A review of artificial intelligence applications for motion tracking in radiotherapy

Running title: AI applied to motion tracking

Adam Mylonas[1,2,†], Jeremy Booth[3,4], and Doan Trang Nguyen[1,2,3]

1 ACRF Image X Institute, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia

2 School of Biomedical Engineering, University of Technology Sydney, Sydney, New South Wales, Australia

3 Northern Sydney Cancer Centre, Royal North Shore Hospital, St Leonards, New South Wales, Australia

4 Institute of Medical Physics, School of Physics, The University of Sydney, Sydney, New South Wales, Australia

† Author to whom correspondence should be addressed. E-mail address: adam.mylonas@sydney.edu.au

Funding: D T Nguyen is funded by an Early Career Research Fellowship from the Australian National Health and Medical Research Council (NHMRC) and the Cancer Institute of New South Wales:

- Cancer Institute NSW Early Career Fellowship 2018-ECF007
- NHMRC Council Early Career Fellowship GNT1138807

Conflict of interest disclosure: Authors A Mylonas and D T Nguyen are inventors on two patent applications related to AI in motion tracking.

Abstract

During radiotherapy, the organs and tumour move as a result of the dynamic nature of the body; this is known as intrafraction motion. Intrafraction motion can result in tumour underdose and healthy tissue overdose, thereby reducing the effectiveness of the treatment while increasing toxicity to the patients. There is a growing appreciation of intrafraction target motion management by the radiation oncology community. Real-time image guided radiation therapy (IGRT) can track the target and account for the motion, improving the radiation dose to the tumour and reducing the dose to healthy tissue. Recently, artificial intelligence (AI) based approaches have been applied to motion management and have shown great potential. In this review, four main categories of motion management using AI are summarised: marker-based tracking, markerless tracking, full anatomy monitoring, and motion prediction. Marker-based and markerless tracking approaches focus on tracking the individual target throughout the treatment. Full anatomy algorithms monitor for intrafraction changes in the full anatomy within the field of view. Motion prediction algorithms can be used to account for the latencies due to the time for the system to localise, process and act.

Key words: radiation oncology, artificial intelligence, motion tracking, deep learning, machine learning

Introduction

Since the early days, radiotherapy has been driven by new technological developments. Image guided radiation therapy (IGRT) improves the treatment accuracy and enables the success of highly conformal treatment delivery techniques such as Intensity Modulated Arc Radiation Therapy and Volumetric Modulated Arc Therapy. With hypofractionated and ultrahypofractionated treatments becoming more mainstream, particularly with Stereotactic Body Radiation Therapy (SBRT) becoming the standard of care for non-small cell lung cancer and prostate cancer, ASTRO recommends not only daily IGRT but also real-time IGRT as the target positioning method for patient treatments (1, 2). This is due to a growing appreciation of intrafraction target motion management by the radiation oncology community.

Although the prostate was assumed to move negligibly, greater understanding of prostate motion led to changes in intrafraction motion management. This was a result of the introduction of kilovoltage (kV) and ultrasound imaging during treatment in the early 2000's driven by the highly sculpted dynamic MLC deliveries. The range and distribution of intrafraction prostate motion observed in patients has been reported in several studies (3-5). An analysis of prostate intrafraction translational motion in 427 patients found that displacement larger than 2 mm occurred in 66% of fractions while 28% of fractions had exceeded 3 mm displacement (6). Most of the prostate motion is reported to be in the anterior-posterior and superior-inferior directions (4). In some cases, intrafraction prostate motion can be up to 15 mm, which typically exceeds the Planning Target Volume (PTV) expansion for radiotherapy treatment of the prostate. In addition to translational motion, the prostate motion also includes rotation. Intrafraction rotation motion is more commonly seen in the Left-Right axis, with an average of 2.5º (±2.3º) reported (5). However, in extreme cases, the rotational motion can reach beyond 10º (7, 8). The physiological origins of prostate intrafraction motion are

frequently bladder and rectal volume changes due to bladder filling, rectal gas passing or rectal filling. These same processes also result in intrafraction motion and deformation seen during other pelvic cancer treatments, such as cervical cancer. Using cine MRI images, motion of the cervical Clinical Target Volume (CTV) was reported to be as high as 10.6 mm (9).

Respiration-induced motion is faster and has greater variability than the random motion of the prostate and has been shown to affect treatment accuracy. The motion of the target within the thorax and abdomen is thus needed to be monitored in real-time with a high frequency. For thoracic and upper abdominal targets, intrafraction motion (translation, rotation and deformation) of several centimetres has been observed in the liver (10, 11), lung (12, 13), and pancreas (14). For targets located close to the heart, cardiac motion is also a concern for motion management for two reasons. Firstly, the target will exhibit respiratory and cardiac induced motion. Secondly, the heart and the associated large vessels are typically dose limiting organs due to high radiation sensitivity. Cardiac motion of the heart and its substructures were reported to be between 7 mm to 15 mm in a MRI study (15). However, the cardiac component of the motion was found to be within 4-5 mm magnitude (16). Althought inclusion of an internal target volume during planning to account for motion is a widely used method, it has been shown to be less optimal than monitoring and accounting for motion in real-time both lung and liver SBRT treatments (17, 18).

The evidence for the dosimetric impact of intrafraction motion is mounting. A study on liver cancer treatment with radiation therapy showed that the dose to the liver could be reduced by as much as 20% due to liver motion (19). Another study on prostate cancer patients showed that prostate motion could result in 19% tumour underdose (20). For lung cancer patients, when motion is compensated for, not only

will the tumour receive the full dosage, but the surrounding organs benefit with a reported reduction in mean lung dose by 17% and in heart dose by 56% (18).

