Ensemble Framework Based Real-time Respiratory Motion Prediction for Adaptive Radiotherapy Applications

Sivanagaraja Tatinati[a], Kalyana C. Veluvolu[a,*], Wei Tech Ang[e], Kianoush Nazarpour[c,d]

[a]*School of Electronics Engineering, College of IT Engineering, Kyungpook National University, Daegu, South Korea*
[b]*School of Life Science and Technology, Xidian University, Xi'an, Shaanxi, 710071, China*
[c]*School of Electrical and Electronic Engineering, Newcastle University, Newcastle-upon-Tyne NE1-7RU, UK*
[d]*Institute of Neuroscience, Newcastle University, Newcastle-upon-Tyne NE2 4HH, UK*
[e]*School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore*

Abstract

Successful treatment of tumors with motion-adaptive radiotherapy requires accurate prediction of respiratory motion, ideally with a prediction horizon larger than the latency in radio-therapy system. Accurate prediction of respiratory motion is however non-trivial task due to the presence of irregularities and intra-trace variabilities, such as baseline drift and temporal changes in fundamental frequency pattern etc. In this paper, to enhance the prediction of the respiratory motion traces, we propose a stacked regression ensemble framework that integrates heterogeneous respiratory motion prediction algorithms. We further address two crucial issues for developing a successful ensemble framework: 1) selection of appropriate prediction methods to ensemble (level-0 methods) among the best existing prediction methods; and 2) finding a suitable generalization approach that can successfully exploit the relative advantages of the chosen level-0 methods. The efficacy of the developed ensemble framework is assessed with real respiratory motion traces acquired from 31 patients undergoing treatment. Results show that the developed ensemble framework improves the prediction performance significantly compared to the best existing methods.

Keywords: motion-adaptive radiotherapy, respiratory motion prediction, ensemble

*Corresponding author
Email address: veluvolu@ee.knu.ac.kr (Kianoush Nazarpour)
[1]Total words: 4976 words

1. Introduction

Radiotherapy is a therapeutic alternative for patients whom are diagnosed with inoperable tumor or who can not to go through surgery due to various complications [1, 2]. Tumor movement due to respiration is one of the most challenging problems in robotic radiotherapy especially when dealing with tumors in the lungs. In Wilbert et al. [1], with correlation analysis, it was shown that movement of external body can be a good predictor of the tumor movement that was calculated using computed tomography. In the current commercial systems, such as CyberKnife and VERO, there is however an inevitable time delay of $70ms$ - $400ms$ between the actual movement of tumor and the movement obtained from the correlation model. Furthermore, this delay is device dependent and mainly due to image data acquisition, processing and mechanical limitations of the radiotherapy systems [9, 2]. To overcome this time delay and hence the positioning error, *prediction* of tumor movement with a prediction horizon equivalent to the radiotherapy system latency was proposed [1, 13]. Accurate prediction of tumor motion is however challenging because the respiratory motion traces are often subject to irregularities and intra-trace variabilities such as baseline, frequency, and temporal changes in their fundamental pattern [14, 18].

Over the last two decades, several signal processing techniques have been developed to predict the respiratory motion [1, 2, 7, 8, 9, 13, 32]. Univariate prediction approaches include methods based on state-space modeling with adaptive algorithms such as least mean squares (LMS) [13] or extended Kalman filtering (EKF) [4], or wavelet-based multi-scale regression (wLMS) [5] and methods based on machine learning techniques such as support vector machines (SVM) [30] or accurate-online SVM (SVRPred) [6] or artificial neural networks (ANN) [19], or ensemble learning [32] have been developed. Recently, to further enhance the prediction performance, multi-variate prediction approaches based on Bayesian inference [10, 11] and Gaussian processes [12] have been developed. A hybrid method was formulated by cascading EKF and SVM focused on the prediction horizons between 200ms and 600ms which are relevant to image-guided radiotherapy [15]. A comparative analysis of these algorithms

with latencies of the VERO and the CyberKnife systems was provided in [9, 10]. Results showed that, at large prediction horizons, machine learning techniques provided better prediction performance for most of the traces compared to state-space modeling methods albeit exhaustive training on large data sets.

Each respiratory motion prediction method has a signal model formulated to represent certain degree of the underlying true respiratory phenomenon. The signal model is solely based on the accumulated prior knowledge which is gained by either stochastic state-space modeling [3, 4] or heuristic machine learning techniques [10, 19]. By the virtue of its subject-dependent and non-stationary nature, accumulated prior knowledge does not represent the whole evolution of the respiratory motion because irregularities and intra-trace variabilities lead to temporal variations. Consequently, no prediction method can be superior than other methods for all subjects [20]. A comparison analysis conducted among all the prediction methods on the data acquired from 31 subjects supports this hypothesis [9, 15].

Motivated by these observations and to enhance the respiratory motion prediction performance, we propose an ensemble learning framework based on stacked regression [28]. Ensemble learning is one of the innovative ideas from the machine learning community and has been a successful technique in classification and time series forecasting [21, 22, 26, 27, 28]. Stacked regression is a way of integrating multiple prediction models to enhance the prediction performance [22, 26]. It has been shown that stacked regression significantly reduces the modeling errors and its variance when compared to the methods choose to ensemble [23, 27, 28]. Based on the bias-variance trade-off analysis for ensemble methods [21, 31], we hypothesized that ensemble of the best existing respiratory motion prediction methods can enhance the prediction performance.

A necessary and sufficient condition to design a successful ensemble learning framework for respiratory motion prediction is "the prediction methods chosen to ensemble (referred as level-0 methods in Breiman et al. [22]) should be accurate and diverse" [26, 27, 28]. Based on the techniques proposed in Breiman et al. [22] we identified level-0 methods from the pool of the best existing respiratory motion prediction methods. To exploit the

relative advantages of these level-0 methods and to ensemble, we employed machine learning techniques. We validated the proposed ensemble learning approaches with a comprehensive analysis conducted for four prediction horizon with publicly available respiratory motion database recorded from 31 patients [9, 29]. The chosen prediction horizons were well in-line with the latencies of commercially available systems.

2. Methods and Materials

2.1. Ensemble Learning: Stacked Regression

The proposed ensemble learning framework for respiratory motion prediction comprises of training stage and testing stage, as shown in Figure 1.

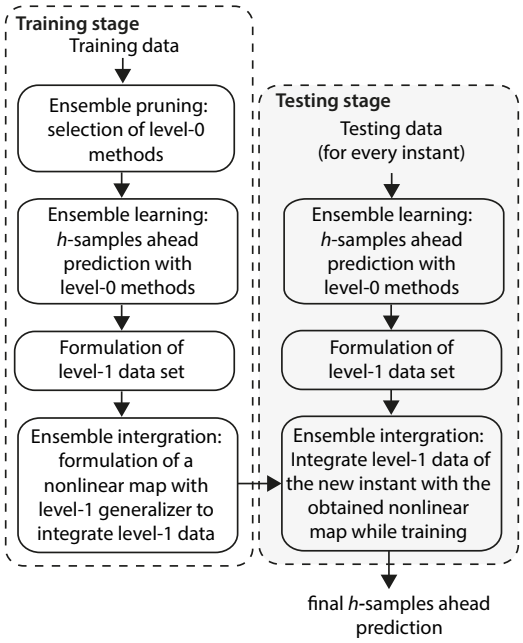

Figure 1: Schematic diagram: Framework of proposed ensemble learning approach for respiratory motion prediction.

