Secondary Pulmonary Tuberculosis Recognition by 4-Direction Varying-Distance GLCM and Fuzzy SVM

Yu-Dong Zhang1,2 · Wei Wang2 · Xin Zhang3 · Shui-Hua Wang1,2

Accepted: 15 November 2021 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Tuberculosis (TB) is an infectious disease caused by Mycobacterium tuberculosis. Our study investigates the recognition of secondary pulmonary (SPTB). A novel F3 model is proposed. The first F means using a four-direction varying-distance gray-level co-occurrence matrix (FDVDGLCM) to analyze the chest CT images; the second F means a five-property feature set (FPFS) from the FDVDGLCM results; the third F means fuzzy support vector machine (FSVM). Besides, a slight adaption of multiple-way data augmentation is used to boost the training set. The 10 runs of 10-fold cross-validation demonstrate that this F3 model achieves a sensitivity of 93.68% ± 1.75%, a specificity of 94.17% ± 1.68%, a precision of 94.17% ± 1.55%, an accuracy of 93.92% ± 1.05%, an F1 score of 93.91% ± 1.07%, an MCC of 87.88% ± 2.09%, and an FMI of 93.92% ± 1.06%. The AUC is 0.9624. The FSVM can give better performance than ordinary SVM. The proposed F3 model is superior to six state-of-the-art SPTB recognition models.

Keywords

Secondary pulmonary tuberculosis · Gray-level co-occurrence matrix · Varying-distance · Support vector machine · Fuzzy membership function · Fuzzy support vector machine · Data augmentation

1 Introduction

Tuberculosis (TB) is an infectious disease caused by Mycobacterium tuberculosis [1]. TB outside of the lung is called extrapulmonary TB, such as bones or other organs. TB infects the lungs typically; thus is named pulmonary TB (PTB) [2].

Typical symptoms of PTB are high temperature [3], night sweats, extreme fatigue, weight loss, poor appetite, etc. Respiratory signs comprise hemoptysis, cough, expectoration, chest tightness, chest pain, or dyspnea in different grades. In the world, PTB is one of the top ten fatality causes and the foremost cause from a single contagious agent.

PTB can be classified into primary PTB and secondary PTB (SPTB) [4]. we fix on SPTB. The most common locations of SPTB lesions: apical and posterior segments of the upper lobe and dorsal segments of the lower lobe. CT manifestations: Exudative patchy shadows of varying density, miliary shadows, tuberculosis bulbs, caseous pneumonia, proliferative nodular shadows, satellite foci, cavities, tree bud signs, bronchial dissemination, cord strips and calcifications, and numerous other forms.

In current years, experts are prone to harness artificial intelligence (AI) methods to create smart SPTB diagnosis models. For instance, Bagci, Kubler, Luna, Jain, Bishai and Mollura (2013) [5] created a computer-aided detection and quantification model for cavitary TB. Based on
a support vector machine (SVM), the authors proposed a shape-based SVM (SSVM) to discover airways and cavities. Li, Huang and Jin (2018) [6] presented an AE-CNN classifier, which stands for autoencoder and convolutional neural network. James-Reynolds, Currie and Gao (2020) [7] proposed a depth-ResNet for TB severity level analysis. Xie, Wu, Han, Wang, Wu, Cui, Feng, Zhu and Chen (2020) [8] offered a computer-aided system (CAS) for the detection of TB in radiographs. Based on convolutional neural network (CNN), Zhang, Nayak, Zhang and Wang (2020) [9] presented an improved CNN (ICNN) for TB recognition. Rajpurkar, O’Connell, Schechter, Asnani, Li, Kiani, Ball, Mendelson, Maartens, van Hoving, Griesel, Ng, Boyles and Lungren (2020) [10] presented a CheXaid model to assist physicians in making the diagnosis of TB.

Nevertheless, the performances of the above methods still have rooms to improve. In order to recognize SPTB more accurately, we propose a new SPTB recognition system in this study. The motivation is that gray-level co-occurrence matrix (GLCM)—an image texture descriptor—may produce a measure of chest CT images to distinguish SPTB from healthy people’s scanning. The contributions of our method are five points:

(i) A 4-direction varying-distance GLCM (FDVDGLCM) is introduced as the feature extractor.
(ii) A 5-property feature set is extracted from FDVDGLCM matrix.
(iii) The fuzzy support vector machine (FSVM) is harnessed as the classifier.
(iv) A slight adaption of multiple-way data augmentation is used to boost the training set.
(v) Multiple runs of cross-validation results show that our model is superior to six state-of-the-art models.

2 Dataset and preprocessing

The raw dataset $V_1$ in this study is from Ref. [9], in which 66 SPTB and 66 healthy control (HC) are enrolled, and 144 SPTB images and 144 HC images are generated from the 3D chest CT scanning. Table 1 shows the demographics of our dataset.

| Class  | No. of subjects | No. of images |
|--------|----------------|--------------|
| SPTB   | 68             | 144          |
| HC     | 66             | 144          |

Figure 2a, c show one sample of SPTB and one sample of HC. Figure 2b delineates the lesions within five red ellipses. Multiple nodule shadows of different sizes are seen in the posterior segment of the upper lobe tip of both lungs.
3 Methodology

3.1 Gray-level co-occurrence matrix

Table 2 presents the acronym list for ease of reading. Suppose we have a gray-level image $I$ with the size of $m \times n$ and $p$ different pixel values, the gray-level co-occurrence matrix (GLCM) computes how often pairs of pixels with a specific value and offset occur. GLCM is a powerful approach to texture analysis with loads of medical applications such as lung cancer classification [11], similar to our task.