A number of different intrafraction real-time guidance methods have been used during radiotherapy treatments. The first real-time tracking radiotherapy system, pioneered by Shirato et al. (21) uses four orthogonal room-mounted X-ray systems to track the prostate position based on segmented positions of implanted markers (21-26). Other commercial systems including CyberKnife (27) (Accuray, CA, USA) and Exactrac (28) (BrainLab, Munich, Germany) use two room-mounted X-ray imagers synchronously to monitor the target motion during treatment. Due to gantry occlusion, an image pair can only be taken at cardinal angles. This limitation is not seen in the four X-ray system developed by Shirato et al. (21).

Non-radiographic methods for target motion monitoring have also been developed in the past two decades. Calypso (Varian, Palo Alto, CA) (4) and RayPilot (Micropos, Gothenburg, Sweden) (29) utilise implanted markers which also act as electromagnetic transponders, transmitting positional signals to an external receiver. Emerging real-time guidance technologies include ultrasonography (30) and integrated magnetic resonance imaging (MRI)-radiation therapy systems (31, 32). While these technologies are becoming more common, this review will focus primarily on kV IGRT.

Up to date, we have witnessed artificial intelligence (AI) lending its power to the ideal goal of monitoring the target during radiotherapy using a standard linear accelerator. Machine learning (ML) and deep learning (DL) are key AI technologies that are being increasingly used in radiotherapy applications. These technologies excel at extracting features from training data and making predictions on new unseen data. Most of the published research to date that use these technologies centre around localising the target locations on 2D kV images such that the 3D position of the target can be reconstructed. The target in this context could be a surrogate such as arbitrary-

124 shaped fiducial markers (33), the gross tumour volume (34), bony anatomy surrounding
125 the target (35) or the soft tissue target with very low radiological contrast (36, 37).

126 AI-related motion tracking publications have been predominately produced over
127 the past two decades. In recent years, there has been a surge of publications related to
128 this field (Figure 1). Almost half of all articles published over the past two decades have
129 been published in just the past three years (Figure 1) and now focus primarily on DL.
130 Therefore, it is important to evaluate the shift in AI-related motion tracking research
131 from ML to DL over the past two decades. Furthermore, a review of this field will
132 summarise initial progress made with AI techniques and provide insight for further
133 research. In this review, we aim to examine the state of the art of AI techniques that
134 have been applied to motion monitoring, starting from marker-based approaches to full
135 anatomy reconstruction. Marker-based approaches are typically used for prostate real-
136 time IGRT since the markers act as a high contrast surrogate for the target. However,
137 the development of markerless tracking technologies is beneficial to the patient and the
138 healthcare system by obviating the need for implanted markers. Full anatomy
139 algorithms can monitor for intrafraction changes in the full anatomy within the field of
140 view. Another important technology is motion prediction algorithms which can be used
141 to account for the latencies due to the time for the system to localise, to process and to
142 act. AI algorithms have helped to advance these key technologies for motion tracking in
143 radiotherapy.

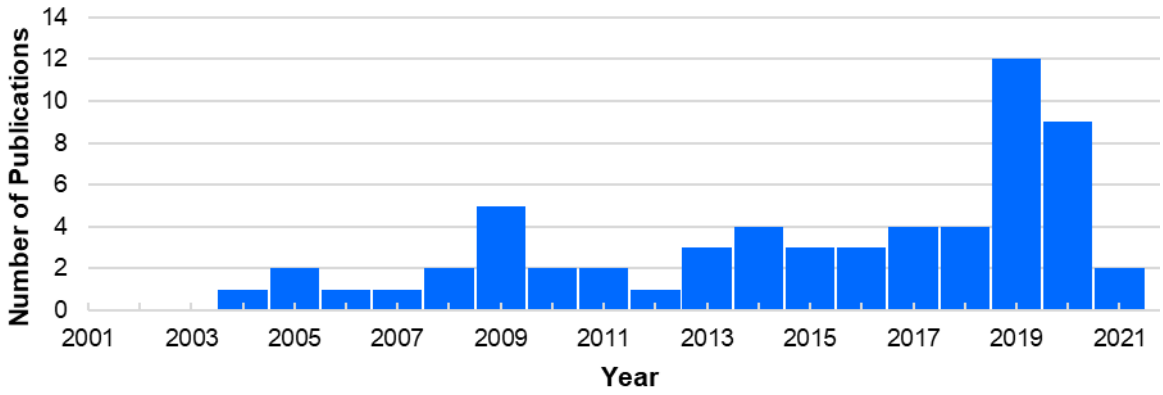

Figure 1. Number of publications over the past two decades for AI in intrafraction motion tracking.

AI technologies

The research discipline of AI aims to develop intelligent systems so advanced that they can solve problems to the same ability or even better than humans. ML is a branch of AI in which systems can automatically learn from data and improve from experience without being programmed to do so. There are a large variety of statistical algorithms used in ML such as artificial neural networks (ANN), support vector machines (SVM), decision trees, decision forests, logistic regression, and Bayesian classifiers (38). DL is a subset of ML which solves problems using neural networks which have a structure similar to the human neural system. The name DL comes from the structure (or architecture) of the networks. These networks have multiple layers in which different operations are applied to train the network, unlike traditional ML algorithms which only had a small number of layers.

Convolutional neural networks (CNN) are one of the main DL approaches with a multiple layer architecture. Traditional CNN architectures are typically used for image classification problems (Figure 2A). Traditional CNNs are designed to receive an input image of a specified size and output a single label classification. CNNs can be adapted

for image segmentation using a common approach known as regions with CNN (R-CNN). R-CNNs first generate region proposals that potentially contain the object of interest and then performs classification on the regions using a CNN. The traditional CNN architecture can be modified to perform detection and semantic segmentation (Figure 2B). Semantic segmentation is the technique of partitioning an image into labelled segments by classifying each pixel. A common architecture used for semantic segmentation is the encoder-decoder network (Figure 2A) which allows for an arbitrarily shaped input and output of the same size.