1) Training stage: The three phases namely ensemble pruning, ensemble learning, and integration form core part in the training stage. In the ensemble pruning stage, the prediction methods that are employed as the level-0 methods, will be selected from the pool of existing respiratory motion prediction methods. To this end, diversity measures, such

as correlation and mutual information, are employed to identify the diverse and accurate prediction methods (the pruning procedure is detailed in the following sub-section). In the ensemble learning phase, the chosen level-0 methods are trained independently on a dataset $\mathcal{L}$ to perform multi-step prediction: $\mathcal{L} = \{\mathbf{s}_k,\ s_{k+h}\}_{k=1}^{N}$ with $\mathbf{s}_k = [s_k, s_{k-1}, \cdots, s_{k-n}]$ as the input vector (the value of n depends on the prediction algorithm), s_{k+h} as the output vector, and N as the number of training samples. Finally, in the ensemble integration phase, for each sample $\mathbf{s}_k$ in $\mathcal{L}$, let $\hat{s}_{k+h}^1, \hat{s}_{k+h}^2, \cdots, \hat{s}_{k+h}^{\mathcal{P}}$ denote h-samples ahead predicted values computed with the level-0 prediction methods ($\mathcal{P}$ represents number of level-0 methods chosen). Consequently for $\mathcal{L}$, the data set formulated by assembling the predicted values of level-0 prediction methods is named as level-1 data, $\mathcal{T} = \{s_{k+h}, [\hat{s}_{k+h}^1\ \hat{s}_{k+h}^2\ \cdots\ \hat{s}_{k+h}^{\mathcal{P}}]\}_{k=1}^{N}$. To derive a nonlinear map ($\mathcal{F}$) which can effectively integrate the level-1 data, machine learning techniques are employed as a level-1 generalizer.

b) Testing stage: In this stage, for every new instant s_i, level-1 data is formulated by assembling the $h-$samples ahead predicted values obtained from level-0 methods, that is $(\hat{s}_{i+h}^1, \hat{s}_{i+h}^2, \cdots, \hat{s}_{i+h}^{\mathcal{P}})$. With the formulated level-1 data and the identified nonlinear map $\mathcal{F}$, the final h- samples ahead prediction for the corresponding instant will be computed, that is $\hat{s}_{i+h} = \mathcal{F}(\hat{s}_{i+h}^1, \hat{s}_{i+h}^2, \cdots, \hat{s}_{i+h}^{\mathcal{P}})$.

2.1.1. Ensemble pruning

The success of an ensemble of prediction methods relies highly upon the *diversity* among the individual prediction methods. Diversity is defined as the degree of disagreement among the individual prediction methods [27]. According to the bias-variance trade-off analysis, an ensemble method that comprises of level-0 methods with much disagreement is more likely to attain a good final generalization performance [31, 27]. Thus, in this work, to identify appropriate level-0 methods for respiratory motion prediction, we formulated a pool of best respiratory motion prediction methods. According to [4] and [9], LCM-EKF and wLMS methods provide better performance for most of the traces compared to other existing respiratory motion prediction methods such as kernel density estimation (KDE) [17], ANN [19] and SVM [30, 9]. Accordingly, we formulated a pool of methods with LCM-EKF,

autoregressive moving average model with fading memory Kalman filter (ARMA-FMKF), wLMS, normalized least mean squares (nLMS) and multi-step linear method (MULIN).

To assess the diversity between two prediction algorithms, we employed the diversity measures proposed in [27, 28]: correlation coefficient (ρ) and mutual information (I). Let, $\hat{\mathbf{s}}^{M1} = [\hat{s}_1^{M1}, \hat{s}_2^{M1}, \cdots, \hat{s}_N^{M1}]$ and $\hat{\mathbf{s}}^{M2} = [\hat{s}_1^{M2}, \hat{s}_2^{M2}, \cdots, \hat{s}_N^{M2}]$ represent N-dimensional output vectors of two respiratory motion prediction algorithms $M1$ and $M2$.

Definition 1. *The correlation coefficient (ρ) between $\hat{\mathbf{s}}^{M1}$ and $\hat{\mathbf{s}}^{M2}$ is defined as:*

$$\rho = \frac{\sum_{i=1}^{N}(\hat{s}_i^{M1} - \mu_{\hat{\mathbf{s}}^{M1}})(\hat{s}_i^{M2} - \mu_{\hat{\mathbf{s}}^{M2}})}{\sqrt{\sum_{i=1}^{N}(\hat{s}_i^{M1} - \mu_{\hat{\mathbf{s}}^{M1}})^2 \sum_{i=1}^{N}(\hat{s}_i^{M2} - \mu_{\hat{\mathbf{s}}^{M2}})^2}}.$$

where $\mu_{\hat{\mathbf{s}}^{M1}} = \frac{1}{N}\sum_{i=1}^{N} \hat{s}_i^{M1}$ and $\mu_{\hat{\mathbf{s}}^{M2}} = \frac{1}{N}\sum_{i=1}^{N} \hat{s}_i^{M2}$ represent the averages of $\hat{\mathbf{s}}^{M1}$ and $\hat{\mathbf{s}}^{M2}$.

The diversity of two predictors is inversely proportional to the correlation between them. As such, two prediction methods with low correlation coefficient between them are preferred over those with high correlation coefficient.

Definition 2. *The mutual information between $\hat{\mathbf{s}}^{M1}$ and $\hat{\mathbf{s}}^{M2}$ is defined as:*

$$I(\hat{\mathbf{s}}^{M1}; \hat{\mathbf{s}}^{M2}) = H(\hat{\mathbf{s}}^{M1}) + H(\hat{\mathbf{s}}^{M2}) - H(\hat{\mathbf{s}}^{M1}; \hat{\mathbf{s}}^{M2})$$

where $H(\hat{\mathbf{s}}^{M1})$ and $H(\hat{\mathbf{s}}^{M2})$ are the entropies of $\hat{\mathbf{s}}^{M1}$ and $\hat{\mathbf{s}}^{M2}$ respectively and $H(\hat{\mathbf{s}}^{M1}; \hat{\mathbf{s}}^{M2})$ represents the joint differential of $\hat{\mathbf{s}}^{M1}$ and $\hat{\mathbf{s}}^{M2}$. If $\hat{\mathbf{s}}^{M1}$ and $\hat{\mathbf{s}}^{M2}$ are Gaussian random variables with variances σ_{M1}^2 and σ_{M2}^2 then $H(\hat{\mathbf{s}}^{M*}) = \frac{1}{2}[1 + log(2\pi\sigma_{M*}^2)]$ and $H(\hat{\mathbf{s}}^{M1}; \hat{\mathbf{s}}^{M2}) = 1 + log(2\pi) + \frac{1}{2}log(\sigma_{M1}^2\sigma_{M2}^2(1-\rho^2))$.

Diversity measures obtained for the chosen five best prediction methods are provided in Figure 2. To assess the diversity measures, a 5×5 matrix is constructed such that the ij^{th} element of the matrix represents the diversity measure (correlation or mutual information) between the i^{th} and j^{th} prediction methods. To construct the matrix, we performed 8 samples ahead prediction with all the methods on the respiratory motion database of 304

traces (described in Section II-B) and the parameters are initialized as provided in Table 1. In Figure 2(a) correlation obtained for all possible combinations between two prediction methods are plotted. In Figure 2(b), the other measure mutual information is plotted.

