Radiomics Based Bayesian Inversion Method for Prediction of Cancer and Pathological Stage

HINA SHAKIR[1], **TARIQ KHAN**[2], **HAROON RASHEED**[1], **(Senior Member, IEEE),**
AND YIMING DENG[3], **(Senior Member, IEEE)**

[1]Department of Electrical Engineering, Bahria University, Karachi 75620, Pakistan
[2]Department of Electrical and Power Engineering, National University of Science and Technology, Islamabad 75350, Pakistan
[3]Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA

CORRESPONDING AUTHOR: H. SHAKIR (hinashakir.bukc@bahria.edu.pk)

ABSTRACT   Objective: To develop a Bayesian inversion framework on longitudinal chest CT scans which can perform efficient multi-class classification of lung cancer. Methods: While the unavailability of large number of training medical images impedes the performance of lung cancer classifiers, the purpose built deep networks have not performed well in multi-class classification. The presented framework employs particle filtering approach to address the non-linear behaviour of radiomic features towards benign and cancerous (stages I, II, III, IV) nodules and performs efficient multi-class classification (benign, early stage cancer, advanced stage cancer) in terms of posterior probability function. A joint likelihood function incorporating diagnostic radiomic features is formulated which can compute likelihood of cancer and its pathological stage. The proposed research study also investigates and validates diagnostic features to discriminate accurately between early stage (I, II) and advanced stage (III, IV) cancer. Results: The proposed stochastic framework achieved 86% accuracy on the benchmark database which is better than the other prominent cancer detection methods. Conclusion: The presented classification framework can aid radiologists in accurate interpretation of lung CT images at an early stage and can lead to timely medical treatment of cancer patients.

INDEX TERMS   Bayesian inversion, cancer stage estimation, nodule classification, particle filter, radiomic features.

I. INTRODUCTION

Image analysis techniques have successfully provided personalized prognosis and treatment plans for cancer patients with a greater accuracy in the recent years. There are several studies available that show the importance of imaging features for treatment, monitoring and outcome prediction in lung cancer as well as other cancer types [1], [2]. Besides prognosis, the quantitative imaging features also known as radiomic features have been extracted from the medical images to be further used in clinical research for lung cancer diagnosis [3], [4]. Most of the prominent radiomics based classification methods for lung cancer have incorporated neural networks and deep learning approaches [5]–[13]. A multi-view convolutional network is proposed in [9] for binary (benign and malignant) and ternary classification (binary, primary malignant and metastatic malignant). The proposed model achieved error rate of 5.41% for binary classification

and 13.91% for ternary classification respectively. Another multi-view knowledge-based collaborative (MV-KBC) deep model classified nodules on limited chest CT data in [11]. The prediction error computed in terms of Root Mean Square Error (RMSE) of the model is 0.62. Authors in [12] investigated the usefulness of a deep convolutional neural network (DCNN) for benign, primary lung cancer and metastatic cancer classification. The best averaged accuracy of 68% was achieved with transfer learning for image size of 224.

Another common approach adopted to address classification problem is the application of machine learning classifiers. A structural co-occurrence matrix (SCM) based method was proposed by [13]. The proposed technique did not only classify the nodules as benign or malignant but also classified the malignant nodules into 5 prescribed levels with an accuracy of 74.5%. Yu *et al.* in [5] presented a study to detect pathological stages I, II, III and IV respectively in non-small

TABLE 1. List of notations.

Annotations	Description
k	Screening time in quarters
U	Set of diagnostic scores as states
U_k	Diagnostic score (state) at time k
K	Total number of quarters
Z_k	Set of diagnostic radiomic feature measurements
z_{qk}	q^{th} diagnostic feature measurement at time k
N_s	Total number of samples
u_i^k	i^{th} particle at screening time k
w_i^k	Weight of ith particle at screening time k
p	m^{th} coefficient of measurement function
M	Order of measurement function
ξ	Predefined convergence threshold
N	Number of iterations

FIGURE 1. Work flow of the Bayesian inversion method implemented in proposed system for multi-class classification of lung nodules.

cell lung cancer (NSCLC) using Random Forest algorithm. Out of the bag (OOB) error was computed to measure classification accuracy of the Random Forest algorithm whereas the predicted cancer stages of test data sets achieved ROCs between 0.6 to 1.

However, there is a dearth of scientific approaches to diagnose the benign or cancerous nature of a lung nodule along with its pathological stage in case of malignancy. Furthermore, the insufficient size of training data sets required for neural networks training is a challenge for accurate cancer stage detection of the tumor. Keeping in view the aforementioned challenges, there is an emerging need to investigate new sophisticated solutions for the detection of malignancy status in lung nodules. A useful approach towards joint estimation of tumor malignancy and its stage can involve tracking of tumor characteristics using temporal analysis of CT images [14], [15]. Such an approach also offers good potential to address the problem of smaller data sets in medical imaging.

In this research article, a Bayesian inversion method is proposed which employs longitudinal radiomics data extracted from CT scans of lung nodules to detect cancer and its pathological stage. The presented system adopts particle filtering approach to compute the posterior Probability Density Function (PDF) of cancer occurrence and cancer pathological stage respectively. For PDF estimation, a joint likelihood function incorporating diagnostic radiomic features is developed which can diagnose malignant nodules and the pathological cancer stage. The cancer stage is determined as early stage cancer (stage I & stage II) or advanced stage cancer (stage III & stage IV) respectively. Furthermore, radiomic features which can discriminate between cancer stages are also investigated and evaluated.