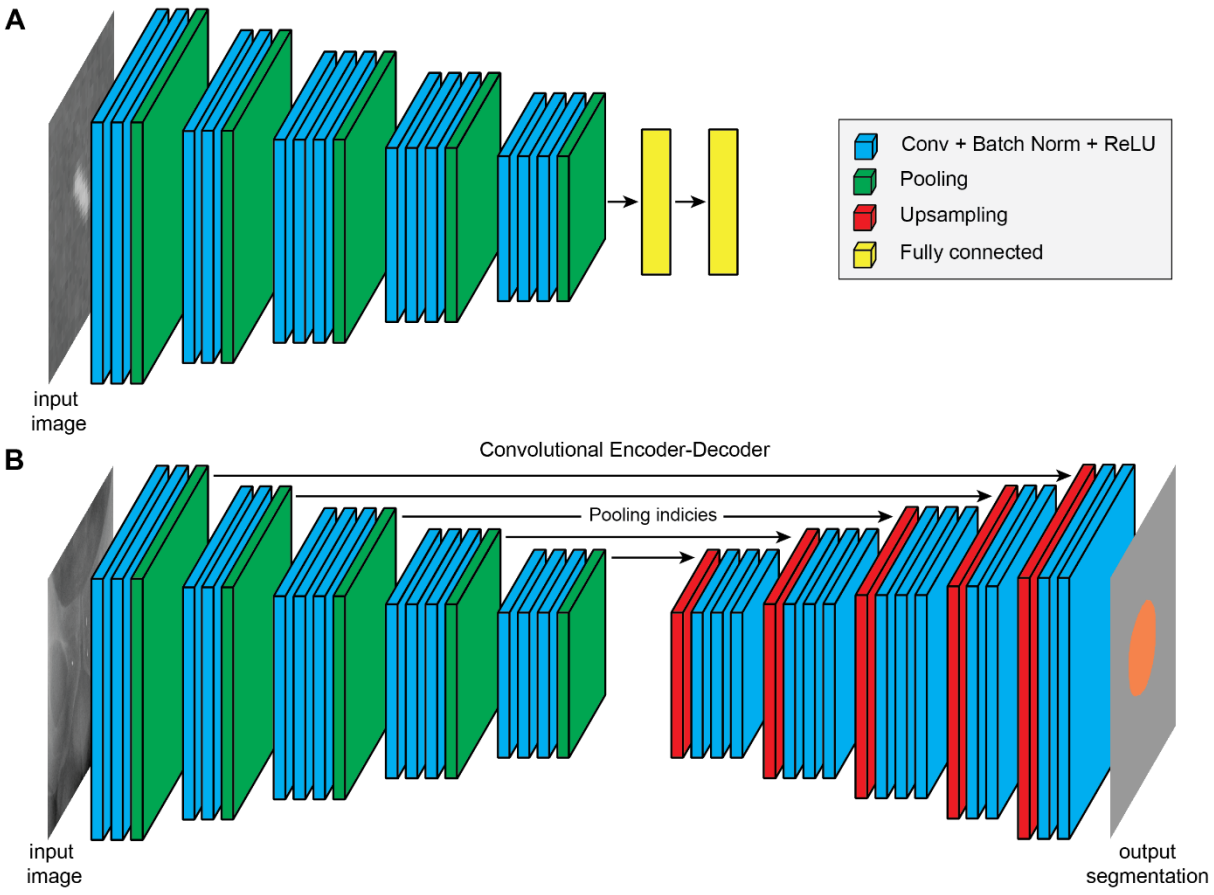

Figure 2. The basic architecture of (A) a CNN for image classification and (B) an encoder-decoder network for semantic segmentation. The network layers include convolutions (Conv), batch normalisation (Batch Norm), rectified linear unit (ReLU), pooling, upsampling, and fully connected layers.

Method

The authors conducted a topical literature review using the online databases PubMed and Web of Science from February to May 2021. The search criteria initially restricted the articles to the field of radiotherapy using: (radiation therapy OR radiotherapy). These search results were refined towards AI using: (machine learning OR deep learning OR neural networks OR artificial intelligence). Finally, the articles were further refined to tracking applications using: (fiducial markers OR markerless), or refined to motion prediction applications using: (motion AND prediction). An additional search was completed to review applications of AI to motion management in MRI-guided radiotherapy. The search criteria initially restricted the articles to the field of MRI-guided radiotherapy using: (mr OR mri) AND linac. These search results were refined towards AI using the aforementioned AI search terms. In addition, the reference lists of selected articles were hand searched for other relevant articles. Articles returned in the search results that were not related to intrafraction motion tracking were excluded from the review.

Marker-based approaches

IGRT for some types of cancer typically relies on markers to determine the target location. For motion tracking in kV images, markers are implanted into the tumour or surrounding tissue to act as a surrogate for the target due to a lack of contrast in the intrafraction images (Figure 3). A range of AI approaches for marker tracking have been demonstrated to have high accuracy (Table 1).