The GLCM matrix $G$ is defined as

$$G_{\Delta x, \Delta y}(i,j) = \sum_{x=1}^{m} \sum_{y=1}^{n} \begin{cases} 1 & \text{if } I(x,y) = i \land I(x+\Delta x,y+\Delta y) = j \\ 0 & \text{otherwise} \end{cases}$$

where $\land$ means and operation; $i$ and $j$ are pixel values; $x$ and $y$ are spatial positions in the image $I$; $(\Delta x, \Delta y)$ are the offsets defining the spatial relation. Note the GLCM matrix is a $p \times p$ matrix, i.e., $G_{\Delta x, \Delta y} \in \mathbb{R}^{p \times p}$. Some advanced GLCM techniques, such as local-GLCM [12], may be tested in our future study.

3.2 Four-direction mechanism

In practice, the $(x, y)$ are often replaced by row and column coordinates $(r, c)$. Subsequently, the offset is expressed in the form of $(\Delta r, \Delta c)$. As can be seen here, the offset $[0, 1]$
corresponds to $0^\circ$, similarly, $[-1, 1]$, $[-1, 0]$, and $[-1, -1]$ correspond to $45^\circ$, $90^\circ$, and $135^\circ$, respectively. The directional angles of GLCM in our study are set to the following four directional angles.

$$\theta = [0^\circ, 45^\circ, 90^\circ, 135^\circ]$$

(2)

Other angles $[180^\circ, 225^\circ, 270^\circ, 315^\circ]$ can be chosen, but they will give the same results as $[0^\circ, 45^\circ, 90^\circ, 135^\circ]$, respectively. That is the reason why we only choose this four-direction mechanism (Fig. 3).

The following Fig. 4 shows an example of four-direction GLCM (4D-GLCM). For example, Fig. 4a shows the element $G_{0, 1}(3, 1) = 1$, because there is one instance in the left-hand image where one horizontally adjacent pixels have the values of 3 and 1. The element $G_{0, 1}(7, 4) = 2$ is since there exist two occurrences in the raw image where two horizontally adjacent pixels have the values of 7 and 4.

### 3.3 The varying-distance mechanism

All the above four offsets have a distance $d$ of 1, as clearly shown in Fig. 3. Srivastava, Rajitha, Agarwal and Singh (2020) [13] introduced the varying-distance mechanism. In their paper, $d$ ranges from 1 to the width of the image. Nevertheless, we find by the trial-and-error method that $d$ assigned with a value larger than three does not improve the performance.

Figure 5 shows the varying distance where the maximum distance if $D = 3$. Hence, we can write the pseudocode of our four-direction varying-distance GLCM (FDVDGLCM) in Algorithm 2. Note, the offset $(\Delta r, \Delta c)$ can be expressed at the form of $(d, \theta)$ and each GLCM matrix $G_{d, \theta} \in \mathbb{R}^{p \times p}$. The final overall GLCM matrix $G \in \mathbb{R}^{p \times p \times 4 \times D}$ can be expressed as

$$G = \sum_{d=1}^{D} C_{\theta=(0,45,90,135)} G_{d, \theta},$$

(3)
where \( C \) means the concatenation function.

**Algorithm 2 Pseudocode of FDVDGLCM**

Input an image \( I \).

for \( d = 1 : 1 : D \)

   for \( \theta = 0, 45, 90, 135 \)
      
      Calculate the corresponding offset \( (d, \theta) \). See Figure 5.

      Calculate and record the corresponding GLCM matrix \( G_{d,\theta} \).
   
end

Concatenate \( G_{d,\theta} \) over all offsets and form the overall GLCM matrix \( \mathcal{G} \). See Eq. (3).

Output \( \mathcal{G} \).

Figure 6 displays a toy example of varying distance in which \( d = 1, 2, 3 \). As we can observe, the larger the distance \( d \) is, the GLCM will analyze pairs of voxels further away from each other.

### 3.4 Five-property feature set of FDVDGLCM

Based on FDVDGLCM, we calculate a five-property feature set (FPFS). This study sets the number of properties as \( N_F = 5 \). First is the contrast, indicating the linear dependency of gray levels of two neighboring pixels. For constant image, the contrast is 0.

**Contrast** \( d,\theta \) is defined as

\[
\text{Contrast}_{d,\theta} = \sum_{i=1}^{p} \sum_{j=1}^{p} (i - j)^2 G_{d,\theta}(i,j)
\]  

Second is the correlation, measuring how correlated a pixel is to its neighbor over the whole GLCM image.

**Correlation** \( d,\theta \) is defined as

\[
\text{Correlation}_{d,\theta} = \sum_{i=1}^{p} \sum_{j=1}^{p} \frac{(i - \mu_i)(j - \mu_j)G_{d,\theta}(i,j)}{\sigma_i \sigma_j}
\]  

where \( \mu_i \) and \( \mu_j \) are the GLCM mean defined as

![Fig. 6 A toy example of varying-distance \( d = 1, 2, 3 \)](image)

![Fig. 7 Schematic of V-fold cross-validation](image)
\[ \mu_i = \sum_{i=1}^{p} \sum_{j=1}^{p} i \times G_{d,\theta}(i,j) \] (6)

\[ \mu_j = \sum_{i=1}^{p} \sum_{j=1}^{p} j \times G_{d,\theta}(i,j) \] (7)

The GLCM variances \( \sigma_i \) and \( \sigma_j \) are defined as:

\[ \sigma_i = \sqrt{\sum_{i=1}^{p} \sum_{j=1}^{p} (i - \mu_i)^2 G_{d,\theta}(i,j)} \] (8)

\[ \sigma_j = \sqrt{\sum_{i=1}^{p} \sum_{j=1}^{p} (j - \mu_j)^2 G_{d,\theta}(i,j)} \] (9)

Note, correlation is within the range of \([-1, 1]\). The \(-1\) stands for a perfectly negative correlation, while \(+1\) is a perfectly positive correlation.