The presented Bayesian inversion method was validated on a total of 200 nodules with 130 malignant and 70 benign nodules. There were 109 malignant and 63 benign nodules which were successfully diagnosed showing better performance of the proposed framework than the other prominent methods in literature.

In the following sections, research problem is formulated and proposed stochastic frame work is described. Next,
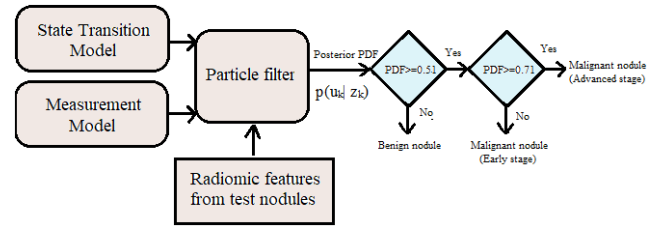
implementation of the presented model is discussed. Then, results and discussion are presented followed by a conclusion. The notations used in this research work are listed in Table 1 for reader's understanding.

II. PROBLEM FORMULATION

A number of radiomic features depicting shape and gray level intensities of lung nodule have shown diagnostic power towards lung, colon and head and neck cancer [16]. The aim of the presented research is to utilize such radiomics features measurements from periodic chest screening images in a stochastic framework to detect lung cancer. The sequential evaluation of nodules' malignancy status given the radiomics data from periodic screening images is an inverse problem and can be solved using Bayesian inversion approach [17]. There are several methods available to solve the aforesaid problem but a recursive Bayesian filter based on sequential Monte Carlo method is proposed since it uses nonlinear models [18]. The need of nonlinear modeling arises due to the nonlinear behavior of measured radiomics features and the associated likelihood function for lung cancer diagnosis [16].

The proposed Bayesian inversion framework requires a likelihood function for PDF estimation. Since multi-class classification using a single mathematical model has not been a preferred approach due to the complex relationship between benign and cancerous (multiple stages) nodules [19], [20], an intuitive approach has been adopted to combine two mathematical likelihood equations into one for PDF estimation. The proposed system for cancer prediction is shown in Fig. 1. In practice, an inverse problem can be represented by two models namely state transition model and measurement model [21] and is shown in Fig. 2 with states represented by u and measurements by z. A description of these two models and their application in Bayesian inversion method for malignancy estimation is given in the following section.

III. DIAGNOSTIC STOCHASTIC FRAMEWORK
A. STOCHASTIC STATE-SPACE MODELS

To start with, a state vector $U = \{u_1, u_2, u_3, u_4, \ldots, u_{k-1}, u_k\}$ is defined where each element u_k is a diagnostic score between 0 and 1 that represents the benign/malignancy state of a lung tumor at a screening interval k. A state transition model represents the evolution of state u_k with respect to time instants and can be expressed as [22]:

$$u_k = x_{k-1}(u_{k-1}, v_{k-1}) \qquad (1)$$
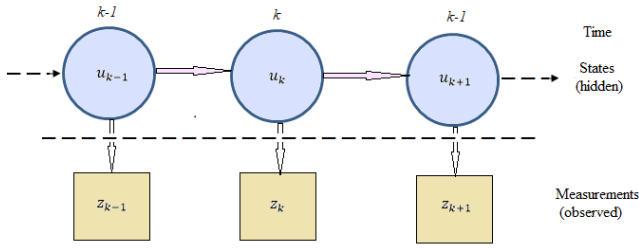
FIGURE 2. Illustration of proposed inverse problem with hidden states and their observed measurements.

Here $x_k(.)$ is a non-linear and time varying function that models the state transition process and v_k is process noise respectively.

The q number of diagnostic radiomic feature measurements of a lung nodule in a CT image at each screening interval k are assumed to be $z_{1k}, z_{2k}, \ldots, z_{qk}$ and are represented by a measurement vector Z such that $Z = \{z_{11}, z_{21}, \ldots, z_{q1}, z_{21}, z_{22}, \ldots z_{q2}, \ldots, z_{1k}, z_{2k}, \ldots, z_{qk}\}$. The measurement function relates the features measurements to the true malignancy states of a lung tumor and can be expressed as:

$$z_k = y_k(u_k, n_k) \tag{2}$$

Here $y_k(.)$ is a time varying and independent function that represents measurement model and n_k is the measurement noise. Since, there has been a lack of discussion on the linearity of radiomics based state transition and measurement functions in the literature, a more generalized approach is adopted to perform the Bayesian inversion estimation which is offered by particle filter to estimate the posterior probability of malignancy and pathological cancer stage respectively in lung nodules.

B. PARTICLE FILTERING

A particle filter is an algorithm which represents the desired posterior PDF of an unknown state as weighted samples, also known as particles. Using a large number of samples (particles), the estimated PDF is approximated to the true PDF. In particle filtering approach to dynamic state estimation of a lung nodule, the posterior PDF of the unknown malignancy state u_k is constructed based on the set of received feature measurements Z_k.