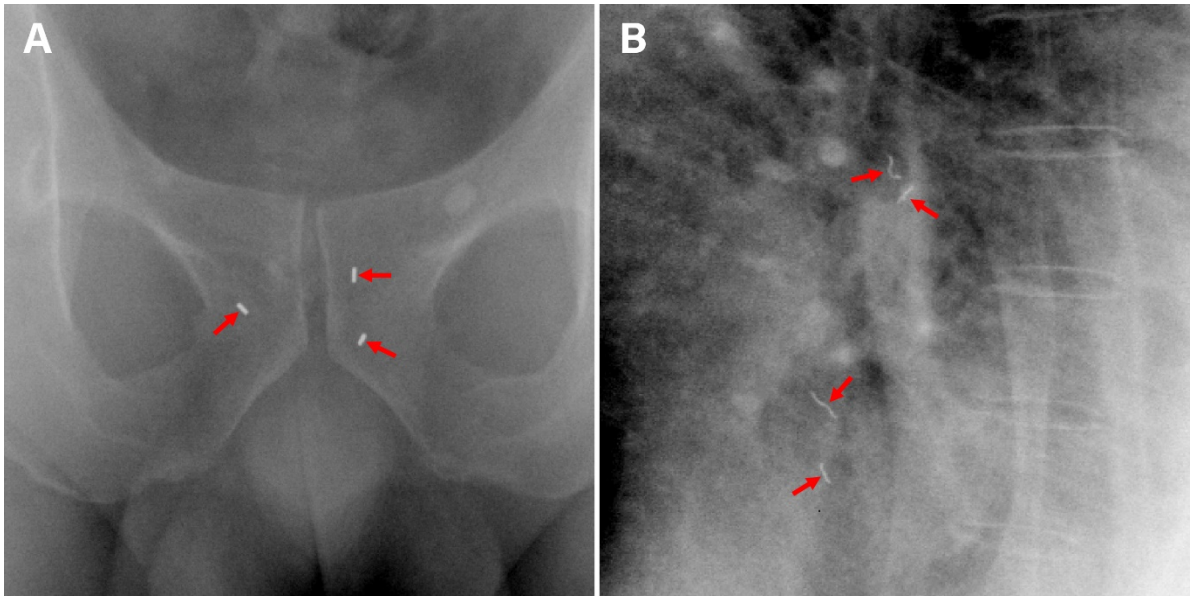

Figure 3. Example kV projections of (a) regular shaped gold fiducial markers implanted in the prostate and (b) arbitrarily shaped coiled markers implanted in the lung.

Many methods for segmenting markers use traditional approaches such as template matching (39-43). Template matching approaches require prior information of the marker. Typically for regular shaped markers (Fig 1A; e.g. cylindrical gold fiducials), the templates are constructed using known marker properties (40, 41). Other types of markers such as coiled markers (Fig 1B; e.g. Visicoil™) may deform into arbitrary shapes upon implantation resulting in unknown marker properties. In order to develop a template, an additional learning period is required to determine the marker properties (39, 42, 43). Cone-beam CT (CBCT) projections are commonly used as learning data for the template generation (41). This learning period has the drawback of subjecting the patient to additional radiation due to the extra imaging dose. DL methods have been successfully used for the automatic detection of markers in x-ray and MRI images to overcome these issues (33, 44, 45).

Table 1. Summary of the methods to track implanted markers.

Study	Imaging modality	Marker type	Method	Error	Other metrics

Lin et al. 2012 (45)	MV	Regular	ML	RMSE error: 0.5 mm	-
Mylonas et al. 2019 (33)	kV	Regular	CNN	Mean (±SD) error: −0.8±1.3 pixels and 0.3±1.2 pixels in the x- and y-directions respectively.	Sensitivity (recall): 99.4% Specificity: 99.3%
Mylonas et al. 2019 (33)	kV	Arbitrary	CNN	Mean (±SD) error: −0.9±2.8 pixels and 0.2±2.2 pixels in the x- and y-directions respectively.	Sensitivity (recall): 98.5% Specificity: 99.6%
Liang et al. 2020 (44)	X-ray	Regular	U-Net	Mean (±SD) error: 0.3±0.5 pixels.	Sensitivity (recall): 95.6% Precision: 98.6%

Tracking fiducial markers using the megavoltage electronic portal imaging device (MV-EPID) is ideal as it eliminates the need for additional imaging equipment and provides target information in the view of the treatment beam. However, accurate tracking is challenging for VMAT treatments due to low contrast and MLC leaves occluding markers. Lin et al. (45) detected the markers using discriminant analysis for the first frame and mean-shift feature space analysis for subsequent frames. To reduce the computation time, the authors exploited the fact that the markers moved slightly between each frame. An average root-mean-square error (RMSE) of 0.5 mm was achieved using this approach.

Mylonas et al. (33) implemented a framework that uses a CNN to automatically detect regular and arbitrarily shaped markers in intrafraction fluoroscopic images. The CNN had three convolutional layers with two pooling layers and one fully connected layer. The CNN was used to classify whether sub-images of the kV projections contained a marker or background. High classification performances were achieved for both types of markers. For the regular shaped markers implanted in the prostate, a sensitivity and specificity of 99.4% and 99.3%, respectively, were achieved. These results improve upon a marker detection rate of 94.8% for a template matching approach reported by Fledelius et al. (41). For the arbitrarily shaped markers implanted in the lung, a sensitivity and specificity of 98.5% and 99.6%, respectively, were achieved. This improves upon the traditional method developed by Bertholet et al. (39)

which reported a sensitivity and specificity of 67% and 94%, respectively. While the CNN approach improves upon the current standard, it should be noted that the systems were tested on different datasets. The CNN was incorporated into a tracking system to segment multiple markers for intrafraction monitoring. The computation time was adequate for real-time applications with a time of 36 ms on a CPU and 9 ms on a GPU. Submillimeter accuracy was achieved for both regular and arbitrarily shaped marker tracking.

A study by Liang et al. (44) presented an automated framework that was used to evaluate intrafraction motion in x-ray images acquired by CyberKnife. The framework includes a fully convolutional network (FCN)-based fiducial marker detection module. The authors utilised the U-Net architecture to perform semantic segmentation of the markers in full-size x-ray images. The U-Net architecture is an implementation of the encoder-decoder network architecture shown in Figure 2B. The network has a contracting path to obtain the feature information and a symmetrical expanding path which increases the resolution of the encoder feature maps. The method achieved a precision and recall of 98.6% and 95.6% respectively, and the mean (±SD) centroid error was 0.3±0.5 pixels. Their proposed method was used to retrospectively analyse the intrafraction motion of markers implanted in the liver. This method could be further developed for use in real-time tumour tracking systems. The computation time of this technique would need to be examined before being implemented in a real-time tracking system.