The third is the energy property, calculating the sum of squared GLCM elements. It measures the textural uniformity of an image. Its value ranges from 0 to 1. The definition is:

\[ \text{Energy}_{d,\theta} = \sum_{i=1}^{p} \sum_{j=1}^{p} G_{d,\theta}^2(i,j) \] (10)

The fourth is homogeneity, a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Again, its range is from 0 to 1. The definition is:

\[ \text{Homogeneity}_{d,\theta} = \sum_{i=1}^{p} \sum_{j=1}^{p} \frac{G_{d,\theta}(i,j)}{1 + |i - j|} \] (11)

The above four properties are commonly used to extract features in GLCM. However, Athertya and Kumar (2021) [14] stated that entropy-based measure obtained from GLCM alone is sufficient to detect focal changes, indicating entropy's importance. Hence, we introduce the final property

\[ \text{Entropy}_{d,\theta} = -\sum_{i=1}^{p} \sum_{j=1}^{p} G_{d,\theta}(i,j) \times \ln G_{d,\theta}(i,j) \] (12)

The five properties run on all the GLCM matrix \( G_{d,\theta} \) are concatenated and form the FPFS \( F \).

\[ F = \sum_{d=1}^{D} C_{\theta=0,45,90,135} \left\{ \begin{array}{l} \text{Contrast} \left( G_{d,\theta} \right) \\ \text{Correlation} \left( G_{d,\theta} \right) \\ \text{Energy} \left( G_{d,\theta} \right) \\ \text{Homogeneity} \left( G_{d,\theta} \right) \\ \text{Entropy} \left( G_{d,\theta} \right) \end{array} \right\}, \] (13)

### 3.5 Support vector machine

The FPFS is sent to the classifier. Note, we do not choose deep learning [15, 16] due to the inherent characteristics and dimensions of FPFS \( F \). Instead, we find that traditional classifier—the support vector machine (SVM) [17]—exhibits excellent performances in handling the FPFS describe in Section 3.4. Additionally, a variant of SVM—fuzzy SVM (FSVM)—is introduced to our method.

The SVM is a standard robust classifier [18]. Suppose there are two classes of positive and negative, and the dataset is

\[ \{ (g_n, t_n) \mid p_n \in \mathbb{R}^d \}, n = 1, \ldots, N, \] (14)

where \( N \) stands for the number of samples, \( g_n \) the feature vector of \( n \)-th sample data, \( d \) the dimension of input features, and \( t_n \) the corresponding target label.

\[ t_n = \begin{cases} +1 & \text{class = positive} \\ -1 & \text{class = negative} \end{cases} \] (15)

The goal of SVM [19] is to build a hyperplane with the dimensions of \((d - 1)\) that can part the positive class from the negative class [20]. The SVM model is alike to logistic regression (LR), because both are driven by linear functions. The hyperplane of SVM can be written as:

\[ w^T g - b = 0, \] (16)

where \( w \) and \( b \) have analogous mathematical meanings as those in LR. \( w \) stands for the weights and \( b \) the bias. The solutions of \( w \) and \( b \) are:

\[ \begin{cases} \min_{\|w\|^2} \|w\|^2 \\ \text{s.t.} \ t_n (w^T g_n - b) \geq 1, n = 1, \ldots, N \end{cases} \] (17)

In cases that input features \( g_n \) encloses noises, we utilize the “soft margin” method, which adds the positive slack vector \( \rho \)

\[ \rho = (\rho_1, \rho_2, \ldots, \rho_n, \ldots, \rho_N). \] (18)

Equation (17) is then rephrased as:

\[ \begin{cases} \min_{w,\rho} \|w\|^2 + Le^T \rho \\ \text{s.t.} \ t_n (w^T g_n - b) \geq 1 - \rho_n, n = 1, \ldots, N \end{cases} \] (19)

where \( e \) stands for a vector of ones of \( N \)-dimension and \( L \) the error penalty [21].

The constraint optimization problem in Eq. (19) is solved by the Lagrange multiplier [22] method as:

\[ \begin{cases} \max_{\rho \geq 0} \left\{ \frac{1}{2} \|w\|^2 + Le^T \rho - \sum_{i=1}^{N} \rho_i \left| t_i g_i - b_i \right| - \rho \right\} \end{cases} \] (20)

The dual form (DF) [23] is utilized to elucidate Eq. (20) because the min-max optimization is challenging. We get:

\[ \begin{cases} \max_{\rho \geq 0} \sum_{n=1}^{N} \rho_n - \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \rho_n \rho_m \left( t_n^T g_m - t_m^T g_n \right) \\ \text{s.t.} \ \sum_{n=1}^{N} t_n \rho_n = 0, n = 1, \ldots, N \end{cases} \] (21)
3.6 Fuzzy membership function

The fuzzy membership function (FMF) [24] is introduced to each training sample \((g_n, t_n)\). Note, the raw training sample includes only the input and target [25], while the fuzzy sample data comprises the input, the target, and the fuzzy altitude (FA) [26]. Mathematically, Eq. (14) is generalized to
\[
\left\{ \left( (p_n, t_n, a_n) \big| p_n \in \mathbb{R}^d \right), n = 1, \ldots, N \right. \\
\text{s.t.} \quad 0 < a_n \leq 1
\]

where \(a_n\) is the FA of the sample \(g_n\) towards class \(t_n\).

