

Respiratory Motion Prediction Using Deep Convolutional Long Short-Term Memory Network

Abstract

Background: Pulmonary movements during radiation therapy can cause damage to healthy tissues. It is necessary to adapt treatment planning based on tumor motion to avoid damage to healthy tissues. A range of approaches has been proposed to monitor the issue. A treatment planning based on fourdimensional computed tomography (4D CT) images can be addressed as one of the most achievable options. Although several methods proposed to predict pulmonary movements based on mathematical algorithms, the use of deep artificial neural networks has recently been considered. **Methods:** In the current study, convolutional long shortterm memory networks are applied to predict and generate images throughout the breathing cycle. A total of 3295 CT images of six patients in three different views was considered as reference images. The proposed method was evaluated in six experiments based on a leaveonepatientout method similar to crossvalidation. **Results:** The weighted average results of the experiments in terms of the rootmeansquared error and structural similarity index measure are 9×10^{-3} and 0.943, respectively. **Conclusion:** Utilizing the proposed method, because of its generative nature, which results in the generation of CT images during the breathing cycle, improves the radiotherapy treatment planning in the lack of access to 4D CT images.

Keywords: Convolutional long short-term memory, deep neural network, lung motion, radiotherapy, respiratory motion prediction

Submitted: 31-Jul-2019

Revised: 04-Sep-2019

Accepted: 09-Oct-2019

Published: 25-Apr-2020

Introduction

According to statistics, lung cancer shows the most significant mortality among cancers in the world,^[1] and this makes it necessary to use modern methods to manage the treatment. Pulmonary movement due to respiration can be addressed as one of the challenges of the treatment of pulmonary tumors, which leads to the displacement of the target and the possibility of damage to healthy tissue surrounded.^[2]

It is essential to adapt treatment planning based on target motion to spare healthy tissues. The tumor position could be monitored over time using four-dimensional computed tomography (4D CT) images, those are achieved by adding time to the 3D CT. The treatment planning based on the use of 4D CT images, so-called 4D radiotherapy, can reduce the dose of radiation transmitted to the healthy tissue around the tumor.^[3] Since estimating the

next position of the tumor at any moment in the breathing cycle is of great importance, the acquisition of CT images continuously throughout the breathing cycle (4D CT imaging), provides useful information about the tumor position at any moment.^[4]

Lung motion prediction models are used to estimate pulmonary movements during a complete cycle of respiration.^[5] Several methods have been proposed as a model for estimating pulmonary movements in the past decades.^[6-11] The general aim of the motion models is to find the relationship between the surrogate data and the amount of motion obtained from the imaging data. Finally, the motion model should be able to estimate motion in all respiratory positions.^[12] However, preliminary data are required for estimating motion that may sometimes not be appropriately available.

The use of machine learning methods has been recently considered in estimating pulmonary movements. Several studies based on artificial neural network approaches have been effectively able to predict

Shahabedin Nabavi¹,
Monireh Abdoos¹,
Mohsen Ebrahimi Moghaddam¹,
Mohammad Mohammadi^{2,3}

¹Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran, ²Department of Medical Physics, Royal Adelaide Hospital, ³Department of Medical Physics, School of Physical Sciences, The University of Adelaide, Adelaide, Australia

Address for correspondence:
Mr. Shahabedin Nabavi,
Faculty of Computer Science
and Engineering,
Shahid Beheshti University,
Tehran, Iran.
E-mail: s_nabavi@sbu.ac.ir

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.JMSS_38_19

Quick Response Code:



How to cite this article: Nabavi S, Abdoos M, Moghaddam ME, Mohammadi M. Respiratory motion prediction using deep convolutional long short-term memory network. *J Med Signals Sens* 2020;10:69-75.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

pulmonary behaviors.^[13-18] With the advent of deep learning, there has been a widespread change in various fields, including medical interventions and medical image analysis, and the emergence of this method has had massive effects on machine learning.^[19] Recently, a range of studies has used deep learning to predict lung movements.^[20-23] Among these, some studies have used the concept of recurrent neural networks (RNN) to develop a model for predicting pulmonary movements.^[20,21,23] Methods proposed have tried predicting the next position of the tumor based on its current position by considering a portion of the data as a training set and the remaining part as a test set. RNNs memorize the relationship between the elements of input sequences as historical information in their hidden neurons and understand how elements transform and behave.^[19]