Markerless-based approaches

The development of markerless IGRT would be beneficial to patients and the healthcare system by eliminating the need for the surgical implantation of markers. A

markerless approach would avoid extra appointments, costs, discomfort, and potential complications associated with marker implantation. Percutaneous marker implantation can have complications such as haemorrhaging (46, 47). Furthermore, some patients are not candidates for markers due to contraindications. This prevents some patients from being treated with real-time IGRT techniques. Marker migration is another potential issue that can result in tracking errors (48) and markers produce metal artefacts in CT images that may result in treatment-planning errors (49). Therefore, the ability to track tumours without markers would be beneficial to cancer patients and enable wide access to real-time IGRT. It is difficult to develop highly accurate markerless tracking methods for organs such as the prostate or pancreas due to the low contrast in intrafraction images (Figure 4). However, AI approaches have shown great potential for markerless tracking (Table 2 and 3).

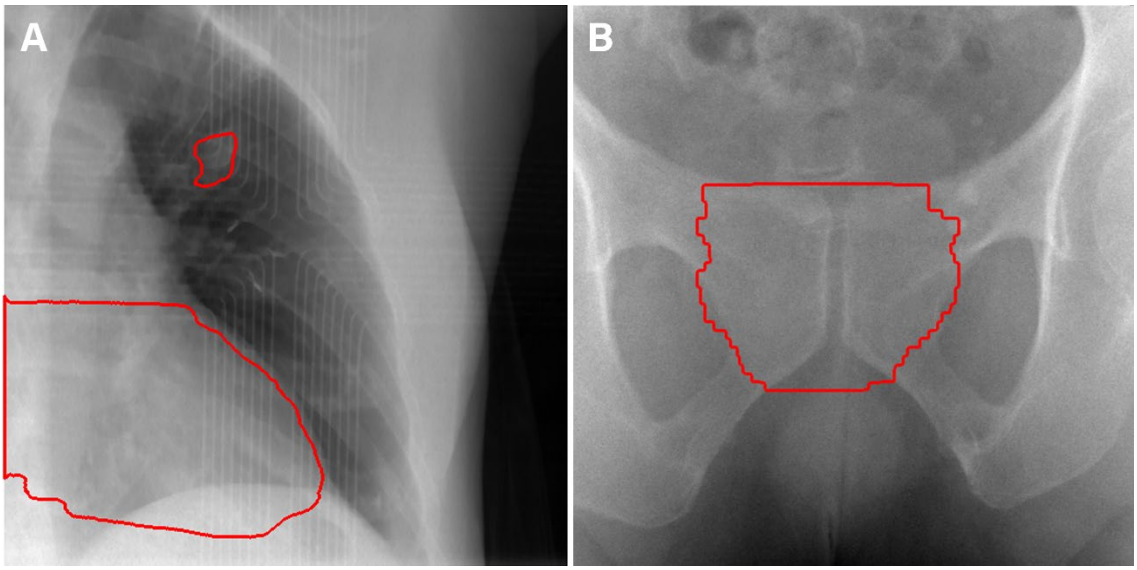

Figure 4. Example segmentations of the (a) heart and lung tumour and (b) the prostate.

277 **Table 2.** Summary of the markerless-based approaches for fluoroscopic guided
278 radiotherapy.

Study	Imaging modality	Method	Site	Results
Tong et al. 2009 (50)	X-ray	ML	Lung	Mean error (95% confidence level): 2.1 (4.6) pixels (pixel size is about 0.5 mm)
Terunuma et al. 2018 (51)	Synthetic X-ray	Encoder-decoder CNN	Lung	Mean error: approximately 1 mm.
Edmunds et al. 2019 (52)	X-ray	Mask R-CNN	Diaphragm	Mean error: 4.4 mm
Hirai et al. 2019 (53)	X-ray	DNN	Liver and lung	Mean (±SD) accuracy: 1.6±0.7 mm
Mori et al. 2019 (54)	X-ray	ML	Chest	Accuracy: <0.5 mm
Zhao et al. 2019 (37)	Synthetic X-ray	Faster R-CNN	Pancreas	MAD: less than 2.6 mm across all directions
Zhao et al. 2019 (36)	Synthetic X-ray	Faster R-CNN	Prostate	MAD (±SD): 1.6±0.4 mm, 1.6±0.4 mm, and 1.7±0.4 mm in anterior-posterior, lateral, and oblique directions, respectively.
Roggen et al. 2020 (35)	X-ray	Faster R-CNN	Vertebrae	Sensitivity of 1.5 mm with a median below 0.5 mm for single vertebrae.
Sakata et al. 2020 (55)	X-ray	ML	Lung	Mean (±SD) error, Euclidean distance: 1.0±0.3 mm.
Takahashi et al. 2020 (56)	X-ray	FCN	Lung	Mean (±SD) error: mean (±SD): 0.5±0.2 mm and 0.3±0.2 mm for 2 cm and 3 cm tumours respectively.

279

280 Several ML approaches have been employed for markerless lung tumour
281 tracking. A novel method developed by Tong et al. (50) uses principal component
282 analysis (PCA) to determine the motion patterns of the surrogate anatomical features.
283 The tumour position is predicted based off the motion patterns through regression
284 using ML approaches. The tracking system was tested on the fluoroscopic images for
285 ten lung cancer patients. Compared to other methods, ANN regression was shown to
286 perform the best with a mean tracking error of 2.1 pixels and mean maximum error
287 (95% confidence level) of 4.6 pixels where one pixel is approximately 0.5 mm.