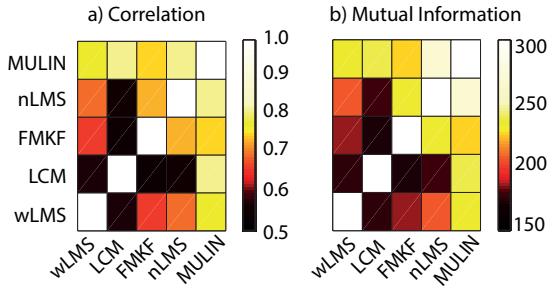

Figure 2: Diversity measures between level-0 methods a) Correlation b) Mutual information.

Correlation analysis showed that LCM-EKF & wLMS, LCM-EKF & ARMA-FMKF, and LCM-EKF & nLMS are the pairs with minimized correlation coefficients, as shown in figure 2. As aforementioned, diversity is inversely proportional to correlation, consequently these four methods are appropriate methods to ensemble. However, Figure 2 further revealed that nLMS is largely correlated with both wLMS and ARMA-FMKF. We therefore concluded that LCM-EKF, ARMA-FMKF, and wLMSshould serve as most appropriate level-0 methods.

2.1.2. Ensemble learning with level-0 prediction methods

Brief descriptions of level-0 methods employed in this work are provided here. For more details the interested reader is referred to [4, 9, 32].

a) Local circular motion with extended Kalman filter (LCM-EKF) [4]

To capture the temporal evolution of the respiratory motion, the LCM method models the respiratory motion as a one-dimensional projection of a circular motion in the x-y plane. The evolution of position $x(k)$ is the projection of the planar circular motion onto the x-axis. The y-axis is an auxiliary axis augmented to define the circular motion. Furthermore, the LCM method includes the angular velocity of the circular motion as a part of the system states to effectively capture the temporal evolution. The state-space model of the LCM

model can be given as:

$$\mathbf{x}(k+1) = f(\mathbf{x}(k)) + \nu(k) \quad (1)$$

$$s(k) = \mathbf{x}(k) + \omega(k) \quad (2)$$

where $\mathbf{x}(k) = [x(k)\ \dot{x}(k)\ \dot{y}(k)\ \Omega(k)]$ and $x(k),\ \dot{x}(k),\ y(k),\ \dot{y}(k)$ represent the position and the velocity along the x-axis and y-axis respectively; $\Omega(k)$ represents the angular velocity, and $f(\cdot)$ represents the state evolution function which is

$$f(\mathbf{x}(k)) = \begin{pmatrix} 1 & \frac{\sin\Omega(k)T}{\Omega(k)} & -\frac{1-\cos\Omega(k)T}{\Omega(k)} & 0 \\ 0 & \cos\Omega(k)T & -\sin\Omega(k)T & 0 \\ 0 & \sin\Omega(k)T & \cos\Omega(k)T & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{x}(k),$$

$\nu(k)$ represents the process noise and $\omega(k)$ represents the measurement noise. Co-variance matrices for both noise variables are modeled as derived in Hong et al. [4]

Multi-step prediction of the respiratory motion with LCM can be given as:

$$\hat{s}_{k+h}^{LCM} = \hat{x}(k) + \frac{\sin\hat{\Omega}(k)hT}{\hat{\Omega}(k)}\hat{\dot{x}}(k) - \frac{1-\cos\hat{\Omega}(k)hT}{\hat{\Omega}(k)}\hat{\dot{y}}(k)$$

where $\hat{s}_{k+h}^{LCM}$ represents the h-samples ahead predicted value computed with LCM-EKF. In [4, 15], a first-order EKF was employed to update the LCM parameters iteratively.

b) Autoregressive moving average with fading-memory Kalman filter (ARMA-FMKF)

ARMA model characterizes the temporal evolution of the respiratory motion by employing a random walk model for regression coefficients [33]. The state-space model of an ARMA system of order (p, q) can be expressed with:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \eta(k) \quad (3)$$

$$s(k) = \mathbf{w}^T(k)\ \Phi(k) + \varepsilon(k) \quad (4)$$

where $\mathbf{w} = [w_1 \cdots w_p \; w_1 \cdots w_q]^T$ represents the weights (regression coefficients), $\Phi(k) = \begin{bmatrix} s(k-1) & \cdots & s(k-p) & e(k-1) & \cdots & e(k-q) \end{bmatrix}^T$ represents delayed inputs, $e(k)$ represents the prediction error, $\eta(k)$ represents state noise and ε_n represents measurement noise. The prediction error $e(k)$ can be computed with as $e(k) = s(k) - \hat{s}(k)$. Furthermore, $\eta(k)$ and $\varepsilon(k)$ are statistically independent and are modeled with zero-mean white-noise with covariances $E[\eta(k)\eta^T(n)] = \mathbf{Q} = \sigma_\eta^2 \times \mathbf{I}$ and $E[\varepsilon(k)\varepsilon^T(k)] = R_1 = \sigma_\varepsilon^2 \; (if \; k = n)$ respectively [33]; where $\mathbf{I}$ represents identity matrix.

Multi-step prediction with ARMA model can be obtained as:

$$\hat{s}_{k+h}^{ARMA} = \hat{\Phi}(k+h)^T \hat{\mathbf{w}}(k+h)$$

where $\hat{s}_{k+h}^{ARMA}$ represents the h-samples ahead prediction obtained with ARMA model, $\hat{\mathbf{w}}(k+h) = \mathbf{w}(k)$ (the weights vector remains constant for k to $(k+h)$ samples) and $\hat{\Phi}(k+l) = [\hat{s}(k-l-1) \; \cdots \; \hat{s}(k-l-p) \; e(k-1) \; \cdots \; e(k-q)]^T; l = 1, 2, \cdots, h$ (the input vector $\hat{\Phi}(k)$ is updated iteratively).

Multi-step prediction model for ARMA performs under the assumption that the signal is stationary in a given prediction horizon. However, for respiratory motion prediction at large prediction horizons this assumption does not hold necessarily. In order to reduce the effect of prior measurements on the state vector, a fading-memory Kalman filter (FMKF) was employed. In FMKF, the covariance matrix is multiplied by a forgetting factor λ to increase the state vector variance and hence decrease its influence on the ARMA coefficients.

c) Wavelet-based multi-scale auto-regression (wLMS) [5]

Respiratory motion prediction with wLMS is carried out in two stages. In the first stage *à* trous wavelet is employed to decompose the signal into $J+1$ scales. In the second stage, with the decomposed scales, multi-scale auto regression is formulated to perform multi-step prediction [5, 9].

The decomposition of respiratory motion signal into $J+1$ scales with *à* trous wavelet

can be described as follows:

$$s_k = c_{J,k} + \sum_{i=1}^{J} \mathbf{W}_{i,k}$$

where $\mathbf{W}_{i,k} = c_{i,k} - c_{i+1,k}$; $c_{i+1,k} = \frac{1}{2}(c_{i,k-2^j} + c_{i,k})$; $c_{0,k} = s_k$ represent the wavelet scales and $c_{J,k}$ represents the smoothed signal.