The study of Kai *et al.*^[20] focuses on the use of RNNs for lung motion estimation, and this approach has been used to predict the trajectories of lung tumors for radiotherapy purposes. The proposed model error value has been reported <1 mm in 3D space in the term of the root-mean-square error (RMSE). In this study, the use of RNNs has been compared with classical neural networks for estimating pulmonary movements, which results in better performance of RNNs. The study by Lin *et al.*^[21] is an attempt to use a special kind of RNN so-called long short-term memory (LSTM). LSTMs are a kind of RNN architecture that includes feedback connections in addition to feedforward connections.^[24] This type of RNNs can be used to predict time series data such as respiratory signal. LSTM was used to predict respiratory signals on 1703 sets of real-time position management data, and RMSE of 0.048 and 0.139 has been reported in internal and external validity data, respectively. The results of the proposed method based on deep LSTM show the superiority of this method over conventional neural networks. In the study of Park *et al.*,^[22] intra- and inter-fraction fuzzy deep learning method has been used to predict tumor motions due to respiration and also decrease computational time. Using this method improved the RMSE by 29.98%. The study of Wang *et al.*^[23] is an effort to improve the effectiveness of radiotherapy treatment based on real-time tumor motion prediction during treatment sessions. A deep bidirectional LSTM network has been proposed for this prediction, and the RMSE of 0.097 mm has been reported on the 103 malignant lung tumor patients' respiratory motion data. The results of this study are about five times better than the results of autoregressive integrated moving average method^[25] for motion estimation.

In the field of computer vision, various methods have been proposed to predict the next frames in the video so that these methods work on the current frame to predict the subsequent frame using a generative approach.^[26-32] Given that the 4D CT images can eventually be viewed as a video, we utilize this feature to take advantage of the capabilities of video predicting methods. The predictive coding

network (PredNet) is one of the state-of-the-art architectures in the field of predicting the next frame in natural videos.^[28]

The current study is an attempt to use PredNet, a convolutional LSTM network architecture, to predict future slices based on the current slices in the breathing cycle, which can be used to determine the amount of pulmonary movement. For this purpose, the 4D CT images are used as reference images, and each slice is generated based on this generative method and compared with its corresponding slice.

The rest of this paper is organized as follows. Section II introduces the structure of the proposed convolutional LSTM Network, the material required and refers to the designed experiments. In section III, the results of the experiments are presented and discussed. Section IV presents the conclusion of this paper.

Material and Methods

Dataset preparation

4D CT images of six patients with pulmonary tumors were collected. 4D CT images were acquired on a 16 Slice Brilliance CT Big Bore Oncology™ configuration (Philips).^[33] These images were examined for more accurate observation of tumor position in three coronal, sagittal, and axial views, and all slices with tumor visibility were stored as reference images. The number of images per view for each patient is presented in Table 1. A total of 3295 CT slices was selected for these six patients in different views.

Methodology

The PredNet architecture consists of four main modules: the input convolutional layer that receives the current frame as an input, the recurrent layer, which represents a recurrent convolutional layer, the prediction layer, and error representation module. Figure 1 presents this architecture.

The RNNs allow the effects of previous inputs to be stored like memory and affect the output of the next step. In terms of time series data, LSTM, an RNN architecture, can be used to predict. The LSTM receives a 1D input then it does not look good to work on video or image datasets.