Mori et al. (54, 57) and Sakata et al. (55) used a ML algorithm to produce a tumour likelihood map using fluoroscopic images. A binary classifier optimized by the extremely randomized tree method was trained to classify images as positive (including tumour) and negative (not including tumour). The classifier was trained using digital reconstructed radiographs (DRR) generated from the 4DCT. During the treatment, sliding window classification was used to generate the tumour likelihood map. The tumour position was predicted using the tumour likelihood map and prior probability tumour positions. The system achieved an accuracy of <0.5 mm when tested on a moving chest phantom (54). When tested on images from eight patients, the mean (±SD) tumour position error was 1.0±0.3 mm in 3D space based off paired x-ray images (55). Furthermore, the ML algorithm was shown to be accurate when trained with DRRs generated from a simulated 4DCT (57).

Zhao et al. implemented a DL approach for markerless tracking for prostate (36) and pancreatic (37) cancer. The faster R-CNN model used in these studies consisted of two networks that share information: a region-proposed network (RPN) and a R-CNN. The authors used a RPN to generate proposals for the R-CNN to reduce the computation time and enable real-time target detection. The feature extractor was the VGG16 ConvNet. A model was trained for three different directions: anteroposterior, left-right, and an oblique. The network input can be either a digitally reconstructed radiography (DRR) or x-ray projection and the output is a bounding box of the PTV. The models were evaluated on simulated kV projection images. For the pancreatic cancer model, the median absolute deviation (MAD) between the predicted and actual position was less than 2.6 mm for all cases and directions (37). For the prostate cancer model, the MAD (±SD) between the predicted and actual positions were 1.6±0.4 mm, 1.6±0.4 mm, and 1.7±0.4 mm in anterior-posterior, lateral, and oblique directions, respectively (36). The prostate cancer model was additionally evaluated on kV images

of one prostate cancer patient. The predicted prostate position was consistent with the position of the markers for the two kV images shown (36). These results demonstrated the feasibility of markerless prostate and pancreas tracking for IGRT. For clinical implementation, it would be beneficial to expand this method for tracking through all angles of a radiotherapy treatment arc. For both methods, the model produced the prediction in less than 200 ms which makes this method feasible for real-time monitoring.

A DL method developed by Hirai et al. (53) successfully estimated lung and liver tumour positions for markerless tracking. A deep neural network (DNN) was trained to produce a target probability map (TPM) that estimated the tumour position. The DNN was based on an encoder and decoder network architecture (Figure 2B). For training, DRR images were cropped to produce sub-images of the target and surrounding anatomy. The network takes the pattern of sub-images to produce the TPM. During tracking, the fluoroscopic images are cropped at the same positions used during the training process on the DRR images and the target positions were calculated based on the TPMs. The framework was implemented on both lung and liver cases. The accuracy was calculated using the Euclidian distance in the 3D space between the calculated and reference tumour position. An overall mean accuracy (±SD) of 1.6±0.7 mm was achieved for all patients. The framework was more accurate on liver cases achieving a mean accuracy (±SD) of 1.4±0.8 mm in comparison to 1.9±0.7 mm for lung cases. This method was shown to be feasible for real-time since the computation time was less than 40 ms.

A study by Roggen et al. (35) evaluated a DL model to detect vertebra motion in 2D x-ray projection images. The authors used the Mask R-CNN framework based on ResNet101 to perform image segmentation. First, the framework detects several regions of interest and then a Faster R-CNN is used to process each region to

340 determine the actual presence of the objects of interest. The framework was trained
341 and tested using the CBCT projections acquired using an on-board kV imaging system.
342 The ground truth was the forward projection of the 3D contour of the vertebrae. The
343 network takes a 2D kV projection image as input and outputs a segmentation mask and
344 bounding box of each vertebrae. This approach was validated on images from one
345 patient and a patient-like full-body phantom. When validated on the patient images,
346 positional changes in the range of 1.5 mm could be detected, with a median below 0.5
347 mm. Sub-millimetre accuracy was achieved when the detection of several vertebrae
348 was combined. Additionally, this method can detect spinal rotations larger than 1°
349 based on the detection of all vertebrae in the image. Further development is still
350 required to achieve the same level of accuracy and provided additional advantages
351 compared to current systems such as the Cyberknife system. For spinal tracking, the
352 Cyberknife system has an accuracy of <0.2 mm for translational and <0.3° for rotational
353 directions (58).

354 A R-CNN was also used by Edmunds et al. (52) to segment the diaphragm for
355 lung tumour tracking with a mean diaphragm apex error of 4.4 mm. Terunuma et al.
356 (51) used a encoder-decoder CNN architecture for lung tumour segmentation in DRRs.
357 When tested on DRRs, the tracking error was approximately 1 mm with a processing
358 time of 25ms/frame for contouring and tracking. A DL approach for segmenting lung
359 tumours was also developed by Takahashi et al. (56). Their approach uses a FCN
360 trained on patient specific DRRs. The system was tested in a phantom study with
361 tumours of 1, 2 and 3 cm in size. Sub-millimetre accuracy was achieved with a mean
362 (±SD) error of 0.5±0.2 mm and 0.3±0.2 mm for 2cm and 3 cm tumours respectively.
363 However, the mean error increased to over 1mm for a smaller 1 cm tumour. The
364 tracking system had an approximate processing time of 33 ms/frame.

365 **Table 3.** Summary of the markerless-based approaches for MRI-guided radiotherapy.