Let the mean of class +1 be \(m_+\) and the mean of class −1 be \(m_-\). We have [27]:
\[
m_+ = \text{mean}(g_n | t_n = +1),
\]
\[
m_- = \text{mean}(g_n | t_n = -1).
\]

We can deduce the radii of class +1 and class −1 as:
\[
r_+ = \max_{\{g_n, t_n = +1\}} \left| m_+ - g_n \right|.
\]
\[
r_- = \max_{\{g_n, t_n = -1\}} \left| m_- - g_n \right|.
\]

The FMF is described on the means and radii of two classes \(t_n = +1\) and \(t_n = -1\):
\[
a_n = \begin{cases} 
1 - \frac{|m_+ - g_n|}{r_+ + \nu} & t_n = +1 \\
1 - \frac{|m_- - g_n|}{r_- + \nu} & t_n = -1
\end{cases}, n = 1, 2, \ldots, N
\]

Here \(\nu\) is a positive parameter, to make sure the FA \(a_n\) is larger than zero, viz., \(a_n > 0\).

3.7 Fuzzy SVM

Assume vector \(a\) is the fuzzy membership vector [28] of all samples:
\[
a = (a_1, a_2, \ldots, a_n, \ldots, a_N)
\]

We can deduce the hyperplane [29] of FSVM as
\[
\min_{w, b, \rho} \left\{ \frac{1}{2} \|w\|^2 + L \rho \right. \\
\left. \text{s.t.} \quad t_n (w^T g_n - b) \geq 1 - \rho_n, n = 1, \ldots, N \right.
\]

Note, a smaller \(a_n\) decreases the influence of the slack vector \(\rho_n\) so that the corresponding sample \(g_n\) is deemed less substantial. The Lagrangian [30] is attained similarly as
\[
\min_{w, b, \rho} \left\{ \frac{1}{2} \|w\|^2 + L \rho - \sum_{n=1}^N \varphi_n [t_n (w^T g_n - b) + \rho_n - 1] - \sum_{n=1}^N \rho_n t_n \right\}
\]

Its DF is attained as
\[
\left\{ \begin{array}{ll}
\max_{\varphi} \sum_{n=1}^N \varphi_n - \frac{1}{2} \sum_{m=1}^N \sum_{n=1}^N \varphi_m \varphi_n t_m^T g_n t_n & \\
\text{s.t.} \quad 0 \leq \varphi_n \leq a_n L, n = 1, \ldots, N
\end{array} \right.
\]

3.8 Implementation of our F3 model

If we summarize all the techniques, the pseudocode of our F3 model can be recapped in Algorithm 3. We dub our model F3. The first F stands for the FDVDGLCM, the second F for the FPFS, and the last for FSVM. From Algorithm 3, we observe that our F3 model can work in two phases. Within Phase I, the input is a dataset \(D\), and the output is the trained classifier \(M\). Within Phase II, the trained classifier \(M\) is harnessed to classify the category of the input image \(I\). The output is the label \(O\).

| Algorithm 3 Pseudocode of proposed F3 Model |
|---------------------------------------------|
| **Phase I** Training |
| Input | Dataset \(D\). |
| Step 1 | Analyze texture information via FDVDGLCM. |
| Step 2 | Extract properties via FPFS. |
| Step 3 | Train the classifier via the FSVM. |
| Output | The trained classifier \(M\). |
| **Phase II** Prediction |
| Input | Image \(I\). |
| Step 1 | Analyze texture information via FDVDGLCM. |
| Step 2 | Extract properties via FPFS. |
| Step 3 | Predict the category of the image \(I\) via trained classifier \(M\). |
| Output | Label \(O\). |

3.9 Cross-validation and multiple-way data augmentation

V-fold cross-validation is employed to run our F3 model due to the relatively small size of our dataset. Hold-out validation is not chosen since it works well for a larger dataset. In \(r\)-th run (\(1 \leq r \leq R\)) of the V-fold cross-validation, the whole dataset \(D\) (144 SPTB image and 144 HC images) is divided into V folds.
\[
D \mapsto \{D_r(1), D_r(2), \ldots, D_r(v), \ldots, D_r(V)\}, r = 1, 2, \ldots, R
\]

where \(D_r(v)\) stands for the \(v\)-th fold of the whole dataset at \(r\)-th run.
At \( v \)-th (\( 1 \leq v \leq V \)) trial, the \( v \)-th fold is pinched out as the test set \([31]\), and the rest \( V-1 \) folds are selected as training set:

\[
\begin{align*}
\text{Training Set} & \{ D_v(1), \ldots, D_v(v-1), D_v(v+1), \ldots, D_v(V) \} \\
\text{Test Set} & \{ D_v(v) \} \\
, v = 1, 2, \ldots, V, r = 1, 2, \ldots, R
\end{align*}
\tag{33}
\]

Note here; the training set is augmented via the multiple-way data augmentation (MDA) method described in Ref. [9]. We perform a slight adaption to the MDA. In the original MDA [9], \( N_D = 7 \) different data augmentation (DA) \([32]\) techniques are employed: rotation, Gamma correction, scaling, random translation, noise injection, horizontal and vertical shear. Nevertheless, each data augmentation in their paper generates a different number of new images. This study makes all the \( ND = 7 \) data augmentation generates the same number of new \( NN = 30 \) images. Suppose a given image is \( d \), the MDA brings an augmented set as

\[
d \mapsto \begin{cases}
  d & h(d) \\
  f_{DA}^1(d) & f_{DA}^1[h(d)] \\
  \vdots & \vdots \\
  f_{DA}^{N_D}(d) & f_{DA}^{N_D}[h(d)] \\
  N_N & N_N 
\end{cases},
\tag{34}
\]

where \( f_{DA} \) denotes a type of DA method and \( h(d) \) is the horizontally mirrored image of \( d \).