Table 1: Number of images per view per patient

Patient number	View			Total
	Coronal	Sagittal	Axial	
Patient 1	189	196	140	525
Patient 2	189	217	182	588
Patient 3	182	210	175	567
Patient 4	210	350	140	700
Patient 5	119	140	140	399
Patient 6	180	180	156	516
Total	1069	1293	933	3295

and can be utilized to determine the accuracy of predictive models. To evaluate the predictive model using the k-fold cross-validation method, the size of all k subsets must be relatively identical. To prevent the impact of a particular patient's data on the model, the images related to this patient that is considered for the test set are not used to train, and thus the size of the folds is not the same.

Six separate training and testing stages were performed. At the first stage, the images of the first patient and five other patients were used for testing and training, respectively. Then, this process was repeated by separating the data of the second patient as a test set and selecting data of other five remaining patients for training and so on. We use the weighted average to compute the average results of the leave-one-patient-out validation method for six different patients in terms of the variable size of each fold. Eq. 6 shows how to calculate the weighted average error.

$$\text{weighted average error} = \sum_{i=1}^6 \frac{|f_i|}{|F|} \times \text{err}(i) \quad \text{Eq. 6}$$

Where $|f_i|$, $|F|$, and $\text{err}(i)$ are the number of images in the i -th test fold, the cardinality of the data set and the error of the i -th test fold, respectively.

Quantitative evaluation metrics

Quantitative evaluation of the predicted results is reported in terms of RMSE and structural similarity index measure (SSIM). SSIM represents an objective method for examining the difference between reference and predicted images and is used because this method is well suited to the characteristics of the human visual system and can quantify the differences perceptually.^[38] The calculation method for these values is given in Eqs. 7 and 8.

$$\text{RMSE}(x_i, y_i) = \left(\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \right)^{\frac{1}{2}} \quad \text{Eq. 7}$$

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad \text{Eq. 8}$$

Where x and y represent the reference image and the synthetic generated image, respectively.

Results and Discussion

Evaluation of experiments

The results of pulmonary motion prediction based on the convolutional LSTM network are presented in Figure 2. The reference images, the predicted subsequent frames and their difference maps in all three views are shown for the next six frames over time. The prediction of images based on the use of convolutional LSTM model can generate the next frames so that pulmonary movements during the breathing cycle can be monitored. The generative nature of this model makes it possible to generate the next frames in the lack of access to 4D CT images based on the trained model.

Application of the model in training and generating images in different views was done to determine the positional changes of pulmonary tissues during the respiratory cycle in all three superior-inferior, anterior-posterior, and lateral directions. Determining the amount of pulmonary movement in all three directions based on generated images can help the radiotherapy team to better plan and protect healthy tissues.

The quantitative results of the comparison of reference and predicted images in each of the six experiments and the weighted average of these results to estimate the accuracy of the method are presented in Table 2. Less RMSE indicates that the reference and the generated image are more similar. It should be noted that the SSIM can take a value from -1 to 1 , and the greater scores indicate greater similarity between reference and predicted images.

Analysis of the proposed method

Although the PredNet method has been used in this study, the innovation of study is in using this state-of-the-art method of predicting future frames in video prediction for generating CT images during the breathing cycle. These images generated can be used to predict the extent of tumor motion and improve radiotherapy treatment planning. It should be noted that the use of state-of-the-art architectures is common in studies that utilize deep learning techniques, and what matters is adapting the network architecture to the desired application. To the best of our knowledge, this is the first study to use convolutional LSTM networks for pulmonary motion prediction. One of the advantages of this technique is the generation of a sequence of CT images to represent pulmonary movements during breathing from a conventional CT image.