Given the likelihood $p(Z_k|u_k)$ of measured diagnostic features Z_k at time k, the posterior density $p(u_k|Z_k)$ of unknown state u_k for the set of received features Z_k can be computed using Bayes' Theorem [23] as follows:

$$p(u_k \mid Z_k) = \frac{p(Z_k \mid u_k)p(u_k \mid Z_{k-1})}{p(z_k \mid Z_{k-1})} \tag{3}$$

Here the normalizing constant $p(z_k \mid Z_{k-1})$ is given by:

$$p(z_k \mid Z_{k-1}) = \int p(z_k \mid U_k)p(U_k \mid Z_{k-1})du_k \tag{4}$$

The normalization constant depends upon the likelihood function $p(Z_k|u_k)$, defined by the measurement model in

Eq. (2). Recursively, the filter predicts and updates to estimate the sequential unknown malignancy states of nodules. In summary, the proposed particle filter expresses the PDF $p(u_k|Z_k)$ of a nodule's unknown state at screening interval k as a set of weighted samples. The total number of samples drawn from an importance density are assumed N_s. We represent a sample point as u_k^i and its associated weight as w_k^i where $i = \{1, 2, \ldots, N_s\}$. A randomly measured weighted sample at any screening interval can be represented as a set $\{u_k^i, w_k^i\}$. Using the principle of importance sampling, the weights are normalized such that $\sum w_k^i = 1$. Then, the posterior density at a screening time k can be approximated as:

$$p(u_k \mid Z_k) = \sum_{i=1}^{N_s} w_k^i \delta(u_k - u_k^i) \tag{5}$$

The samples can be drawn from the importance density $q(Z_k^i|U_k)$ and the weights can be assigned using the following association:

$$w_k^i \propto \frac{p(Z_k^i \mid U_k)}{q(Z_k^i \mid U_k)} \tag{6}$$

If the samples at screening time $k-1$ are available constituting to probability $p(z_{1:k-1} \mid u_{1:k-1})$, then $p(z_{1:k} \mid u_{1:k})$ is approximated with a new set of samples once the observation z_k is known at the nodule screening time k.

The prior or state transition probability $p(u_k|u_{k-1})$ of malignancy in a nodule is modeled using exponential distribution because the progression rate of lung cancer follows exponential distribution [24]. Since the transition density depends upon the malignancy states of cancer u_{k-1} and u_k at any consecutive times $k-1$ and k respectively, it can be represented as follows:

$$p(u_k \mid u_{k-1}) = e^{-\frac{|u_k - u_{k-1}|^\alpha}{\alpha}} \tag{7}$$

where α is a scalar value that controls the variability in the predicted state value.

IV. IMPLEMENTATION OF PARTICLE FILTER
The test datasets which are given as input to the proposed model comprises of longitudinal radiomics data of patients' CT images acquired from National Lung Screening Trial (NLST) [25].

A. NLST DATASETS
NLST is a randomized multi-site trial which enrolled 53,454 people with high cancer risk to assess the impact of using low dose CT screening in mortality reduction from lung cancer. In NLST project, low dose CT screenings of the enrolled people were performed for three consecutive years which are termed as T0, T1 and T2 respectively. For the implementation of proposed filter, longitudinal data of participants who had cancer detected after the third year (T2) of screening were included. This ensured availability of CT data for all the three years in order to apply the prediction model. A total of 130 cases with positive screening at either

year T0 or T1 followed by screen detected cancer at year T2 were acquired. Another 70 CT data sets with benign nodules detected at year T2 and with positive screening at either year T0 or T1 were accessed from the database.

The detected nodules from all the 3 consecutive CT screenings of each patient were segmented and 105 3-D features were computed. This resulted in a set of 3 longitudinal radiomic readings for each patient. In order to tackle the constraint of fewer data points, single data point was interpolated using cubic-spline method [17] between data obtained from year T0 and year T1, and between year T1 and year T2 respectively. The total number of screening intervals after interpolation were $5(K = 5)$ within a duration of 3 years. Starting with an initial CT scan, there are 4 sample points following and each point represents a half-yearly CT screening data of the patient which was fed into the proposed system for PDF estimation.

B. LIKELIHOOD FUNCTION FOR CANCER AND CANCER STAGE DETECTION

For PDF estimation using particle filter, a likelihood function has to be developed. In the literature reviewed, authors in [16] formulated a likelihood function for cancer detection (benign or malign) incorporating two highly diagnostic radiomic features namely Surface volume ratio and Sum entropy values of a nodule and is given as:

$$p_1(Z_k|u_k) = a + b(z_{1k}) + c(z_{1k}^2) + d(z_{1k}^3) + e(z_{1k}^4) + f\ln(z_{2k}) \tag{8}$$

Here z_{1k} represents the Surface volume ratio and z_{2k} denotes the Sum entropy value of nodule at k^{th} quarter. The value of $p_1(Z_k|u_k)$ is below 0.51 for benign state and equal or above 0.51 for malignancy in the nodule. The coefficients of Eq. (8) assume the following values [16]:

$$a = 0.7478; \ b = 2.2268; \ c = -5.5856;$$
$$d = 3.6318; \ e = -0.73065; \ f = 1.2814E - 02$$

As reported in a research study [26], a single multi-class likelihood function is prone to large errors hence cancer detection and cancer stage detection have to be performed using two distinct mathematical functions.