The F3 model is trained on the augmented training set. The trained model is dubbed \( M(r, v) \), and the corresponding confusion matrix is dubbed \( C(r, v) \). After all the \( V \)-fold trials, the confusion matrix of \( r \)-th run is summarized as

\[
C(r) = \sum_{v=1}^{V} C(r, v)
\tag{35}
\]

where \( K \) indicators \( I(r, k) \), \( k = 1, 2, \ldots, K \) are deduced from the confusion matrix, which will be explained in the next section. Based on \( R \) runs, the mean and standard deviation (MSD) \([33]\) of all \( K \) indicators are calculated as the form of \( a(k) \pm b(k) \), which is defined as:

\[
\begin{align*}
  a(k) &= \frac{1}{R} \times \sum_{r=1}^{R} I(r, k) \\
  b(k) &= \sqrt{\frac{1}{R-1} \times \sum_{r=1}^{R} [I(r, k) - a(k)]^2}
\end{align*}
\tag{36}
\]

Figure 7 shows the schematic of \( V \)-fold cross validation. This study let \( V = 10 \), viz., a ten-fold cross-validation is run. Besides, this \( V \)-fold cross-validation runs \( R \) times. At each run, the data division is reset randomly. Algorithm 4 recaps the pseudocode of \( R \)-run of \( V \)-fold cross-validation procedure.
### 3.10 Measures and indicators

$K=7$ indicators are defined. The SPTB is a positive class, while the HC is a negative class. The confusion matrix is defined as

$$C = \begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

(37)

The definitions of TP, FN, FP, and TN are listed in Table 3. Note, P stands for the actual positive class, so $P = TP + FN$. Similarly, N stands for the actual negative class; hence $N = FP + TN$.

Three common indicators—Sensitivity, Specificity, and Precision—are defined below:

$$\begin{align*}
\text{Sen} &= \frac{c_{11}}{c_{11} + c_{12}}, \\
\text{Spc} &= \frac{c_{22}}{c_{21} + c_{22}}, \\
\text{Prc} &= \frac{c_{11}}{c_{11} + c_{21}}.
\end{align*}$$

(38)

Accuracy [34] is defined as:

$$\text{Acc} = \frac{c_{11} + c_{22}}{c_{11} + c_{12} + c_{21} + c_{22}}.$$  

(39)

F1 score considers both the precision and the sensitivity [35]. It is the harmonic mean of the previous two measures: precision and sensitivity. F1 score [36] is defined as

$$F1 = \left( \frac{\text{Sen}^{-1} + \text{Prc}^{-1}}{2} \right)^{-1} = \frac{2 \times c_{11}}{2 \times c_{11} + c_{12} + c_{21}}.$$  

(40)

Two other indicators—Matthews correlation coefficient (MCC) [37] and Fowlkes–Mallows index (FMI)—are defined below:

$$\text{MCC} = \frac{c_{11} \times c_{22} - c_{21} \times c_{12}}{\sqrt{c_{11} + c_{21}} \times c_{11} + c_{12} \times c_{22} + c_{12} \times c_{22}}.$$  

(41)

$$\text{FMI} = \frac{c_{11}}{c_{11} + c_{21}} \times c_{11}.$$  

(42)

The receiver operating characteristic (ROC) curve [38] and the area under the curve (AUC) [39] are introduced to provide a graphical plot and a quantitative value of measuring the proposed F3 model, respectively. ROC and AUC are calculated via the following: (i) ROC plot is produced by charting the TP rate against the FP rate at different threshold degrees. (ii) AUC [40] is calculated by measuring the entire 2D area underneath the ROC curve from $(0, 0)$ to $(1, 1)$.

### 4 Experiment design

The experiments run on the programming platform of MATLAB 2021a. The parameters are listed in Table 4. The directional angles in GLCM are $\theta = [0^\circ, 45^\circ, 90^\circ, 135^\circ]$. The varying-distance ranges from 1 to $D = 3$. The number of properties extracted from FDVDGLCM is $N_F = 5$. Besides, $K=7$ indicators are deduced from the confusion matrix. The experiment runs a 10-fold cross-validation ten runs. $N_D = 7$ different data augmentation methods are utilized. Each DA generates $N_N = 30$ new images.