A brief history shows that the implementation of 4D CT technology is growing each year in the United States since 2002/03 when 4D CT clinical applications were reported. 44% until 2009 and currently, more than 60% of United States cancer centers have been equipped with 4D CT imaging technology. The utilization is still growing 7% per year approximately.^[4] Although there are no accurate statistics, observation shows that the application of 4D CT is not popular enough in Asia rather than either the United States or developed countries. For example, at the

Table 2: Quantitative evaluation of the next-frame prediction

	RMSE	SSIM
Experiment 1	6×10^{-3}	0.951
Experiment 2	11×10^{-3}	0.935
Experiment 3	9×10^{-3}	0.949
Experiment 4	8×10^{-3}	0.943
Experiment 5	5×10^{-3}	0.944
Experiment 6	12×10^{-3}	0.942
Weighted average	9×10^{-3}	0.943

RMSE – Root-mean-squared error; SSIM – Structural similarity index measure

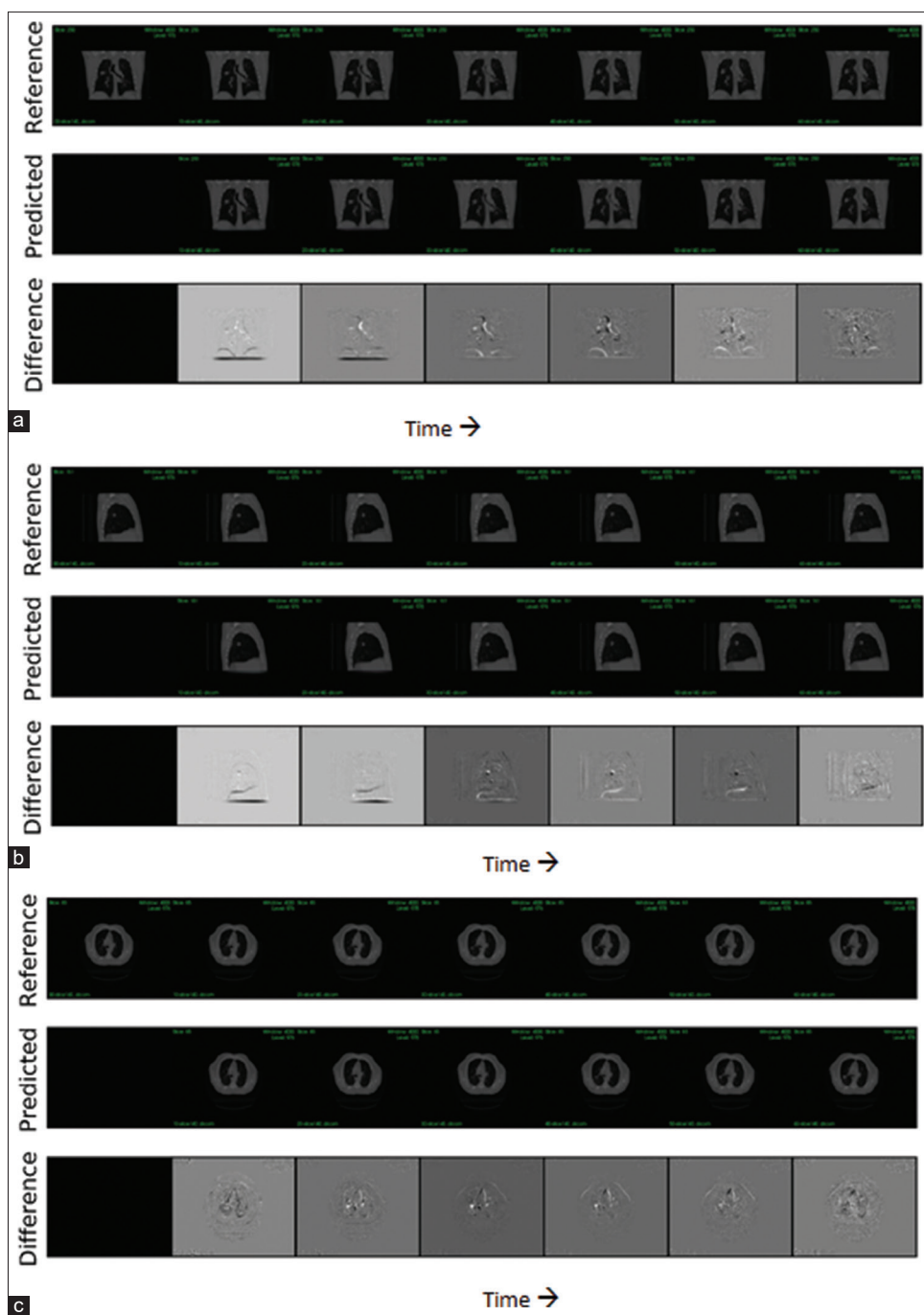


Figure 2: Prediction of images during the respiratory cycle. The first line represents the reference images; the second line represents the predicted images, and the third one contains the difference map in each sub-image (a) Coronal view (b) Sagittal view (c) Axial view