The adaptive multi-scale auto-regression model formulated with the decomposed wavelet scales, can be given as:

$$\hat{s}_{k+h}^{wLMS} = \omega_{J+1,k}^T \tilde{c}_{J,k}^M + \sum_{i=1}^{J} \omega_{i,k}^T \tilde{\mathbf{W}}_{i,k}^M$$

where $\tilde{\mathbf{W}}_{i,k}^M = [\mathbf{W}_{i,k}, \cdots, \mathbf{W}_{i,k-M_i+1}]$; $C_{J,k}^M = [c_{J,k}, \cdots, c_{J,k-M_{J+1}+1}]$; M represents the signal history and $\omega_{j,k}$ represents the weights vectors. In general, multi-step prediction is based on the information from the signal history M. It is however possible that the information not available in the signal history can influence the prediction. To include this information, the following weights update was proposed [9] :

$$\omega_{j,k+1} = (1-\mu)\omega_{j,k} + \mu(\tilde{\mathbf{W}}_{i,k}^M)^{-1} s_k; \qquad \mu \in [0,1]$$

where μ is an exponential averaging parameter.

2.1.3. Ensemble Integration with level-1 generalization method

For respiratory motion prediction, LS-SVM and ANN were chosen as the level-1 generalization algorithm to learn the prediction error dynamics of LCM-EKF, ARMA-FMKF and wLMS (level-0 methods). These level-1 generalization methods formulate a nonlinear map to reduce the ensemble generalization error and consequently improve the prediction performance.

The final prediction model for h-samples ahead with ensemble framework can be given as:

$$\hat{s}_{k+h} = \mathcal{F}(\hat{s}_{k+h}^{LCM}, \hat{s}_{k+h}^{ARMA}, \hat{s}_{k+h}^{wLMS})$$

where $\mathcal{F}$ represents the nonlinear map learnt by the LS-SVM or ANN while offline training

on the level-1 data set $\mathcal{T} = \{\mathbf{s}_i,\ t_i\}_{i=1}^{N}$ with $\mathbf{s}_i = [\hat{s}_{k+h}^{LCM}, \hat{s}_{k+h}^{ARMA}, \hat{s}_{k+h}^{wLMS}]$ as input vector and $t_i = s_{k+h}$ as the corresponding output.

The ensemble framework prediction model with LS-SVM as the level-1 generalizer method and $\mathcal{T}$ the level-1 data can be given as:

$$\begin{aligned} \hat{s}_{k+h} = &\ \mathcal{F}_{LSSVM}(\hat{s}_{k+h}^{LCM}, \hat{s}_{k+h}^{ARMA}, \hat{s}_{k+h}^{wLMS}) \\ = &\ \textstyle\sum_{i=1}^{N} \alpha_i K(\mathbf{s}_i, \mathbf{s}_k) + b;\ k = N+1, \cdots, l. \end{aligned}$$

where $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_N]$ represents the Lagrangian multipliers, K denotes a radial bias function (RBF) kernel, and b represents the bias.

Hereafter, for sake of easy notations, the ensemble framework formulated with three level-0 methods and LS-SVM as level-1 generalizer is named as Ensemble-LSSVM(LCM, ARMA, wLMS). If level-1 generalizer is ANN then the ensemble framework is named as Ensemble-ANN(LCM, ARMA, wLMS)

2.2. Respiratory Motion Traces Database

The respiratory motion database employed in this paper was recorded from 31 patients during the radioactive therapy with CyberKnife at Georgetown university hospital. The database contains 304 motion traces in total. To track the respiratory motion and hence the tumor movement, three passive markers were placed on the patients' chest and abdomen areas. Each marker provided three-dimensional ($3D$) traces corresponding to the abdomen movement. The traces were acquired by using the Synchrony respiratory motion tracking by (Accuracy, Inc). The principal components obtained with principal component analysis (PCA) from the $3D$ motion traces of a marker were considered as the motion trace of the corresponding marker [9]. Thereby, three traces were acquired from three markers, namely m_1, m_2 and m_3 for each fraction. The sampling frequency was 26 Hz. For more information on the recording procedure and pre-processing of the motion traces, see [9] and [29].

In Ernst et al. [9] it was reported that the performance of a prediction algorithm over the PCA-processed traces gives a plausible estimate of its prediction performance for three-

dimensional traces. In this work, we thus evaluated and analyzed the advantages of the proposed ensemble framework to the existing best prediction methods over the PCA-processed traces. The proposed ensemble framework however can be applied to the three axes in parallel to perform the prediction in three-dimensional space.

2.3. Performance Indices

To quantify the prediction performance of methods whilst comparison analysis, we employed the root mean square error (RMSE) and the relative RMSE (Rel.RMSE) metrics.

Definition 3. *RMSE can be defined as*

$$RMSE(\hat{s}_k, s_k) = \sqrt{\frac{1}{N_s}\sum_{k=1}^{N_s} ||\hat{s}_k - s_k||^2}$$

where N_s is the number of samples, s_k is the actual signal at instant k and $\hat{s}_k$ is the predicted signal.

Definition 4. *Rel.RMSE can be defined as*

$$Rel.RMSE(\hat{s}_h^{\mathcal{K}}, s) = (1 - \frac{RMSE(\hat{s}_h^{\mathcal{K}} - s)}{RMSE(\breve{s}_h - s)}) \times 100$$

where s represents actual signal, $\hat{s}_h$ represents h-samples ahead predicted value with method $\mathcal{K}$, and $\breve{s}_h$ represents predicted signal with no prediction method. In the 'no prediction' case, the current position measurement will be employed as the predicted position measurement for the desired horizon i.e, for instance h-samples ahead prediction can be given as $\breve{s}_h = \hat{s}_{k+h} = s_k$. Thus, $RMSE(\breve{s}_h - s)$ represents the upper bound of the prediction error.

We computed mean and standard deviation for each performance index to highlight the robustness of the proposed ensemble learning over the irregularities. For completeness, we also performed a paired Student's t-test and reported the p-values to highlight the statistical significance in improvement of prediction performance with the proposed ensemble approaches when compared with the best existing respiratory motion prediction methods.

3. Results

In this section, the optimal parameter selection for the proposed method is described. We the reported a comprehensive performance analysis of the proposed ensemble framework at four prediction horizons.

3.1. Optimal Parameter Selection

Level-0 Methods

The initialization of LCM-EKF and wLMS are documented in Hong et al. [4] and Ernst et al. [9] respectively. For ARMA-FMKF model, the Akaike information criterion (AIC) [37] identified ARMA$(16, 1)$ as the optimal order for the respiration motion traces. Table 1 reports all parameters for level-0 algorithms.