Malignancy with cancer pathological stage I or II was collectively termed as early stage cancer whereas stage III and IV cancer was termed as advanced stage cancer [27]. The training cohort comprised of 73 nodules with early stage cancer and 127 nodules with advanced stage cancer from Lung1 [28] database. A pre-processing step of features extraction and features reliability test [16] was applied to the training data sets which resulted in 51 stable and reliable features reported in Table SI (Supplementary material). Furthermore, the redundant features were removed and highly diagnostic features were selected using Least absolute shrinkage and selection operator (LASSO) [10], a method of regression analysis. LASSO technique performs L1 regularization of the features via penalized estimation function to achieve a
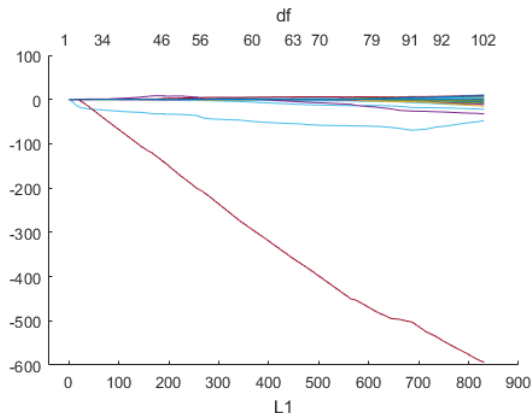
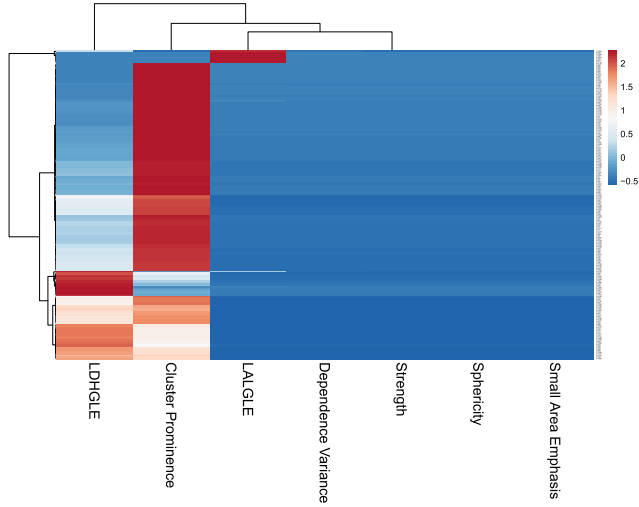**FIGURE 3.** LASSO plot of coefficients fit.

reduced set of diagnostic features which were validated using 5-fold cross validation. Fig. 3 shows the LASSO coefficients of predictors on y-axis and the L1 norm (the penalization parameter also known as lambda) of coefficients on x-axis for a regression fit. The top x-axis of the plot represents degree of freedom denoted by df and it shows the non-zero coefficients of predictors. Following 7 diagnostic features were obtained after the application of LASSO technique (lambda = 0.0282 with minimum MSE) on pre-processed features:

1) Sphericity
2) Dependence Variance (DV)
3) Large Dependence High Gray Level Emphasis (LDHGLE)
4) Cluster Prominence (CP)
5) Large Area Low Gray Level Emphasis (LALGLE)
6) Small Area Emphasis (SAE)
7) Strength

The LASSO coefficients of the aforementioned diagnostic features are reported in Table SII. Their diagnostic power ranking was computed using mean scores of 3 supervised feature ranking approaches namely feature based Neighborhood Component Analysis (fNCA), ReliefF network and Infinite Latent Feature Selection (ILFS) respectively [29]. Out of these 7 coefficients, five most discriminating features were used to formulate the likelihood function for early stage or advanced stage cancer detection as follows:

$$p_2(Z_k \mid u_k)$$
$$= 1.2 - 1.1348(z_{3k}) - 7.4597e^{-07}(z_{4k})$$
$$+ 3.0780e^{-08}(z_{5k}) + 0.3917(z_{6k}) - 0.00361(z_{7k}) \tag{9}$$

Here $p_2(Z_k|u_k)$ indicates likelihood of cancer stage as early stage or advanced stage cancer. It produces a value between 0.51 and below 0.70 for early stage cancer and assumes a value between 0.70 and 0.90 for the detection of advanced stage cancer. The features z_{3k}, z_{4k}, z_{5k}, z_{6k} and z_{7k} represent Sphericity, LDHGLE, Cluster Prominence, Small Area Emphasis and Strength respectively for k^{th} CT screening. The coefficients z_{4k}, z_{5k} of features LDHGLE and Cluster

FIGURE 4. Hierarchical clustering of selected radiomic features as a heatmap.

FIGURE 5. Work flow of the radiomics based Bayesian inversion method for multi-class classification of lung nodules.

Prominence respectively are very small and can be excluded. The proposed likelihood function was validated on 47 nodules with early stage cancer and 143 nodules with advanced stage cancer. The function detected 40 nodules with early stage cancer and 125 nodules with advance stage cancer following the designated score ranges. Overall, the proposed function achieved an accuracy of 86.84% (CI:81.19% to 91.30%), specificity of 85.11% (CI:71.69% to 93.80%) and sensitivity of 87.41% (CI:80.84% to 92.37%) respectively with an AUCROC of 84.61%. The ROC plot of the nodule pathological stage classification using the presented likelihood function is shown in Fig. S1. The correlation of these top 5 diagnostic features (LASSO method) with two cancer groups (early stage and advanced stage) is further demonstrated using hierarchical clustering in a heatmap in Fig. 4. The numerical values of plotted radiomic features are standardized with mean 0 and standard deviation 1. The features are represented on x-axis and their hierarchical clustering is represented on y-axis. Two distinct clusters were formed showing strong association of CP and LDHGLE with the defined cancer groups. However, the heatmap could not fully capture the association of other features with two cancer stages besides CP and LDHGLE.

