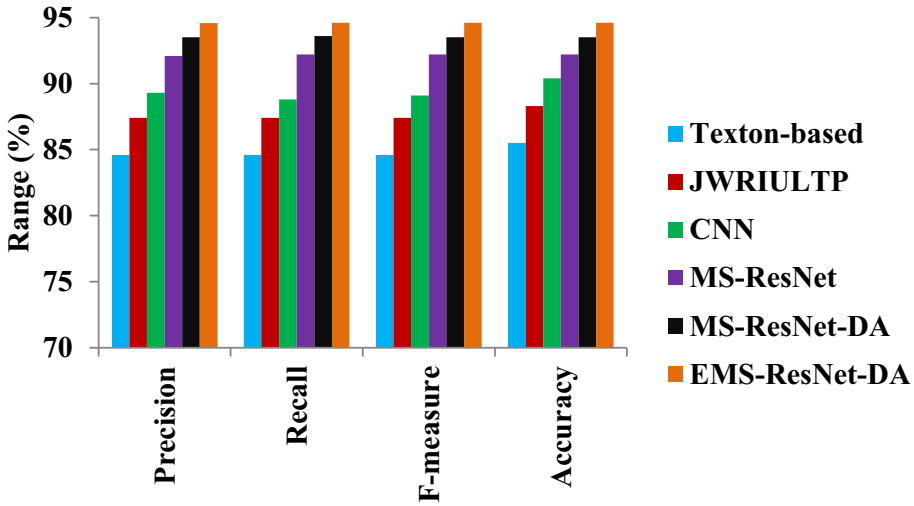


Fig. 12 Summary of the overall performance of proposed and existing models

Moreover, Fig. 12 displays the evaluation between the EMS-ResNet-DA model and a few existing models: a Texton-based model [11], JWRIULTP [8], CNN [14], MS-ResNet [16] and MS-ResNet-DA to interpret its improvement on overall classification accuracy. It is summarized that the overall classification accuracy of EMS-ResNet-DA is 10.6% higher than the Texton-based model, 7.13% higher than the JWRIULTP, 4.65% higher than the CNN, 2.6% higher than the MS-ResNet and 1.2% higher than the MS-ResNet-DA-based classification models. So, it is concluded that the EMS-ResNet-DA model achieves higher efficiency to classify the types of emphysema CT scans.

8 Conclusion

In this study, an MS-ResNet-DA model is proposed, wherein GAN is introduced to increase the four different types of emphysema CT images together with the random normal vector as inputs to the network. The created training samples are again augmented by a variety of methods such as cropping, duplication, rotation and scaling. Then, this model is enhanced by an EPLME algorithm to estimate the relative and absolute locations of emphysema pixels in an image according to the region, edge and vertex. The relative location information offers the current location of the emphysema pixel by adding the relative dislocation measures extracted from two lung disease images. Additionally, the absolute location estimation is performed using the position encoding network for matching the diseased image with the reference emphysema images and validating whether location information is implicitly learned when trained on categorical labels. Once all the pixel location maps of emphysema images are obtained, these are fed to the MS-ResNet along with the actual CT scans for classifying the categories of emphysema accurately. To conclude, the experimental results proved that the EMS-ResNet-DA model has an average accuracy of 94.6% compared to all existing classification models. But, it does not consider the feature's significance during concatenating the features extracted by different scales of input images. Therefore, the future extension of this work will focus on finding the significance of the features obtained from the convolutional layers for enhancing classification effectiveness.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Salvi S, Kumar GA, Dhaliwal RS, Paulson K, Agrawal A, Koul PA, Christopher DJ (2018) The burden of chronic respiratory diseases and their heterogeneity across the states of India: the global burden of disease study 1990–2016. Lancet Glob Health 6(12):1363–1374
2. Hochhegger B, Langer FW, Irion K, Souza A, Moreira J, Baldisserotto M, Pallaoro Y, Muller E, Medeiros TM, Altmayer S, Marchiori E (2019) Pulmonary acinus: understanding the computed tomography findings from an acinar perspective. Lung 197(3):259–265
3. Lynch DA, Austin JH, Hogg JC, Grenier PA, Kauczor HU, Bankier AA, Coxson HO (2015) CT-definable subtypes of chronic obstructive pulmonary disease: a statement of the Fleischner society. Radiology 277(1):192–205
4. Lynch DA, Moore CM, Wilson C, Nevrekar D, Jennermann T, Humphries SM, Regan EA (2018) CT-based visual classification of emphysema: association with mortality in the COPDGene study. Radiology 288(3):859–866