Table 1: Optimal initialization values for level-0 methods

Method	Model parameters
LCM-EKF [4]	$q_1 = q_2 = 0.1$, $q_3 = 10^{-4}$, $q_4 = 10^{-5}$, $R = 10^{-4}$, $\mathbf{P}_0 = 10^{-2} \times \mathbf{I}$
ARMA-FMKF [32]	$p = 16$, $q = 1$, $R_1 = 10^{-2}$, $\mathbf{Q} = 10^{-2}\mathbf{I}$, $\mathbf{P}_0 = 10^{-2}\mathbf{I}$
wLMS [5]	$j = 3$ $\mu = 0.0204$, $M = 200$
nLMS [9]	$\mu = 0.0204$, $n = 10$
MULIN [9]	$l = 1$, $\mu = 0.5$

Ensemble Framework

The performance of LS-SVM is influenced by the number of training samples N and the hyper-parameters i.e., regularization constant C, and RBF kernel variance σ:

a) Number of training samples (N)

In general N determines the learning and generalization capability of LS-SVM. Furthermore, N helps level-1 generalizer exploit the underlying relationship between level-0 methods by formulating a nonlinear map. Thus, to identify the optimal initialization for N, we conducted a study on the whole database for various values of N and obtained the corresponding

Table 2: Summary of trace-wise performance analysis for all methods at all prediction horizons [*The figures for Ensemble*(·) represents figures for both Ensemble-LSSVM(·) and Ensemble-ANN(·) as Ensemble-LSSVM(·) (Ensemble-ANN(·))]

Methods	RMSE (mm) Mean	RMSE (mm) SD	Rel.RMSE
(a) Prediction horizon: 77ms (2 samples)			
No prediction	0.257	0.151	–
ARMA-FMKF [32]	0.165	0.121	38.42%
LCM-EKF [4]	0.159	0.105	39.38%
wLMS [5]	0.135	0.089	47.47%
SVRpred [6]	0.193	0.140	24.91%
Ensemble*(LCM, ARMA)	0.151 (0.157)	0.139 (0.116)	41.08% (38.91 %)
Ensemble*(LCM, wLMS)	0.206 (0.171)	0.149 (0.134)	21.31% (33.46 %)
Ensemble*(wLMS, ARMA)	0.185 (0.158)	0.131 (0.129)	29.62% (38.52 %)
Ensemble*(LCM, ARMA, wLMS)	0.120 (0.145)	0.095 (0.109)	53.31% (43.59 %)
(b) Prediction horizon: 115ms (3 samples)			
No prediction	0.381	0.225	–
ARMA-FMKF [32]	0.243	0.182	37.58%
LCM-EKF [4]	0.217	0.144	43.31%
wLMS [5]	0.191	0.128	49.86%
SVRpred [6]	0.221	0.161	41.99%
Ensemble*(LCM, ARMA)	0.208 (0.209)	0.204 (0.147)	48.25% (46.31%)
Ensemble*(LCM, wLMS)	0.215 (0.218)	0.176 (0.165)	43.56% (43.25%)
Ensemble*(wLMS, ARMA)	0.227 (0.201)	0.167 (0.143)	41.52% (48.64 %)
Ensemble*(LCM, ARMA, wLMS)	0.171 (0.179)	0.156 (0.131)	55.24 % (53.01%)
(c) Prediction horizon: 154ms (4 samples)			
No prediction	0.484	0.291	–
ARMA-FMKF [32]	0.339	0.253	33.86%
LCM-EKF [4]	0.278	0.184	44.35%
wLMS [5]	0.253	0.165	47.72%
SVRpred [6]	0.288	0.191	37.21%
Ensemble*(LCM, FMKF)	0.266 (0.264)	0.221 (0.179)	48.99% (48.52%)
Ensemble*(LCM, wLMS)	0.292 (0.269)	0.199 (0.196)	41.93% (46.46%)
Ensemble*(wLMS, FMKF)	0.283 (0.249)	0.202 (0.17)	45.32% (51.55%)
Ensemble*(LCM, FMKF, wLMS)	0.232 (0.238)	0.187 (0.161)	52.01% (50.82%)
(d) Prediction horizon: 308ms (8 samples)			
No prediction	0.905	0.541	–
ARMA-FMKF [32]	0.863	0.638	11.38%
LCM-EKF [4]	0.545	0.351	42.41%
wLMS [5]	0.517	0.316	44.71%
SVRpred [6]	0.508	0.317	45.66%
Ensemble*(LCM, ARMA)	0.526 (0.503)	0.371 (0.306)	41.87% (44.41%)
Ensemble*(LCM, wLMS)	0.519 (0.479)	0.332 (0.309)	42.65% (49.81%)
Ensemble*(wLMS, ARMA)	0.599 (0.467)	0.471 (0.286)	39.25% (48.39%)
Ensemble*(LCM, ARMA, wLMS)	0.442 (0.464)	0.275 (0.323)	51.11% (49.83%)

average RMSE for prediction horizon of 308ms (8 samples). Results show that $N = 1000$ provides an optimal trade-odd between the prediction error (RMSE) and the computational complexity (number of operations) as shown in Fig. 3.

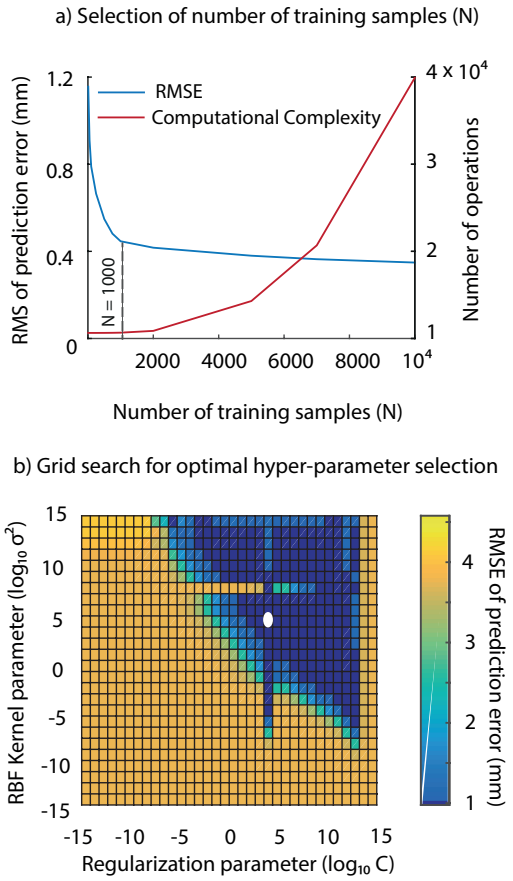

Figure 3: Parameter selection for level-1 generalizer of Ensemble-LSSVM (LCM, FMK-F,wLMS).

Hyper-parameters selection (C, σ^2)

To identify the optimal initialization of LS-SVM hyper-parameters and then to train, each trace was divided into two parts: 1) training samples (including the validation set) and 2) testing samples. The initial 1100 samples of each trace were considered as the training data set. The nonlinear map with level-1 method was then estimated by a 10-fold cross validation (90% for training and 10% for validation) based on a stacked generalization procedure to combine the level-0 methods. A grid search was conducted with a wide range of values

($0 \leq C \leq 1000$ and $0.1 \leq \sigma^2 \leq 20$). The search was conducted for each subject and each marker independently. Results for a typical subject (marker m_1) is provided in Figure 3(b). The region with lowest RMSE (shown in blue) is the desired region. One can pick any set of values for C and σ^2 from this region.

ANN being the ensemble learning algorithm, the parameters that require optimal initialization were: number of hidden layers h_n and the learning rate of the back-propagation algorithm. Optimal initialization was carried out with a grid search conducted on the first 5 breathing cycles in each respiratory motion trace ($N = 1000$) with wide range of values ($2 \leq h_n \leq 10$ and $10^{-4} \leq \sigma^2 \leq 50^{-1}$). The parameter selection was both subject-and marker-specific.

We followed the same approach to identify the optimal initialization for level-1 generalizers of ensemble methods formulated with two level-0 methods.