Finally, in order to integrate cancer and its stage detection equations into one likelihood function, highly diagnostic features namely Surface volume ratio (SVR) and Sum entropy (SE) were chosen as the selection criterion [16]. It was observed through an empirical study carried out on the training data that values of SVR (z_{1k}) and SE (z_{2k}) are always above threshold values T_a and T_b respectively for malignant nodules and therefore can be used to estimate likelihood of early and advanced stage cancer from Eq. (9). On the contrary, the values of z_{1k} and z_{2k} below the threshold T_a and T_b respectively indicate that nodule is benign/non-malignant and its likelihood can be computed using Eq. (8) Using the above mentioned criterion and Eq. (8) and Eq. (9), the proposed
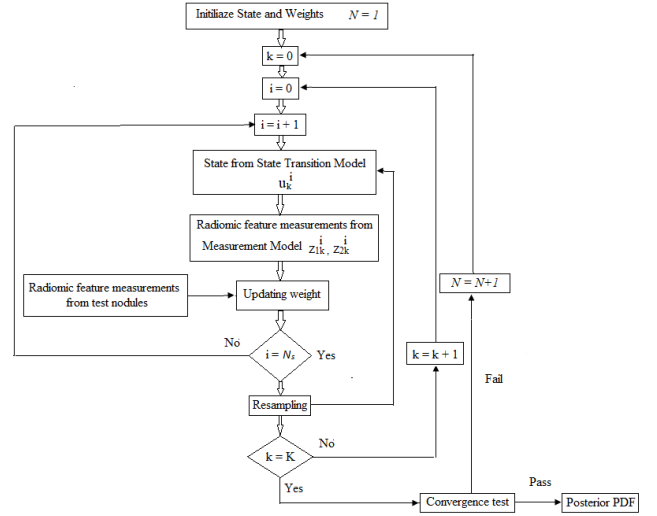
likelihood function for multi-class classification is given as follows:

$$p(Z_k \mid u_k) = \begin{cases} p_2(Z_k \mid u_k), & \text{if } z_{1k} < T_a \text{ AND } z_{2k} < T_b \\ OR \\ p_1(Z_k \mid u_k), & \text{if } z_{1k} \geq T_a \text{ AND } z_{2k} \geq T_b \end{cases}$$
(10)

C. ALGORITHM

The sequence of steps carried out for the proposed diagnostic framework implementation is demonstrated in Fig. 5 and is summarized as follows:

1) To start, N_s samples are generated from the prior/state transition PDF in Eq. (7) for the first time screening of a lung nodule. The value of N_s is equal to 4000.
2) Next, the features are computed using the measurement function for the known malignancy state from Step 1. For this, an Mth order polynomial is proposed as a measurement function as follows:

$$z_{qk} = \sum_{m=0}^{M} p_m u_k^m$$
(11)

Here z_{qk} represents q^{th} diagnostic radiomic feature at k^{th} chest screening. The coefficients p_m of polynomial were determined from a training database of known states and the corresponding diagnostic feature measurements. As discussed, two features (SVR(z_{1k}), SE(z_{2k})) were required to estimate benign /malignant state of a nodule [16] and their measurement functions were formulated as:

$$z_{1k} = \sum_{m=0}^{M} p_{1m} u_k^m$$
(12)

$$z_{2k} = \sum_{m=0}^{M} p_{2m} u_k^m$$
(13)

FIGURE 6. Confusion matrix of cancer detection and stage detection using (a) $M = 2$ and (b) $M = 3$.

FIGURE 7. Estimated diagnostic scores of test nodules versus their malignancy status for $M = 3$ and $M = 2$.

The coefficients p_{1m} and p_{2m} were computed for z_{1k} and z_{2k} respectively. Similarly, measurement functions of cancer stage estimation were also developed.

3) Using the features computed in Step 2 and the test radiomics features, two likelihood PDFs were estimated from Eq. (10). The closeness of the two likelihood values obtained determines the sample weight. Large difference refers to small weight assignment and vice versa. Subsequently, the weights were updated using association in Eq.(6).

4) The problem of degeneracy with particles is addressed by performing resampling [30]. After resampling, the posterior PDF of malignancy occurrence was estimated using samples and their associated weights.

5) The PDF estimation process outlined above was repeated for all the cancer screening times from 1 up till K.

6) The expectation of state computed from the PDF estimates at k was represented as posterior mean estimate (PME) as follows:

$$u_k^{PME} = E[u_k \mid Z_k] = \int u_k p(u_k \mid Z_k) du_k \quad (14)$$

7) If the error in single point PDF estimates in two consecutive iterations was within a preset threshold ξ, each iteration was denoted by N and the estimated malignancy state of a nodule (K screening times) evaluated at N was $u_{1:K}^N$, then the convergence criterion was given by:

$$\frac{\sum_{k=1}^{K} (u_k(N+1) - u_k(N))^2}{K} \leq \xi \quad (15)$$

V. DATA SETS SUMMARY

The proposed particle filtering approach was validated using 200 low dose CT data sets collected from NLST database. From these, there are 130 longitudinal datasets suffering from cancer stage I, II, III or IV respectively at the third year (T3) of screening. Other 70 datasets include patients' longitudinal CT data diagnosed with benign nodules at the third year (T3) of screening. The dimensions of the acquired images were 512 x 512 pixels. A summary of the test nodules is reported in Table SIII.