Study	Imaging modality	Method	Site	Results
Cervino et al. 2011 (59)	cine MRI	ML	Lung	Mean tracking error of 1.5 mm
Yun et al. 2015 (60, 61)	cine MRI	PCNN	Lung	Centroid tracking accuracy of 1.0-1.4 mm
Bourque et al 2017 (62)	cine MRI	ML	Lung	RMSE: 1.3±0.5 mm
Fast et al. 2017 (63)	cine MRI	PCNN	Lung	Median Dice similarity index: ≥ 0.91
Friedrich et al. 2021 (64)	cine MRI	U-net	Liver	Average Dice coefficient ≥ 83% Mean contour distance ≤ 0.5 pixel

Currently, MR images can only be acquired in 2D modality during treatment delivery (65). Therefore, the AI-approaches for motion management during MRI-guided radiotherapy summarised in Table 3 are designed for 2D images (59-64). A study by Cervino et al. evaluated two approaches for tracking lung tumours in cine MRI images (59). Cervino et al. compared an ANN model and a template matching approach combined with surrogate tracking using the diaphragm. It was observed that the template matching approach detected the tumour location more accurately than the ANN. Yun et al. (60) and Fast et al. (63) used a pulse-coupled neural network (PCNN) to track lung tumours in cine MRI images. Yun et al. reported a centroid tracking accuracy of 1.0-1.4 mm (60). However, the PCNN was outperformed by multi-template matching and deformable image registration in the study by Fast et al (63). A recent study compared B-spline deformable image registration and U-net for segmentation of liver tumours in cine MRI images (64). The U-net was found to outperform B-spline and achieved a mean contour distance of less than 0.5 pixels. A ML approach that incorporated motion prediction for tracking lung tumours was proposed by Bourque et al (62). A RMSE of 1.3±0.5 mm and 2.0±0.8 mm with and without prediction, respectively, shows the benefit of including motion prediction into the tracking algorithm.

385

386 **Algorithms for monitoring changes in the full anatomy**

387 Another approach for motion tracking is to monitor for shifts in the full anatomy or
388 features of the surrounding anatomy. It can be useful to monitor for shifts in the full
389 anatomy during the treatment to ensure that the target volume and surrounding organs
390 at risk remain in position. If there are significant shifts, gating can be performed and the
391 patient can be repositioned. There are a variety of AI solutions for gating based on the
392 full anatomy (34, 66, 67).

393 An approach used in several studies is to use a binary classifier for online
394 treatment target verification by determining when to turn the beam ON or OFF (34, 67,
395 68). Tang et al. (67) proposed using an ANN binary classifier to classify the lung
396 tumour as inside or outside of the beam aperture in cine EPID images. The ANN was
397 trained on DRRs with simulated tumour shifts to represent cine EPID images. The MLC
398 contour was superimposed on the DRRs at different locations to simulate different
399 tumour positions. The CTV defined by the physician was projected onto the DRRs. The
400 percentage of the CTV in the beam aperture was calculated for each image. A
401 percentage threshold was set to divide the images into the two classes: inside or
402 outside of the beam aperture. The ANN was validated retrospectively on five lung
403 patients. The ANN had a classification accuracy of 98.0%, recall of 97.6%, and
404 precision of 99.7%. The proposed method could be implemented for real-time gating
405 during radiotherapy when the ANN determines the tumour has moved outside of the
406 beam aperture.

407 Similarly, Lin et al. (34) developed a solution for gated radiotherapy using a
408 binary classifier and evaluated the performance of several types of ML algorithms. The
409 authors used a combination of dimensionality reduction technique and ML algorithms.

The binary classifier was developed to classify images as beam ON or beam OFF based on the gating window size set during treatment planning. The method was validated on fluoroscopic images of nine lung cancer patients. The ANN combined with PCA was determined to be the best candidate for real-time gated radiotherapy. The ANN with PCA method had a mean (±SD) classification accuracy, recall, and target coverage of 96.3±1.6%, 89.9±7.6%, and 97.8±1.7% respectively. Cui et al. (68) investigated the use of a SVM for the classification problem. However, the ANN approach was shown to be more accurate than SVM in the study completed by Lin et al. (34).

An approach for volumetric reconstruction using DL developed Shen et al. (66) could be used for image-guided radiotherapy. An encoder-decoder model was trained to generate 3D volumetric CTs from 2D x-ray projections. The model can be trained on the planning CT to make the approach patient specific. The training dataset of DRR-CT pairs was generated by producing DRR images based on the CT. The authors investigated using 1, 2, 5, and 10 projection views as input to obtain the CT. It was found that the underlying semantic representations of the 3D scene were similar when different 2D projection views are utilised. The model was validated on both abdominal and lung CTs. The mean absolute error between the predicted and ground truth images for single-view reconstruction was 0.018 and 0.025 for abdominal and lung CTs respectively. This single-view imaging approach is a potential solution to image-guided interventional procedures. Another study proposed a method for producing a volumetric image using a single x-ray projection using PCA (69).

Similarly for MRI-guided radiotherapy, DL methods have the potential to reduce the computation time for image reconstruction from 2D cine MRI images. Real-time reconstruction would allow for volumetric segmentation of the target and treatment adaption. Terpstra et al. compared four conventional and DL approaches for image

reconstruction and to estimate the motion of abdominal cancer (70). The authors used a combination of methods to achieve a RMSE up to 1 mm, with a processing time of 60 ms.

Ultrasound-guided radiotherapy can be challenging due to the major changes in the image content due to motion of the anatomy and the position of the ultrasound probe. Zhu et al. (71) used a pretrained CNN for image registration of 3D ultrasound prostate images. The registration errors using the CNN were smaller than 5 mm in 81% of the cases which was an improvement over a manual approach that had errors smaller than 5 mm in only 61% of the cases.

Prediction of tumour motion

There will be a latency in any system from the time to localise, to process and to act. For any action to be considered real-time, then the system must predict (estimate) localisation at a time in the future equal to the system latency. Common system latencies for say an MLC tracking system are 100-400 ms. Prediction algorithms are not new, and many have been applied to the motion of targets in radiotherapy using statistical approaches. More recently, DL approaches have been applied to this problem.