moment, no radiotherapy center has been equipped with 4D CT imaging devices in Iran. Therefore, based on the information provided, the necessity of the application of 4D CT imaging technology should be extremely considered in developing countries.

There are several key differences between 3D and 4D CT applications. For instance, 4D CT images require additional acquisition time up to 2–10 min depending on scanning device configuration, which means more radiation dose will be delivered to patients.^[4] For image artifacts, both 3D and

4D CT suffer from streak, ring, metal artifacts, and blurring. For example, blurring can occur in all types of images due to patient movement, internal organ motion or resolution of scanning during image acquisition. Some additional artifacts including, incomplete structure, duplicated structure, and overlapping structure may degrade the CT images when organ motion affects two adjacent slices, because of iteration frequency of one slice acquisition. Literature shows 90% of 4D CT images have at least one structural artifact, except blurring, which illustrates a warning and it

is worth mentioning that overcoming these artifacts needs time-consuming postprocessing procedure.^[39]

The proposed method can be a satisfactory alternative for better and accurate margin delineation in the radiotherapy centers where do not have access to 4D CT imaging devices. Less imaging radiation delivery and artifact-free images represent also remarkable features of the proposed method.

In spite of the fact that several studies have been presented in recent years about the introduction of respiratory prediction methods based on the use of deep artificial neural networks,^[20-23] and in particular, the use of deep recurrent networks,^[20,21,23] the current study with the generative nature has the advantage of being able to detect tumor motions in all directions. The error of proposed method and the similarity of the reference and generated images significantly reflect the ability of this kind of network to predict pulmonary movements in the lack of proper preliminary data. A larger data set with the more number of images can definitely improve the calculated error. Furthermore, the use of surrogate signals along with images, generated by the proposed method, can make it possible to track the tumor in real-time.

Conclusion

A deep artificial neural network architecture based on convolutional LSTM was introduced. The method predicts the subsequent frame in the video mode. Whereas the 4D CT images represent a sequence of images during the patient's breathing cycle, this architecture was used to predict future frames of 4D CT images from the primary image. This architecture was evaluated on a dataset prepared in all three directions, and the results indicate that the generated images are extremely consistent with the corresponding reference ones. Therefore, to be more reliable in radiotherapy treatment when proper preliminary data are unavailable, this approach can be used to predict tumor position. As a further study, the use of surrogate signals along with the images, produced by this study, can be evaluated for real-time radiotherapy of the tumor.

Acknowledgments

The authors would like to thank the Léon Bérard Cancer Center and CREATIS Laboratory, Lyon, France, for sharing their imaging data to support other studies.

Financial support and sponsorship

None.