VI. RESULTS

The 3-D segmentation of nodules in lung CT volumes was carried out using Grow Cut algorithm after the verification
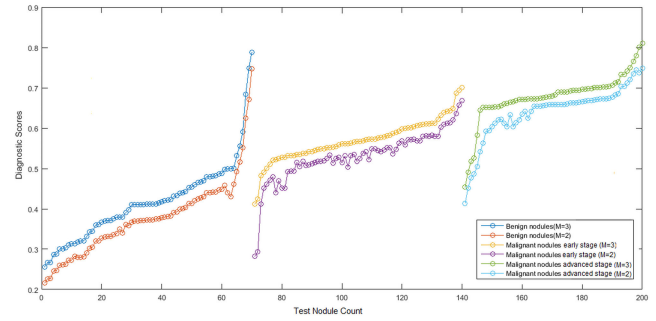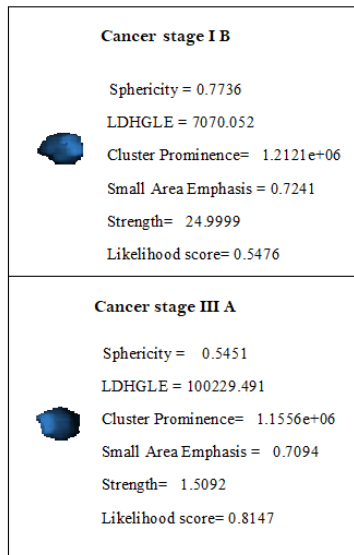
from a senior radiologist. Pyradomics [4] was used for the extraction of 3-D radiomic feature extraction. The diagnosis results of the filter were evaluated by setting the value of M as 2 and 3 respectively in measurement functions, described in subsection IV-C, Step 2.

With the order of measurement equations set to 2 ($M = 2$), the particle filter diagnosed 65 out of 70 benign tumors successfully. In malignant cases, the proposed framework diagnosed 51 out of 70 nodules having early stage cancer and 37 out of 60 nodules with advanced stage cancer correctly. For a comparison, the performance of the proposed framework was assessed using measurement equations of order 3 ($M = 3$). A total of 63 out of 70 benign tumors were successfully detected whereas in malignant cases, there were 60 out of 70 nodules with early stage cancer and 49 out of 60 nodules with advanced cancer stage which were correctly diagnosed. Increasing the order M to 4 did not show any significant improvement in the detection results and are excluded in the discussion.

The described diagnostic performance of the filter is expressed in terms of precision and recall for 95% confidence interval along with confusion matrix in Fig. 6. The recall values determine the sensitivity of classification model and precision values compute the positive predictive value of model. The average precision of the filter decreased from 92.86% to 90.0% for benign nodule classification but increased from 72.857% to 85.714% and 61.667% to 86.667% respectively for early and advanced stage cancer estimation when M value was incremented from 2 to 3. The recall metric improved significantly for both benign as well as early stage classification with $M = 3$. The overall accuracy of the filter is 86%.

The estimated diagnostic scores of benign as well as malignant test nodules at final CT screening (k = 5) which is the 4th half yearly reading are plotted against each nodule count for $M = 2$ & 3 in Fig. 7. Through visual inspection, an overall improvement in the state estimation of malignant lung nodules can be seen using measurement functions of order 3 in PF. To summarize, the diagnostic scores using $M = 2$ showed a tendency to shift towards the lower diagnostic ranges causing a spill over in the adjacent classification group. For example, 5 more nodules with early stage cancer obtained a score below 0.51 with $M = 2$, hence were

Cancer stage I B

Sphericity = 0.7736

LDHGLE = 7070.052

Cluster Prominence= 1.2121e+06

Small Area Emphasis = 0.7241

Strength= 24.9999

Likelihood score= 0.5476

Cancer stage III A

Sphericity = 0.5451

LDHGLE = 100229.491

Cluster Prominence= 1.1556e+06

Small Area Emphasis = 0.7094

Strength= 1.5092

Likelihood score= 0.8147

FIGURE 8. Diagnostic radiomic features and likelihood scores of two sample nodules having cancer stage I B and stage III A.

diagnosed as benign nodules. Similar trend was observed in nodules with advanced stage cancer where 9 more nodules got diagnosed having early stage cancer stage due to low scores for $M = 2$. The above discussed trend in state estimation is illustrated by plotting the state evolution (diagnostic scores) of 1 benign, 1 malignant (cancer stage IB) and 1 malignant (cancer stage IIIA) against the half-yearly CT screenings in Fig. S2. Furthermore, the ROC curves demonstrating the performance of proposed filter for classification of benign, early stage and advanced stage cancer are shown in Fig. S3. The AUCs achieved for benign nodule, early stage cancer and advanced stage cancer detection are 0.8962, 0.8849 and 0.8691 respectively. Computed values of diagnostic features and likelihood scores for stage I B(early stage) and stage III A(advanced stage) cancer are shown in Fig. 8.