### 5 Results and discussions

#### 5.1 Result of MDA

Grab Fig. 2a as the raw training image $d$, Fig. 8a-g shows the results of seven different DA methods. The horizontally mirrored image $h(d)$ and its corresponding DA results are not displayed because of the page limit. As is observed from Fig. 8, MDA provides loads of diversity to the training set, thus improving the performance of the whole recognition system.

| Abbreviation | Full Name | Symbol | Meaning |
|--------------|-----------|--------|---------|
| P            | Positive Class | $c_{11} + c_{12}$ | SPTB |
| N            | Negative Class | $c_{21} + c_{22}$ | HC |
| TP           | True Positive  | $c_{11}$ | SPTB is correctly classified into SPTB. |
| FN           | False Negative | $c_{12}$ | SPTB is wrongly classified into HC. |
| FP           | False Positive | $c_{21}$ | HC is wrongly classified into SPTB. |
| TN           | True Negative  | $c_{22}$ | HC is correctly classified into HC. |

#### Table 3 Definitions in the confusion matrix

| Abbreviation | Full Name | Symbol |
|--------------|-----------|--------|
| P            | Positive Class | $c_{11} + c_{12}$ |
| N            | Negative Class | $c_{21} + c_{22}$ |
| TP           | True Positive  | $c_{11}$ |
| FN           | False Negative | $c_{12}$ |
| FP           | False Positive | $c_{21}$ |
| TN           | True Negative  | $c_{22}$ |
Fig. 8 Results of 14-way MDA
5.2 Statistical performance of our F3 model

The performances of 10 runs of 10-fold cross-validation are reported in Table 5. The MSD values of the seven indicators are listed as: sensitivity (93.68% ± 1.75%), specificity (94.17% ± 1.68%), precision (94.17% ± 1.55%), accuracy (93.92% ± 1.05%), F1 score (93.91% ± 1.07%), MCC (87.88% ± 2.09%), and FMI (93.92% ± 1.06%).

If we sum up the confusion matrixes of ten runs, the resultant confusion matrix is shown in Fig. 9. We can observe that 91 SPTBs are wrongly classified to HC while 84 HCs are wrongly classified to SPTB at ten runs.

5.3 SVM versus FSVM

The effectiveness of FSVM is validated in this ablation experiment. If we replace the FSVM classifier with the ordinary SVM classifier, the ten-run results are itemized in Table 6. The comparison is plotted in Fig. 10, where it clearly indicates our FSVM is superior to ordinary SVM. The ROC curves of both classifiers are drawn in Fig. 11.
Table 7  Comparison to State-of-the-art SPTB recognition models based on ten runs of ten-fold CV

| Model          | Sen       | Spc       | Prc       | Acc       | F1        | MCC       | FMI        |
|----------------|-----------|-----------|-----------|-----------|-----------|-----------|------------|
| SSVM [5]       | 83.82 ± 2.07 | 86.53 ± 1.80 | 86.18 ± 1.51 | 85.17 ± 1.05 | 84.96 ± 1.13 | 70.41 ± 2.09 | 84.98 ± 1.13 |
| AE-CNN [6]     | 87.57 ± 2.06 | 88.33 ± 2.47 | 88.30 ± 2.09 | 87.95 ± 1.37 | 87.91 ± 1.35 | 75.95 ± 2.72 | 87.92 ± 1.34 |
| depth-ResNet [7]| 88.75 ± 1.42 | 91.60 ± 1.48 | 91.38 ± 1.32 | 90.17 ± 0.61 | 90.03 ± 0.63 | 80.40 ± 1.23 | 90.05 ± 0.63 |
| CAS [8]        | 88.47 ± 1.40 | 91.32 ± 1.19 | 91.08 ± 1.07 | 89.90 ± 0.68 | 89.75 ± 0.72 | 79.84 ± 1.36 | 89.76 ± 0.71 |
| ICNN [9]       | 93.75 ± 0.98 | 93.19 ± 2.33 | 93.29 ± 2.07 | 93.47 ± 0.95 | 93.50 ± 0.87 | 86.98 ± 1.86 | 93.51 ± 0.86 |
| CheXaid [10]   | 89.93 ± 2.32 | 90.83 ± 0.72 | 90.76 ± 0.58 | 90.38 ± 1.04 | 90.33 ± 1.17 | 80.79 ± 2.03 | 90.33 ± 1.16 |
| F3 (Ours)      | 93.68 ± 1.75 | 94.17 ± 1.68 | 94.17 ± 1.55 | 93.92 ± 1.05 | 93.91 ± 1.07 | 87.88 ± 2.09 | 93.92 ± 1.06 |

Fig. 11  ROC curve comparison between SVM and our FSVM

Fig. 12  Error bar comparison of state-of-the-art SPTB recognition models

Springer
where we can observe the FSVM has an AUC of 0.9624, greater than that of SVM of 0.9483.

### 5.4 Comparison to state-of-the-art models

The experiment compares the proposed F3 model with six state-of-the-art SPTB recognition models: SSVM [5], AE-CNN [6], depth-ResNet [7], CAS [8], ICNN [9], and CheXaid [10]. The results are itemized in Table 7. We can observe that our F3 model achieves the best at six indicators out of all the seven indicators except sensitivity, of which the F3 model is merely 0.07% lower than the best model as ICNN [9]. The error bar comparison is depicted in Fig. 12.

The reasons why the proposed F3 model is the best are three points: (a) The 5-property from FDVDGLCM is used as the feature descriptor. (b) The FSVM is chosen, which has been proved to give better performance than ordinary SVM. (c) Slight adaption of multiple-way data augmentation is used to boost the training set.

Note ICNN [9] has better sensitivity than our F3 model. The reason is the ICNN used two advanced technologies: stochastic pooling and hyperparameter optimization (HO). The HO strategy may help our F3 model; hence, we shall try to integrate HO into our model in future studies. (Bold means the best.)