3.2. Performance analysis

Predictions were performed at four prediction horizon lengths: 77ms (2 samples), 115ms (3 samples), 154ms (4 samples) and 308ms (8 samples). These lengths were selected considering the typical latencies of commercial robotic radiotherapy systems: $\sim$ 77ms in VERO systems [2], $\sim$ 115ms in CyberKnife [13], $\sim$ 170ms in Tomotherapy [38] and latencies up to several hundred milliseconds for typical couch or multi-leaf collimator (MLC) tracking devices [13, 9]. A comparative analysis was then carried out among 1) *no prediction*, 2) level-0 methods (LCM-EKF, ARMA-FMKF, and wLMS), and 3) Ensemble methods: Ensemble-LSSVM (LCM, ARMA, wLMS), Ensemble-ANN(LCM, ARMA, wLMS). For completeness, we performed the comparative analysis with ensemble methods formulated by two level-0 methods: Ensemble-LSSVM (LCM, ARMA), Ensemble-LSSVM(ARMA, wLMS), Ensemble-LSSVM(LCM, wLSM) and its counterparts with ANN.

3.2.1. Performance comparison of Ensemble-LSSVM and Ensemble-ANN

The level-0 methods parameters are optimally initialized with the values provided in Table 1. Trace-wise analysis was performed for all ensemble methods and tabulated the

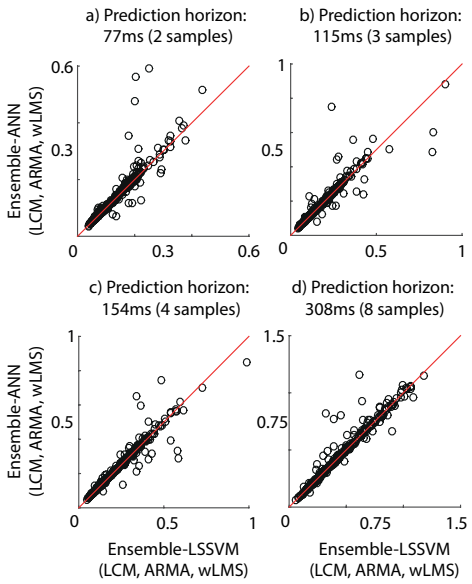

Figure 4: Scatter plots for RMSE of Ensemble-LSSVM(LCM, FMKF, wLMS) against RMSE of Ensemble-ANN(LCM, FMKF, wLMS): a) for prediction horizon of 77ms; b) for prediction horizon of 115ms; c) for prediction horizon of 154ms; and d) for prediction horizon of 308ms.

aforementioned statistics for prediction error in Table 2. This statistical analysis showed that for all traces, irrespective of prediction horizon, Ensemble-LSSVM(LCM, FMKF, wLMS) yields the least prediction error compared to other variants of the proposed ensemble learning methods. For instance, at the prediction horizon 115ms Ensemble-LSSVM(LCM, FMKF, wLMS) reduces the prediction error by 5% compared to RMSE of Ensemble-ANN(LCM, FMKF, wLMS). For all prediction horizons, the same trend was observed for RMSE and Rel.RMSE (Table. 2).

To further highlight the reduction in prediction error with Ensemble-LSSVM(LCM, FMKF, wLMS) compared to Ensemble-ANN(LCM, FMKF, wLMS), scatter plots for trace-wise RMSE of Ensemble-LSSVM(LCM, FMKF, wLMS) against RMSE of Ensemble-ANN(LCM, FMKF, wLMS) for all prediction horizons are shown in Figure 4. If the marker lies above the diagonal line, it denotes that for that particular subject Ensemble-LSSVM(LCM, FMKF, wLMS) is providing less prediction error compared to Ensemble-ANN(LCM, FMKF, wLMS). For most of the subjects the Ensemble-LSSVM(LCM, FMKF, wLMS) provided less prediction error compared to its counterpart, as shown in Figure 4. Furthermore, with the increase

in prediction horizon, the number of makers residing above the diagonal line increases. For completeness, we report that prediction performances of all ensemble methods formulated with two level-0 methods are comparable or less than the prediction performance of ensemble methods formulated with three level-0 methods. The null hypothesis was rejected with strong evidence $p < 10^{-1}$ (paired Student's t-test) for all prediction horizons. The results hence confirm that the proposed ensemble learning approaches reduce the prediction error significantly by adapting to the intra-trace variabilities.

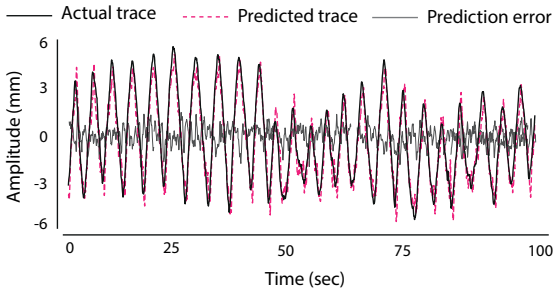

Figure 5: Prediction performance of the Ensemble-LSSVM(LCM, ARMA, wLMS) method on a typical respiratory motion trace for prediction horizon of 8 samples

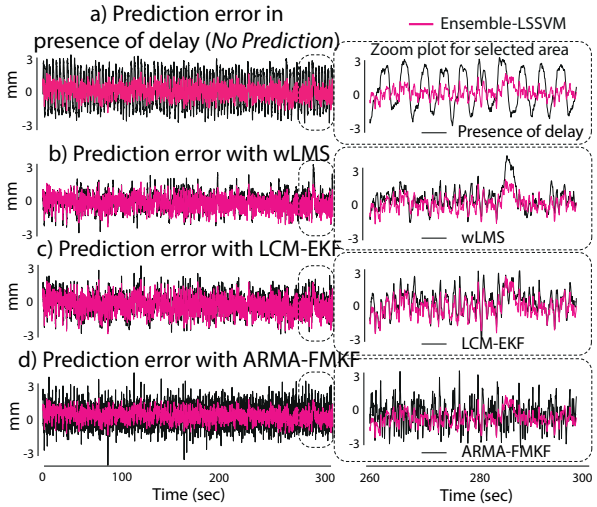

Figure 6: Comparison analysis on a typical respiratory motion trace for prediction horizon of 8 samples

3.2.2. Performance comparison of ensemble learning approaches with existing methods

The prediction performance of Ensemble-LSSVM(LCM, ARMA, wLMS) for prediction horizon of 8 samples (308ms) on an example respiratory motion trace is shown in Figure 5. To underscore the improvement in prediction performance compared to level-0 methods, comparison between the prediction error obtained with each level-0 method to the prediction error obtained with ensemble method for the above example trace is shown in Figure 6. From Figure 5 and Figure 6, we concluded that the ensemble method improves the prediction performance compared to chosen level-0 methods.