VII. DISCUSSION

The binary classification performance of proposed model into benign and malignant nodules is compared with two other significant models tested on NLST database in Table SIV. The research work in [31] investigated the performance of different classifiers to predict malignancy in screen detected cancer at first year (T0) and second year (T1) of Lung screening trial. A total of 208 nodules were detected among which 104 were diagnosed with cancer. An accuracy (ACC) of 80%, specificity (SPS) of 91% and $AUC = 0.83$ were obtained. Authors in [6] analyzed deep learning algorithm generated heat-maps to predict the malignancy in nodules. The presented explainable AI approach to detect malignancy achieved a weighted accuracy of 85%. In comparison, the proposed PF achieved an accuracy of 92.0%, sensitivity of 87.50% and specificity of 94.5% respectively showing excellent performance of the filter in binary classification.

Since no multi-class classification techniques have been reported on NLST datasets in the literature, multi-class

TABLE 2. Comparison of multi-class classification performance using SVM, Random Forest, KNN and PF.

Prediction Model	No. of CT datasets	Prec. %	Recall %	Acc. %	MCC %
SVM	200	79.08	79.09	79.08	71.12
KNN	200	77.20	78.25	78.0	70.51
Random Forest	200	82.79	83.23	83.10	76.44
PF	200	85.79	86.11	86.0	80.08

classification was performed using SVM with Radial Basis Function (RBF) kernel, Random Forest and K nearest neighbors KNN) classifiers for a comparative performance analysis. The data sets were chosen from third year (T2) screening with 180 datasets used for training and the remaining 20 datasets were chosen for the testing of classifiers. The number of benign and malignant nodules in the experiments were same as used in the proposed filter and reported in Table SIII. The classifier was trained and tested 10 times with 10-fold cross validation. The evaluation metric included averages (for 10 iterations) of precision, recall, accuracy and Mathew's Correlation Coefficient (MCC) and is reported in Table 2. Classification of nodules as benign, with early stage cancer and advanced stage cancer using SVM, KNN and Random Forest classifiers achieved average precision of 79.08%, 77.20% and 82.79%, recall of 79.09%, 78.25% and 83.23% and an accuracy of 79%, 78% and 83% respectively. In comparison, the proposed particle filter resulted in an average precision of 85.79%, recall of 86.11% and an accuracy of 86% respectively. Furthermore, MCC [32] was computed which achieved values of 71.12%, 70.51%, 76.44% and 80.08% for SVM, KNN, Random Forest and PF respectively. Evidently, the achieved multi-class diagnosis results using diagnostic radiomic features and particle filter are better when compared with the trained SVM, KNN and Random Forest classifier on low dose datasets. The proposed filter has classified 86% of the nodules correctly with fewer input longitudinal data points proving particle filter a good choice for multi-class classification. However, the classification results could be improved with the availability of more chest CT data but is non-attainable due to health hazards attached with multiple CT screenings of a patient. Moreover, a limitation faced to implement the proposed approach is poor management of medical images database currently at many hospitals.

With proper management of patients' longitudinal CT data, this research work can be extended to cancer prognostics. Besides current status, the future malignancy status of a nodule can also be estimated using PF approach which re-iterates its efficacy.

VIII. CONCLUSION

An accurate and efficient radiomics based Bayesian inversion framework is proposed to detect cancer and its pathological stage as early or advanced stage. This contribution is superior to other prominent cancer detection algorithms whose

performances are dependent on large data sets training, a limitation in medical imaging. Moreover, the published contemporary methods do not provide a complete picture of malignancy and its stage using a single research approach as presented in this research work. Furthermore, highly discriminative features towards cancer stages were also identified. The high accuracy of the presented method shows that the proposed model can be used by clinicians in cancer diagnostics.

REFERENCES

[1] V. Nardone, P. Tini, P. Pastina, C. Botta, and A. Reginelli, "Radiomics predicts survival of patients with advanced non-small cell lung cancer undergoing PD-1 blockade using Nivolumab," *Oncol. Lett.*, vol. 9, no. 2, pp. 1559–1566, 2020.

[2] Y. Zhang, A. Oikonomou, A. Wong, M. A. Haider, and F. Khalvati, "Radiomics-based prognosis analysis for non-small cell lung cancer," *Sci. Rep.*, vol. 7, no. 1, p. 46349, Jun. 2017.

[3] Z. Zhou, S. Li, G. Qin, M. Folkert, S. Jiang, and J. Wang, "Multi-objective-based radiomic feature selection for lesion malignancy classification," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 1, pp. 194–204, Jan. 2020.

[4] P. Afshar, A. Mohammadi, K. N. Plataniotis, A. Oikonomou, and H. Benali, "From handcrafted to deep-learning-based cancer radiomics: Challenges and opportunities," *IEEE Signal Process. Mag.*, vol. 36, no. 4, pp. 132–160, Jul. 2019.

[5] L. Yu *et al.*, "Prediction of pathologic stage in non-small cell lung cancer using machine learning algorithm based on CT image feature analysis," *BMC Cancer*, vol. 19, no. 1, p. 464, Dec. 2019.

[6] V. K. Venugopal *et al.*, "Unboxing AI–radiological insights into a deep neural network for lung nodule characterization," *Academic Radiol.*, vol. 27, no. 1, pp. 88–95, Jan. 2020.