### 6 Conclusion

This study proposes a novel F3 model for SPTB recognition. The F3 model means (i) using FDVDGLCM to analyze the chest CT images, (ii) using FPFS to extract five properties, and (iii) using FSVM as the classifier. Besides, a slight adaption of multiple-way data augmentation is used to boost the training set. The ten runs of 10-fold cross-validation demonstrate that this F3 model is superior to six state-of-the-art models.

After reviewing our model, we observe four shortcomings: (i) The dataset is moderately small. (ii) The classification task is only a binary classification problem. (iii) Some advanced classification tools are yet to be verified. (iv) The hyperparameters can be optimized to improve the performances.

In future studies, we intend to endeavor to gather more SPTB subjects to expand the experimental dataset. Chest CT images of other infectious diseases will be added, such as community-acquired pneumonia and COVID-19. We shall test the recently proposed graph convolutional network, attention network, etc. Finally, the HO strategy will be used to integrate into our model.

**Acknowledgments** This paper is partially supported by Medical Research Council Confidence in Concept Award, UK (MC_PC_17171), Royal Society International Exchanges Cost Share Award, UK (RP202G0230), Hope Foundation for Cancer Research, UK (RM60G0680), Global Challenges Research Fund (GCRF), UK (P202PF11), Sino-UK Industrial Fund, UK (RP202G0289), and British Heart Foundation Accelerator Award, UK (AA/18/3/34220).

**References**

1. Pradhan M, Nanda B, Kar P, Nanda BB (2022) Intermolecular interactions of anti-tuberculosis drugs with different solvents: a review. Biointerface Research in Applied Chemistry 12(1):883–892
2. El Malik ARA, El Wahab A, Eltrawy HH (2021) Retrospective study of pulmonary and extrapulmonary tuberculosis cases in Giza Governorate during the period of 2014–2018. Egyptian Journal of Bronchology 15(1) Article ID. 24
3. Chirenda J, Chipinduro M, de Kock M, Spies C, Sakubani CT, Warren RM, Sampson SL, Streicher EM (2019) Recovery of mycobacterium tuberculosis from positive mycobacterium growth indicator tubes stored at room temperature for up to 6 years in low-income and high-tuberculosis-burden country. International Journal of Mycobacteriology 8(2):185–189
4. Anton AI, Andrei TP, Nicoleta M, Bechir ES, Ecaterina TD (2020) Tongue, tonsil and lung - rare synchronous digestive and respiratory involvement - a case series and short literature review. Acta Medica Mediterranea 36(2):793–796
5. Bagci U, Kubler A, Luna B, Jain S, Bishai WR, Mollura DJ (2013) Computer-aided detection and quantification of cavitary tuberculosis from CT scans. Med Phys 40(11):14 Article ID. 113701
6. Li LJ, Huang HY, Jin XY (2018) AE-CNN Classification of Pulmonary Tuberculosis Based on CT images. In: Ninth International Conference on Information Technology in Medicine and Education. IEEE, Hangzhou, China, pp 39–42
7. James-Reynolds C, Currie E, Gao XHW (2020) Analysis of tuberculosis severity levels from CT pulmonary images based on enhanced residual deep learning architecture. Neurocomputing 392:233–244
8. Xie YL, Wu ZY, Han X, Wang HY, Wu YF, Cui L, Feng J, Zhu ZH, Chen ZYL (2020) Computer-Aided System for the Detection of Multicategory Pulmonary Tuberculosis in Radiographs. J Healthc Eng 2020:12 Article ID. 9205082
9. Zhang Y-D, Nayak DR, Zhang X, Wang S-H (2020) Diagnosis of secondary pulmonary tuberculosis by an eight-layer improved convolutional neural network with stochastic pooling and hyper-parameter optimization. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-020-02612-9
10. Rajpurkar P, O’Connell C, Schechter A, Asnani N, Li JS, Kiania A, Ball RL, Mendelson M, Maartens G, van Hoving DJ, Griesel R, Ng AY, Boyles TH, Lungren MP (2020) CheXaid: deep learning assistance for physician diagnosis of tuberculosis using chest x-rays in patients with HIV. NPJ Digital Medicine 3(1):8 Article ID. 115
11. Patil SA, Udupi VR (2010) Chest X-ray features extraction for lung cancer classification. J Sci Ind Res 69(4):271–277
12. Songpan W (2018) Improved Skin Lesion Image Classification Using Clustering with Local-GLCM Normalization. In: 2nd European Conference on Electrical Engineering and Computer Science (EECS). IEEE, Bern, SWITZERLAND, pp 206–210
13. Srivastava D, Rajitha B, Agarwal S, Singh S (2020) Pattern-based image retrieval using GLCM. Neural Comput Appl 32(15):10819–10832
14. Athertya JS, Kumar GS (2021) Classification of certain vertebral degenerations using MRI image features. Biomed Phys Eng Exp 7(4) Article ID. 045013
15. Liu X, Chen S, Song L, Woźniak M, Liu S (2021) Self-attention negative feedback network for real-time image super-resolution. Journal of King Saud University - Computer and Information Sciences. https://doi.org/10.1016/j.jsuci.2021.07.014

16. Liu X, Song L, Liu S, Zhang Y (2021) A review of deep-learning-based medical image segmentation methods. Sustainability 13(3):1224

17. Rogic S, Kascelan L (2021) Class balancing in customer segments classification using support vector machine rule extraction and ensemble learning. Comput Sci Inf Syst 18(3):893–925