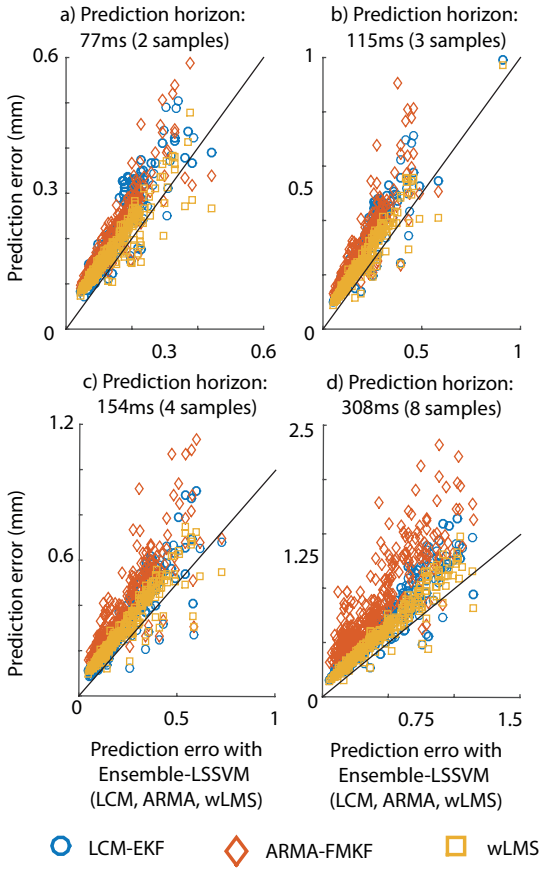

Figure 7: Scatter plots for RMSE of Ensemble-LSSVM(LCM, FMKF, wLMS) against RMSE of level-0 methods (LCM-EKF, ARMA-FMKF, and wLMS) in a) at prediction horizon of 77ms in b) at prediction horizon of 115ms in c) at prediction horizon of 154ms and in d) at prediction horizon of 308ms.

Scatter plots for trace-wise RMSE of Ensemble-LSSVM(LCM, FMKF, wLMS) against

RMSE of LCM-EKF, RMSE of wLMS and RMSE of ARMA-FMKF highlights the reduction in prediction error with ensemble learning when compared to the level-0 methods, shown in Fig. 7. RMSE obtained with Ensemble-LSSVM(LCM, FMKF, wLMS) is set as reference for this analysis. In these scatter plots, if any marker lies above the diagonal line, it denotes that Ensemble-LSSVM(LCM, FMKF, wLMS) provides less prediction error compared to the corresponding method. Scatter plots show that for most of the subjects, the proposed ensemble learning method provides less prediction error compared to the level-0 methods.

The statistics provided in Table 2 show that for all prediction horizons Ensemble-LSSVM (LCM, FMKF, wLMS) yields less prediction error compared to its level-0 methods. For instance, at prediction horizon 308ms, Ensemble-LSSVM(LCM, FMKF, wLMS) betters the prediction performance by 40%, 9%, 7% and 6% compared to *no prediction*, ARMA-FMKF, LCM-EKF, wLMS and SVRpred respectively. For all prediction horizons the same trend was observed. To confirm the statistical significance of reduction in prediction error with the employment of Ensemble-LSSVM(LCM, FMKF, wLMS) compared to other ensemble approaches, paired Student's t-tests with Bon-ferroni correction for multiple comparisons were performed. For all combinations and for all prediction horizons, the null hypothesis was rejected with $p < 10^{-2}$.

4. Discussions

Respiratory motion prediction with the horizon of system latencies was proposed to reduce the tumor positioning error in motion-adaptive radiotherapy. In this work, we developed an ensemble learning framework based on stacked regression to enhance the respiratory motion prediction capabilities and hence reduce the positioning error. A trade-off between bias and variance of MSE was required to determine the model complexity. By increasing the learning model complexity, variance of the model increases whereas the bias steadily decreases and vice versa. However, with ensemble learning techniques, the variance of the MSE is reduced without affecting the bias and model complexity. Consequently, ensemble method provided better prediction performance than level-0 methods. This supports the hypothesis of ensemble to improve the respiratory motion prediction performance.

Based on the diversity measures, we chose LCM-EKF, ARMA-FMKF, and wLMS as the level-0 methods and LS-SVM as level-1 generalizer. Each level-0 method has a unique signal model to predict respiratory motion. The LCM-EKF method characterizes the respiratory motion as a circular motion in augmented delayed axis (refer to equations (1) and (2)). To update the LCM model parameters according to the evolution of circular path an adaptive algorithm first-order EKF was employed. The ARMA model is a stochastic modeling technique and is solely based on the signal history (refer to equations (3) and (4)). To adaptively update the ARMA model, we employed FMKF. The wLMS replaces the signal history in ARMA model with the decomposed wavelet scales (refer to equations (6) and (7)). Furthermore, this model updates its weights based on the variations in the signal history (decomposed wavelet scales) rather than the error signal. With these chosen three diverse level-0 methods, we developed variants for the ensemble framework. Analysis conducted on the respiratory motion data base (304 traces) showed that for a $115ms$ prediction horizon our developed ensemble method (Ensemble-LSSVM (ARMA, LCM, wLMS)) improved the prediction performance (Rel.RMSE) by 55%, 18%, 12%, 13%, and 6% compared to *no prediction*, ARMA-FMKF, LCM-EKF, SVRpred, and wLMS methods respectively.

Computational complexity is one of the parameters that determines the efficacy of a method in real-time implementations. In the proposed ensemble framework, the training of level-0 methods is ideal for parallel computation because the chosen methods are diverse and independent to each other. Furthermore, the level-1 generalizer (LS-SVM or ANN) is trained off-line to identify the nonlinear map that can ensemble the level-0 methods. A trained LS-SVM or ANN can be implemented in real-time applications.

If $C_{LCM},\ C_{ARMA},$ and, C_{wLMS} are the computational complexities of chosen level-0 methods which are feasible for real-time implementation and $C_{LSSVM} = \mathcal{O}(N)$ is the computational complexity of the LS-SVM for generalizing the level-0 methods, then the total computational complexity of proposed ensemble modeling is $Ce = max(C_{LCM},\ C_{ARMA}, C_{wLMS}) + C_{LSSVM}$. For the proposed ensemble approach it is $Ce = C_{wLMS} + C_{LSSVM}$. Consequently, the proposed ensemble learning method is feasible for real-time respiratory motion predic-

tion.

In the context of respiratory motion prediction with ensemble learning, selection of accurate and diverse level-0 methods and level-1 generalizer is vital. We chose possibly less computationally complex yet diverse state-space modeling methods as level-0 methods. One can select machine learning techniques such as ANN or SVM as potential level-0 method(s) too. However, these techniques require computationally expensive re-training procedure to update the algorithm parameters according to the available new database at regular intervals. Recently in [10, 12], probabilistic approaches based multi-variate prediction and gating of radiation beam were proposed for motion-adaptive radiotherapy. In this work, we limited our scope to uni-variate prediction. The ensemble framework based on online re-trained machine learning techniques and probabilistic approaches as level-0 methods will be discussed elsewhere.

5. Conclusions

To enhance the performance of respiratory motion traces prediction, an ensemble learning framework is proposed in this paper. We addressed two crucial issues for successful implementation of ensemble learning (stacked regression) for respiratory motion prediction: first, identification of LCM-EKF, wLMS and ARMA-FMKF as the appropriate level-0 methods; second, selection of LS-SVM as the level-1 generalizer to accurately ensemble the level-0 methods. Eight variants of ensemble methods were developed for respiratory motion prediction. To evaluate the performance of proposed methods, analysis was conducted on a database collected form 31 patients. The analysis was performed for four prediction horizons 77ms, 115ms, 154ms, and 308ms that are in line with the commercially available robotic radiotherapy devices. Results showed that the Ensemble-LSSVM(LCM, wLMS, FMKF) algorithm provides most accurate prediction of respiratory motion traces for all prediction horizons. The improvement in prediction performance obtained with other variants of ensemble learning algorithm further supports our hypothesis that the ensemble of best existing prediction methods yields even more accurate respiratory motion prediction.