Conflicts of interest

There are no conflicts of interest.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A, *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394-424.
2. Korreman SS. Image-guided radiotherapy and motion management in lung cancer. *Br J Radiol* 2015;88:20150100.
3. Keall P. 4-dimensional computed tomography imaging and treatment planning. *Semin Radiat Oncol* 2004;14:81-90.
4. Ehrhardt J. 4D modeling and estimation of respiratory motion for radiation therapy. Lorenz C, editor. Berlin: Springer; 2013.
5. Verma P, Wu H, Langer M, Das I, Sandison G. Survey: Real-time tumor motion prediction for image-guided radiation treatment. *Comput Sci Eng* 2010;13:24-35.
6. Cerviño LI, Jiang Y, Sandhu A, Jiang SB. Tumor motion prediction with the diaphragm as a surrogate: A feasibility study. *Phys Med Biol* 2010;55:N221-9.
7. Riaz N, Shanker P, Wiersma R, Gudmundsson O, Mao W, Widrow B, *et al.* Predicting respiratory tumor motion with multi-dimensional adaptive filters and support vector regression. *Phys Med Biol* 2009;54:5735-48.
8. Ruan D, Fessler JA, Balter JM. Real-time prediction of respiratory motion based on local regression methods. *Phys Med Biol* 2007;52:7137-52.
9. Sharp GC, Jiang SB, Shimizu S, Shirato H. Prediction of respiratory tumour motion for real-time image-guided radiotherapy. *Phys Med Biol* 2004;49:425-40.
10. Vedam SS, Keall PJ, Docef A, Todor DA, Kini VR, Mohan R. Predicting respiratory motion for four-dimensional radiotherapy. *Med Phys* 2004;31:2274-83.
11. Werner R, Ehrhardt J, Schmidt R, Handels H. Patient-specific finite element modeling of respiratory lung motion using 4D CT image data. *Med Phys* 2009;36:1500-11.
12. McClelland JR, Hawkes DJ, Schaeffter T, King AP. Respiratory motion models: A review. *Med Image Anal* 2013;17:19-42.
13. Isaksson M, Jalden J, Murphy MJ. On using an adaptive neural network to predict lung tumor motion during respiration for radiotherapy applications. *Med Phys* 2005;32:3801-9.
14. Ren Q, Nishioka S, Shirato H, Berbeco RI. Adaptive prediction of respiratory motion for motion compensation radiotherapy. *Phys Med Biol* 2007;52:6651-61.
15. Goodband JH, Haas OC, Mills JA. A comparison of neural network approaches for on-line prediction in IGRT. *Med Phys* 2008;35:1113-22.
16. Murphy MJ, Pokhrel D. Optimization of an adaptive neural network to predict breathing. *Med Phys* 2009;36:40-7.
17. Lee SJ, Motai Y, Murphy M. Respiratory motion estimation with hybrid implementation of extended Kalman filter. *IEEE Trans Industr Electron* 2012;59:4421-32.
18. Rostampour N, Jabbari K, Esmaeili M, Mohammadi M, Nabavi S. Markerless respiratory tumor motion prediction using an adaptive neuro-fuzzy approach. *J Med Signals Sens* 2018;8:25-30.
19. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436-44.
20. Kai J, Fujii F, Shiinoki T, editors. Prediction of Lung Tumor Motion Based on Recurrent Neural Network. 2018 IEEE International Conference on Mechatronics and Automation (ICMA). IEEE; 2018.
21. Lin H, Shi C, Wang B, Chan MF, Tang X, Ji W, *et al.* Towards real-time respiratory motion prediction based on long short-term memory neural networks. *Phys Med Biol* 2019;64:085010.
22. Park S, Lee SJ, Weiss E, Motai Y. Intra- and inter-fractional variation prediction of lung tumors using fuzzy deep learning. *IEEE J Transl Eng Health Med* 2016;4:1-12.
23. Wang R, Liang X, Zhu X, Xie Y. A feasibility of respiration prediction based on deep Bi-LSTM for real-time tumor tracking. *IEEE Access* 2018;6:51262-8.