5. Stockley RA, Halpin DM, Celli BR, Singh D (2019) Chronic obstructive pulmonary disease biomarkers and their interpretation. Am J Respir Crit Care Med 199(10):1195–1204
6. Yang J, Angelini ED, Smith BM, Austin JH, Hoffman EA, Bluemke DA, Laine AF (2016) Explaining radiological emphysema subtypes with unsupervised texture prototypes: MESA COPD study. Medical Computer Vision and Bayesian and Graphical Models for Biomedical Imaging. Springer, Cham, pp 69–80
7. Binder P, Batmanghelich NK, Estepar RSJ, Golland P (2016) Unsupervised discovery of emphysema subtypes in a large clinical cohort. International Workshop on Machine Learning in Medical Imaging. Springer, Cham, pp 180–187
8. Peng L, Lin L, Hu H, Ling X, Wang D, Han X, Chen YW (2017) Joint weber-based rotation invariant uniform local ternary pattern for classification of pulmonary emphysema in CT images. In: IEEE international conference on image processing, pp 2050–2054
9. Sorensen L, Shaker SB, De Bruijne M (2010) Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE Trans Med Imaging 29(2):559–569
10. Mendoza CS, Washko GR, Ross JC, Diaz AA, Lynch DA, Crapo JD, Estépar RSJ (2012) Emphysema quantification in a multi-scanner HRCT cohort using local intensity distributions. In 9th IEEE international symposium on biomedical imaging, pp 474–477
11. Yang J, Feng X, Angelini ED, Laine AF (2016) Texton and sparse representation based texture classification of lung parenchyma in CT images. In: 38th Annual international conference of the IEEE engineering in medicine and biology society, pp 1276–1279
12. Xu Y, van Beek EJ, Hwanjo Y, Guo J, McLennan G, Hoffman EA (2006) Computer-aided classification of interstitial lung diseases via MDCT: 3D adaptive multiple feature method (3D AMFM). Acad Radiol 13(8):969–978
13. Vasconcelos V, Barroso J, Marques L, Silvestre Silva J (2015) Enhanced classification of interstitial lung disease patterns in HRCT images using differential lacunarity. Biomed Res Int 2015:1–9
14. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S (2016) Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imaging 35(5):1207–1216
15. Cheplygina V, Pena IP, Pedersen JH, Lynch DA, Sørensen L, de Bruijne M (2017) Transfer learning for multicenter classification of chronic obstructive pulmonary disease. IEEE J Biomed Health Inform 22(5):1486–1496
16. Peng L, Lin L, Hu H, Li H, Chen Q, Ling X, Chen YW (2019) Classification and quantification of emphysema using a multi-scale residual network. IEEE J Biomed Health Inform 23(6):2526–2536
17. Gao M, Bagci U, Lu L, Wu A, Buty M, Shin HC, Xu Z (2018) Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. Comput Methods Biomech Biomed Eng Imaging Vis 6(1):1–6
18. Humphries SM, Notary AM, Centeno JP, Lynch DA (2018) Automatic classification of centrilobular emphysema on CT using deep learning: comparison with visual scoring. Image analysis for moving organ, breast, and thoracic images. Springer, Cham, pp 319–325
19. Pino Peña I, Cheplygina V, Paschaloudi S, Vuust M, Carl J, Weinreich UM, de Bruijne M (2018) Automatic emphysema detection using weakly labeled HRCT lung images. PLoS ONE 13(10):1–16