Acknowledgments

This research was supported by the Basic Science Research Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2014R1A1A2A10056145).

Conflicts of Interest

The authors declare no conflict of interest.

Ethical Approval

Not required.

References

[1] Wilbert J, Meyer J, Baier K, Guckenberger M, Herrmann C, Hess RJ, Ma L, Mersebach T, Richter A, Roth M, Schilling K, Flentje M. Tumor tracking and motion compensation with an adaptive tumor tracking system (ATTS): System description and prototype testing. Medical Physics 2008; 35: 3911–21.

[2] Poulsen PR, Cho BC, Sawant A, Ruan D, Keall PJ. Detailed analysis of latencies in image-based dynamic MLC tracking. Medical Physics 2010; 37:4998–05.

[3] McCall KC, Jeraj R. Dual component model of respiratory motion based on periodic autoregressive moving average (periodic ARMA) method. Physics in Medicine and Biology 2007; 52:3455–66.

[4] Hong SM, Jung BH, Ruan D. Real-time prediction of respiratory motion based on local dynamic model in an augmented space. Physics in Medicine and Biology 2011; 56:1775–89.

[5] Ernst F, Schlaefer A, Schweikard A. Prediction of respiratory motion with wavelet-based multi-scale autoregression. In: Ayache N, Ourselin S, MA, Editors, Lecture Notes in Computer Science, Springer. MICCAI 2007. p.668–75.

[6] Ernst F, Schweikard A. Forecasting respiratory motion with accurate online support vector regression (SVRpred). International Journal of Computer Assisted Radiology and Surgery 2009; 4:439–447.

[7] Lujan AE, Larsen EW, Balter JM, Ten Haken RK. A method for incorporating organ motion due to breathing into 3D dose calculations. Medical Physics 1999; 26:715–20.

[8] Keall P, Kini VR, Vedam SS, Mohan R. Potential radiotherapy improvements with respiratory gating. Australasian Physical and Engineering Sciences in Medicine 2002; 25:1–6.

[9] Ernst F, Durichen R, Schlaefer A, Schweikard A. Evaluating and comparing algorithms for respiratory motion prediction. Physics in Medicine and Biology 2013; 58:3911–29.

[10] Durichen R, Wissel T, Ernst F, Schlaefer A, Schweikard A. Multivariate respiratory motion prediction. Physics in Medicine and Biology 2014; 59:6043–6060.

[11] Durichen R, Wissel T, Ernst F, Schlaefer A. Respiratory motion compensation with relevance vector machines. In *M*edical Image Computing and Computer-Assisted Intervention ? MICCAI 2013, pp 108-115, 2013.

[12] Durichen R, Pimentel M A F, Ernst F, Clifton L, Schweikard A, Clifton D A. Multitask Gaussian processes for multivariate physiological time-series analysis. IEEE Transactions on Biomedical Engineering 2015; 62:314–322.

[13] Seppenwoolde Y, Berbeco RI, Nishioka S, Shirato H, Heijmen B. Accuracy of tumor motion compensation algorithms from a robotic respiratory tracking system: a simulation study. Physics in Medicine and Biology 2007; 34:2774–84.

[14] Ruan D, Fessler JA, Balter JM, Keall P. Real-time profiling of respiratory motion: baseline drift, frequency variation and fundamental pattern change. Physics in Medicine and Biology 2009; 54:4777–92.

[15] Hong SM, Bukhari W. Real-time prediction of respiratory motion using a cascade structure of an extended Kalman filter and support vector regression. Physics in Medicine and Biology 2014; 59:3555–73.

[16] Putra D, Haas OC, Mills JA, Burnham KJ. A multiple model approach to respiratory motion prediction for real-time IGRT. Physics in Medicine and Biology 2008; 53:1651–53.

[17] Ruan D. Kernel density estimation-based real-time prediction for respiratory motion. Physics in Medicine and Biology 2010; 55:1311–26.

[18] Hamalainen RP, Kettunen A. Stability of Fourier coefficients in relation to changes in respiratory air flow patterns. Medical Engineering and Physics 2000; 22:733–39.

[19] Murphy MJ, Pokhrel D. Optimization of an adaptive neural network to predict breathing. Physics in Medicine and Biology 2009; 34:40–7.

[20] Wolpert DH, Macready WG. No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation 1997; 1:67–82.

[21] Dzeroski S, Zenko B. Is combining classifiers with stacking better than selecting the best one?. Machine Learning 2004; 54:255–73.

[22] Breiman L. Stacked regressions. Machine Learning 1996; 24:49–64.

[23] Ho TK, Hull JJ, Srihari SN. Decision combination in multiple classifier systems. IEEE Transactions on Pattern Analysis and Machine Intelligence 1994; 16:66–75.

[24] Yu L, Lai KK, Wang S, Huang W. A bias-variance-complexity trade-off framework for complex system modeling. In *Computational Science and Its Applications-ICCSA 2006*, pp. 518-527, 2006.

[25] Rodrigues PL, Rodrigues NF, Pinho ACM, Fonseca JC, Corriea-Pinto J, Vilaca JL. Automatic modeling of pectus excavatum corrective prosthesis using artificial neural networks. Medical Engineering and Physics 2014; 36:1338–45.

[26] Ting KM, Witten IH. Issues in stacked generalization. Journal of Artificial Intelligence Research 1999; 10:271–89.

[27] Brown G, Wyatt JL, Tino P. Managing diversity in regression ensembles. Journal of Machine Learning Research 2005; 6:1621–50.

[28] Rooney N, Patterson D, Nugent C. Non-strict heterogeneous stacking. Pattern Recognition Letters 2007; 28:1050–61.

[29] Ernst F. Compensating for quasi periodic motion in robotic radiosurgery. New York, NY: Springer; 2011.

[30] Riaz N, Shanker P, Wiersma R, Gudmundsson O, Mao W, Widrow B, Xing L. Predicting respiratory tumor motion with multi-dimensional adaptive filter and support vector regression. Physics in Medicine and Biology 2009; 54: 5735–48.

[31] Valentini G, Dietterich TG. Bias-variance analysis of support vector machines for the development of SVM-based ensemble methods. Journal of Machine Learning Research 2000; 1:1–48.

[32] Tatinati S, Veluvolu KC, Hong SM, Nazarpour K. Real-time prediction of respiratory motion traces for radiotherapy with ensemble learning. In proceedings of 36th Annual International Conference of IEEE Engineering in Medicine and Biology Society 2014; p.4204–07.

[33] Haykin S. Adaptive Filter Theory. Upper Saddle River, NJ: Prentice-Hall; 2002.

[34] Qijun X, Ming R, Yiqun Y, Xuemin S. Adaptive fading Kalman filter with an application. Automatica 1994; 30:1333–38.

[35] Romero E, Toppo D. Comparing support vector machines and feed-forward neural network with similar hidden layer weights. IEEE Transactions on Neural Networks 2007; 18:959–63.

[36] Suykens JAK, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J. Least Squares Support Vector Machines. World scientific; 2002.

[37] Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control 1974; 19:716–23.

[38] Lu W. Real-time motion-adaptive delivery (MAD) using binary MLC: II rotational beam (tomotherapy) delivery. Physics in Medicine and Biology 2008; 53:6513–31.