24. Greff K, Srivastava RK, Koutnik J, Steunebrink BR, Schmidhuber J. LSTM: A Search space Odyssey. *IEEE Trans Neural Netw Learn Syst* 2017;28:2222-32.
25. Shirato H, Shimizu S, Kunieda T, Kitamura K, van Herk M, Kagei K, *et al.* Physical aspects of a real-time tumor-tracking system for gated radiotherapy. *Int J Radiat Oncol Biol Phys* 2000;48:1187-95.
26. Finn C, Goodfellow I, Levine S. Unsupervised learning for physical interaction through video prediction. In *Advances in neural information processing systems*; 2016. p. 64-72.
27. Lotter W, Kreiman G, Cox D. Unsupervised Learning of Visual Structure Using Predictive Generative Networks. *CoRR*; 2015.
28. Lotter W, Kreiman G, Cox D. Deep Predictive Coding Networks for Video Prediction and Unsupervised Learning. *International Conference on Learning Representations. (ICLR)*; 2017.
29. Mathieu M, Couprie C, LeCun Y. Deep Multi-Scale Video Prediction Beyond Mean Square Error. *International Conference on Learning Representations (ICLR)*; 2015.
30. Ranzato M, Szelam A, Bruna J, Mathieu M, Collobert R, Chopra S. Video (Language) Modeling: A Baseline for Generative Models of Natural Videos. *CoRR*; 2014.
31. Srivastava N, Mansimov E, Salakhudinov R, editors. *Unsupervised Learning of Video Representations Using Lstms*. *International Conference on Machine Learning*; 2015.
32. Xue T, Wu J, Bouman K, Freeman B, editors. *Visual Dynamics: Probabilistic Future Frame Synthesis Via Cross Convolutional Networks*. *Advances in Neural Information Processing Systems*; 2016.
33. Vandemeulebroucke J, Sarrut D, Clarysse P, editors. *The POPI-Model, A Point-Validated Pixel-Based Breathing Thorax Model*. *XVth International Conference on the use of Computers in Radiation Therapy. (ICCR)*; 2007.
34. Xingjian S, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC, editors. *Convolutional LSTM Network: A Machine Learning Approach for Precipitation Now Casting*. *Advances in Neural Information Processing Systems*; 2015.
35. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. *CoRR*; 2014.
36. Keras CF; 2016. Available from: <http://keras.io/>. [Last accessed on 2019 Apr 20].
37. Kohavi R, editor *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*. Montreal, Canada: IJCAI; 1995.
38. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process* 2004;13:600-12.
39. Yamamoto T, Langner U, Loo BW Jr., Shen J, Keall PJ. Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients. *Int J Radiat Oncol Biol Phys* 2008;72:1250-8.

BIOGRAPHIES



Shahabedin Nabavi received his M.Sc. degree in Software Engineering from Shahid Beheshti University (National University), Tehran, Iran. Currently, he is a PhD Candidate at the Faculty of Computer Science and Engineering, Shahid Beheshti University (National University), Tehran, Iran. His main research interests are Medical

Image Analysis and Deep Learning.

Email: s_nabavi@sbu.ac.ir



Monireh Abdoos has obtained his PhD degree in Computer Engineering from Iran University of Science and Technology. She is an Assistant Professor at the Faculty of Computer Science and Engineering, Shahid Beheshti University (National University), Tehran, Iran. Her main research interests are Multi-Agent Systems, Machine

Learning, and Intelligent Transportation Systems.

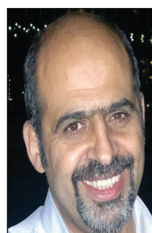
Email: m_abdoos@sbu.ac.ir



Mohsen Ebrahimi Moghaddam received his PhD and M.Sc. degree in Software Engineering from Sharif University of Technology, Tehran, Iran. Currently, he is an associate professor at the Faculty of Computer Science and Engineering, Shahid Beheshti University (National University), Tehran, Iran. His main research interests are

Image Processing, Machine Vision, Data Structures, and Algorithm design.

Email: m_moghdam@sbu.ac.ir



Mohammad Mohammadi has obtained his M.S. degree in Medical Physics from Tehran University of Medical Sciences, Tehran, Iran and his PhD degree in Medical Physics from the University of Adelaide, Adelaide, Australia. He is a medical physicist at the Royal Adelaide Hospital and senior lecturer at the School of Physical Sciences, The University of Adelaide, Adelaide, Australia. His main research interests are Radiation Dosimetry, Radiotherapy Physics, and Medical Imaging.

Email: mohammad.mohammadi@sa.gov.au