20. Ge Y, Wang Q, Wang L, Wu H, Peng C, Wang J, Yi Y (2019) Predicting post-stroke pneumonia using deep neural network approaches. Int J Med Inform 132:1–32
21. Gupta N, Gupta D, Khanna A, RebouçasFilho PP, de Albuquerque VHC (2019) Evolutionary algorithms for automatic lung disease detection. Measurement 140:590–608
22. Jimenez-Carretero D, Bermejo-Peláez D, Nardelli P, Fraga P, Fraile E, Estépar RSJ, Ledesma-Carbayo MJ (2019) A graph-cut approach for pulmonary artery-vein segmentation in noncontrast CT images. Med Image Anal 52:144–159
23. Joyseeree R, Otálora S, Müller H, Depeursinge A (2019) Fusing learned representations from Riesz filters and deep CNN for lung tissue classification. Med Image Anal 56:172–183
24. Xu M, Qi S, Yue Y, Teng Y, Xu L, Yao Y, Qian W (2019) Segmentation of lung parenchyma in CT images using CNN trained with the clustering algorithm generated dataset. Biomed Eng Online 18(1):1–21
25. Surendar P, Ponni M (2021) Diagnosis of lung cancer using hybrid deep neural network with adaptive sine cosine crow search algorithm. J Comput Sci 53:101374. https://doi.org/10.1016/j.jocs.2021.101374
26. Tang C, Plasek JM, Zhang H, Kang MJ, Sheng H, Xiong Y, Zhou L (2019) A temporal visualization of chronic obstructive pulmonary disease progression using deep learning and unstructured clinical notes. BMC Med Inform Decis Mak 19(8):1–9
27. Liu H, Wang L, Nan Y, Jin F, Wang Q, Pu J (2019) SDFN: segmentation-based deep fusion network for thoracic disease classification in chest X-ray images. Comput Med Imaging Graph 75:66–73
28. Balamurugan D, Aravinth SS, Reddy PCS et al (2022) Multiview objects recognition using deep learning-based Wrap-CNN with voting scheme. Neural Process Lett 54:1495–1521. https://doi.org/10.1007/s11063-021-10679-4
29. Cazzolato MT, Scabora LC, Nesso-Jr MR, Milano-Oliveira LF, Costa AF, Kaster DS, Traina AJ (2019) dp-breath: Heat maps and probabilistic classification assisting the analysis of abnormal lung regions. Comput Methods Programs Biomed 173:27–34
30. Ke Q, Zhang J, Wei W, Połap D, Woźniak M, Kośmider L, Damaševičius R (2019) A neuro-heuristic approach for recognition of lung diseases from X-ray images. Expert Syst Appl 126:218–232
31. Gerard SE, Herrmann J, Kaczka DW, Musch G, Fernandez-Bustamante A, Reinhardt JM (2020) Multi-resolution convolutional neural networks for fully automated segmentation of acutely injured lungs in multiple species. Med Image Anal 60:1–41
32. Annamalai M, Muthiah P (2022) An early prediction of tumor in heart by cardiac masses classification in echocardiogram images using robust back propagation neural network classifier. Braz Arch Biol Technol. https://doi.org/10.1590/1678-4324-2022210316
33. Nagaraj Y, Wisselink HJ, Rook M, Cai J, Nagaraj SB, Sidorenkov G, Veldhuis R, Oudkerk M, Vliegenthart R, van Ooijen P (2022) AI-driven model for automatic emphysema detection in low-dose computed tomography using disease-specific augmentation. J Digit Imaging 35:1–13
34. Lor KL, Chang YC, Yu CJ, Wang CY, Chen CM (2022) Bullous parametric response map for functional localization of COPD. J Digit Imaging 35:1–12
35. https://lauge-soerensen.github.io/emphysema-database/

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