[7] W. Choi *et al.*, "Radiomics analysis of pulmonary nodules in low-dose CT for early detection of lung cancer," *Med. Phys.*, vol. 45, no. 4, pp. 1537–1549, Apr. 2018.

[8] C.-H. Chen *et al.*, "Radiomic features analysis in computed tomography images of lung nodule classification," *PLoS ONE*, vol. 13, no. 2, Feb. 2018, Art. no. e0192002.

[9] K. Liu and G. Kang, "Multiview convolutional neural networks for lung nodule classification," *Int. J. Imag. Syst. Technol.*, vol. 27, no. 1, pp. 12–22, Mar. 2017.

[10] Z. Liu *et al.*, "Radiomics of multiparametric MRI for pretreatment prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer: A multicenter study," *Clin. Cancer Res.*, vol. 25, no. 12, pp. 3538–3547, Jun. 2019.

[11] Y. Xie *et al.*, "Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT," *IEEE Trans. Med. Imag.*, vol. 38, no. 4, pp. 991–1004, Apr. 2019.

[12] M. Nishio *et al.*, "Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning," *PLoS ONE*, vol. 13, no. 7, Jul. 2018, Art. no. e0200721.

[13] M. B. Rodrigues *et al.*, "Health of things algorithms for malignancy level classification of lung nodules," *IEEE Access*, vol. 6, pp. 18592–18601, 2018.

[14] J. Teng *et al.*, "Bayesian inference of lymph node ratio estimation and survival prognosis for breast cancer patients," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 2, pp. 354–364, Feb. 2020.

[15] E. Hajiramezanali, M. Imani, U. Braga-Neto, X. Qian, and E. R. Dougherty, "Scalable optimal Bayesian classification of single-cell trajectories under regulatory model uncertainty," *BMC Genomics*, vol. 20, no. S6, pp. 1–11, Jun. 2019.

[16] H. Shakir, Y. Deng, H. Rasheed, and T. M. R. Khan, "Radiomics based likelihood functions for cancer diagnosis," *Sci. Rep.*, vol. 9, no. 1, p. 9501, Dec. 2019.

[17] W. B. Yousuf, T. Khan, and T. Ali, "Prognostic algorithms for flaw growth prediction in an aircraft wing," *IEEE Trans. Rel.*, vol. 66, no. 2, pp. 478–486, Jun. 2017.

[18] G. Welch and G. Bishop, "An introduction to the Kalman filter," Dept. Comput. Sci., Univ. North Carolina Chapel Hill, Chapel Hill, NC, USA, Tech. Rep., 1995.

[19] X. Zhu *et al.*, "Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer," *Eur. Radiol.*, vol. 28, no. 7, pp. 2772–2778, Jul. 2018.

[20] S. Liu *et al.*, "Exploratory study of a CT radiomics model for the classification of small cell lung cancer and non-small-cell lung cancer," *Frontiers Oncol.*, vol. 10, p. 1268, Sep. 2020.

[21] T. Khan, P. Ramuhalli, and S. C. Dass, "Particle-filter-based multisensor fusion for solving low-frequency electromagnetic NDE inverse problems," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 6, pp. 2142–2153, Jun. 2011.

[22] Y. Boers and P. K. Mandal, "Optimal particle-filter-based detector," *IEEE Signal Process. Lett.*, vol. 26, no. 3, pp. 435–439, Mar. 2019.

[23] Z. Feng, X. Wen-fang, and L. Xi, "Overview of nonlinear Bayesian filtering algorithm," *Procedia Eng.*, vol. 15, pp. 489–495, 2011.

[24] A. A. Suleiman *et al.*, "Modeling tumor dynamics and overall survival in advanced non–small-cell lung cancer treated with erlotinib," *J. Thoracic Oncol.*, vol. 10, no. 1, pp. 84–92, 2015.

[25] D. R. Aberle *et al.*, "The national lung screening trial: Overview and study design," *Radiology*, vol. 258, no. 1, pp. 243–253, Jan. 2011.

[26] Z. Zhou *et al.*, "Constructing multi-modality and multi-classifier radiomics predictive models through reliable classifier fusion," 2017, *arXiv:1710.01614*. [Online]. Available: http://arxiv.org/abs/1710.01614

[27] M. M. Koo *et al.*, "Presenting symptoms of cancer and stage at diagnosis: Evidence from a cross-sectional, population-based study," *Lancet Oncol.*, vol. 21, no. 1, pp. 73–79, Jan. 2020.

[28] H. J. W. L. Aerts *et al.*, "Data from NSCLC-radiomics [data set]. The cancer imaging archive," 2019, doi: 10.7937/K9/TCIA.2015.PF0M9REI.

[29] H. Shakir, H. Rasheed, and T. M. R. Khan, "Radiomic feature selection for lung cancer classifiers," *J. Intell. Fuzzy Syst.*, vol. 38, no. 5, pp. 5847–5855, 2020.

[30] L. Tiancheng, B. Miodrag, and D. M. Petar, "Resampling methods for particle filtering: Classification, implementation, and strategies," *Signal Process. Mag.*, vol. 32, no. 3, pp. 70–86, May 2015.

[31] S. Hawkins *et al.*, "Predicting malignant nodules from screening CT scans," *J. Thoracic Oncol.*, vol. 11, no. 12, pp. 2120–2128, Dec. 2016.

[32] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over f1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 1–13, Dec. 2020.