18. Osapoetra LO, Chan W, Tran W, Kollios MC, Czarnecka GJ (2020) Comparison of methods for texture analysis of QUS parametric images in the characterization of breast lesions. PLoS ONE 15(12) Article ID. e0244965

19. Abasi N, Sohrabi MR, Motiee F, Davallo M (2021) The application of artificial neural network and least support vector machine methods based on spectrophotometry method for the rapid simultaneous estimation of triamcinolone, neomycin, and nystatin in skin ointment formulation. Optik 241 Article ID. 167210

20. Jimenez-Cordero A, Morales JM, Pineda S (2021) A novel embedded min-max approach for feature selection in nonlinear support vector machine classification. Eur J Oper Res 293(1):24–35

21. Thakkar K, Ambekar SS, Hudnurkar M (2021) Prediction of longitudinal facial crack in steel thin slabs funnel mold using different machine learning algorithms. Int J Innov Sci 13(1):67–86

22. Karim, E., E. Abdellatif, A. Touhafi, and A. Nabil. An Improved Density Based Support Vector Machine (DBSVM). in 5th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications. 2020. Marrakesh, Morocco: IEEE; p. 335–341

23. Kiran T, Kushal T (2016) Facial Expression Classification using Support Vector Machine Based on Bidirectional Local Binary Pattern Histogram Feature Descriptor. In: 17th Ieee/Acis International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing. IEEE, Shanghai, PEOPLES R CHINA, pp 115–120

24. Raphael AB, Bhattacharya S, M. Assoc Comp (2020) A study on the effect of fuzzy membership function on fuzzified RIPPER for stock market prediction. In: ICMLSC 2020: Proceedings of the 4th International Conference on Machine Learning and Soft Computing. IEEE, Haiphong City, VIETNAM, pp 61–65

25. Liu S, Wang S, Liu X, Lin C, Lv Z (2021) Fuzzy detection aided real-time and robust visual tracking under complex environments. IEEE Trans Fuzzy Syst 29:90–102. https://doi.org/10.1109/TFUZZ.2020.3006520

26. Rout A, Deepak B, Biswal BB, Mahanta GB (2021) Trajectory generation of an industrial robot with constrained kinematic and dynamic variations for improving positional accuracy. International Journal of Applied Metaheuristic Computing 12(3):163–179

27. Garin SV (2020) ENDOXA: from Aristotle’s topics to fuzzy logic. Vestnik Tomskogo Gosudarstvenogo Universiteta-Filosofiya-Sotsiologiya-Politologiya-Tomsk State University Journal of Philosophy Sociology and Political Science 57:153–162

28. Hesamian G, Akbari MG (2021) Support vector logistic regression model with exact predictors and fuzzy responses. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-021-03333-3

29. Mohammadi M, Sarmad M (2021) A bilateral fuzzy support vector machine hybridizing the Gaussian mixture model. Iran J Fuzzy Syst 18(3):161–177

30. Maldonado S, Lopez J, Vairetti C (2021) Time-weighted fuzzy support vector machines for classification in changing environments. Inf Sci 559:97–110

31. Wieczorek J, Lei J (2021) Model selection properties of forward selection and sequential cross-validation for high-dimensional regression. Canadian Journal of Statistics-Revue Canadienne De Statistique. https://doi.org/10.1002/cjs.11635

32. Leach S, Xue YH, Srirdhar R, Paal S, Wang ZY, Murphy R (2021) Data Augmentation for Improving Deep Learning Models in Building Inspections or Postdisaster Evaluation. J Perform Constr Facil 35(4) Article ID. 04021029

33. Siafsi S, Rodolico A, Ciray O, Murphy DG, Parellada M, Arango C, Leucht S (2021) Imitating the Number of Responders from the Mean and Standard Deviation of CGI-Improvement in Clinical Trials Investigating Medications for Autism Spectrum Disorder. Brain Sci 11(7) Article ID. 908

34. Kim JY, Cho SB (2021) Deep CNN transferred from VAE and GAN for classifying irritating noise in automobile. Neurocomputing 452:395–403

35. Park DJ, Park MW, Lee H, Kim YJ, Kim Y, Park YH (2021) Development of machine learning model for diagnostic disease prediction based on laboratory tests. Sci Rep 11(1) Article ID. 7567

36. Agarwal A, Sharma P, Alshehri M, Mohamed AA, Alfarraj O (2021) Classification model for accuracy and intrusion detection using machine learning approach. Peerj Comput Sci Article ID e437

37. Dreizin D, Goldmann F, LeBedis C, Boscak A, Dattwyler M, Bodanapally U, Li G, Anderson S, Maier A, Unberath M (2021) An automated deep learning method for tile AO/OTA pelvic fracture severity grading from trauma whole-body CT. J Digit Imaging 13. https://doi.org/10.1007/s10278-020-00399-x

38. Aql MMF, Bahget SAE, Kholoussi N, Abdel-Salam GME, Raouf HA, Eid MM, Esmail REB (2021) Telomerase dysfunction in the tumorigenesis of genetic disorders. Int J Mol Cell Med 13(1):56–67

39. Leslie WD, Epp R, Morin SN, Lix LM (2021) Assessment of site-specific X-ray procedure codes for fracture ascertainment: a registry-based cohort study. Arch Osteoporos 16(1) Article ID. 107

40. Ali A, Khan N, Abu-Tair M, Noppen J, McClean S, McChesney I (2021) Discriminating features-based cost-sensitive approach for software defect prediction. Autom Softw Eng 28(2) Article ID. 11

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