Several studies have used ML algorithms for predicting tumour motion based on past motion (72-89), including in MRI-guided radiotherapy (90-92) and ultrasound guided radiotherapy (93). A comparison study of ML algorithms was made by Sharp et al. and showed that most ML algorithms have a lower localisation error compared to no prediction (94). Murphy and Dieterich (95) evaluated a linear adaptive filter and an adaptive nonlinear neural networks (NN) for the prediction of highly irregular breathing patterns. The ANN was found to be more robust as it successfully predicted all

461 respiratory patterns, while the linear filter failed for several patterns. For the successful
462 cases, the ANN achieved comparable or better accuracy compared to the linear filter.
463 As the latency was increased, the ANN become significantly more accurate than the
464 linear filter. This indicates the ANN is superior at predicting the future respiratory signal
465 pattern. Ruan et al. (96) showed that a local regression method can outperform ANN
466 for predicting motion 1 s in advance. Based on initial studies, Goodband et al. (76) and
467 Murphy et al. (82) proceed to optimise the performance of NNs to further improve the
468 prediction accuracy. These studies show that different configurations of parameters
469 should be tested to optimise the ANN model. A study by Kraus et al. (97) compared the
470 performance of respiratory motion predictors based on NN, linear regression (LR),
471 kernel density estimation (KDE), and support vector regression (SVR). The authors
472 developed patient-independent models. The predictors were validated on various
473 sampling rates and latencies using 3D lung tumour motion traces and the errors were
474 averaged over all sampling rates and latencies. The prediction errors were normalised
475 to the errors using no prediction to give 0.44, 0.46, 0.49 and 0.55 for NN, SVR, LR, and
476 KDE respectively. This shows that all the predictors approximately halved the
477 prediction errors compared to using no prediction. There were only small differences
478 between the predictors with the ANN approach being the best predictor.

479 The majority of the aforementioned studies focus on predicting respiratory
480 motion for lung cancer treatments. Dick et al. (98, 99) used an ANN to predict the
481 motion of abdominal tumours based on the lung-diaphragm border. The studies
482 demonstrated that it is feasible to predict the location of liver tumours based off the
483 lung-diaphragm border. This method shows the potential for markerless tracking of
484 abdominal tumours by tracking a surrogate organ.

485 More recent approaches use DL to predict the motion trajectories of tumours.
486 These methods commonly use CNNs, long short-term memory (LSTM), or fuzzy DL

(100-105). Lee et al. (100) showed that a CNN model performs better than a RNN. Lin et al. (101) developed a LSTM-based generalised model to predict the patient's respiratory signal. The neural network was trained and tested using real-time position management (RPM) data. When tested on unseen respiratory data with a 500 ms prediction window, the average relative mean absolute error was 0.112 and the root mean square error was 0.139. Wang et al. (105) compared the use of LSTM networks and SVR to predict external respiratory motion and internal liver motion. It was found that the LSTM network performed better than SVR. These results demonstrate the potential of LSTM networks for respiratory signal prediction. Romaguera et al. (103) incorporated convolutional LSTMs into an recurrent encoder-decoder model. The model proposed by Romaguera et al. finds the deformation between input images and then predicts the future motion. This method has shown the potential to predict motion on MRI, ultrasound and CT datasets.

In the study of Park et al. (102), a fuzzy DL model was used to predict intra- and inter-fractional variation of lung tumours. The fuzzy DL model combines fuzzy logic and an ANN component. The ANN component gives the model the ability to self-learn features and set network parameters during training. The fuzzy logic component improves the reasoning ability for uncertainty. This feature is useful for predicting inconsistent inter-fractional variation. Using 192.3 ms prediction, the fuzzy method had an RMSE of 0.5±0.8 mm, representing an improvement of 30.0% over a CNN and Hybrid motion estimation based on extended Kalman filter. The average computation time for the model was 1.5 ms. Torshabi et al. (104) also successfully used an adaptive neuro-fuzzy model to predict the tumour motion based of external surrogate. Therefore, these studies show the potential for combining fuzzy logic and NNs to achieve real-time prediction.

A unique method developed by Kim et al. (106) involved using a CNN to predict motion of the patient in advance based off their facial expressions. A camera was installed in the treatment room to monitor the patient's facial expressions. The system had an accuracy of 85.6% for detecting uncomfortable expressions which can be used to provide warning about potential sudden movements. A limitation of this approach is that only movements associated with an uncomfortable expression will be detected.

Discussion

In the past 5 years, the world has undergone a dramatic change with the advances of AI. Self-driving vehicles are becoming a reality and AI defeated a world-champion at the last traditional board game that had been unconquered by computers (107, 108). Medicine, too, has started to feel the impact of AI. In particular, AI has been used for a range of medical imaging applications including organ segmentation and localisation, lesion detection and segmentation, image registration, and predicting clinical outcomes (109). DL approaches are now commonly used for medical image analysis due to their success in recent years (109, 110). There is a gap for intrafraction motion management that had the potential to be filled by AI-based approaches. In radiotherapy, intrafraction motion is still without a satisfactory commercial solution on a standard linear accelerator. The requirement of accuracy and precision in real-time for a variety of targets is often a barrier for in-house developed solutions to become mainstream: it is far easier to ensure a solution work 90% of the time in a controlled clinical trial single centre environment compared to a commercial solution where it needs to be accurate 99.5% for every centre and every patient. As discussed in the introduction, the clinical benefit of real-time intrafraction motion monitoring is clear with mounting evidence. In the past 5 years, we have witnessed the growing application of AI for target segmentation on kV images to address the gap in intrafraction monitoring solutions.

With AI, more robust solutions for implanted marker tracking was developed with claimed sensitivity up to 99.4% (33). More excitingly, the power of AI was harnessed for markerless tracking of the target without markers despite the low soft tissue contrast on kV images, typically acquired on a standard linear accelerator. Our literature search also shows innovations of AI methods to monitor for changes in the full anatomy. Motion of a target with some periodicity can also be predicted ahead of time such that the frequency of measurements can be reduced with no decrease in tumour detection accuracy.

The use of AI for marker-based tracking approaches has been demonstrated to be highly accurate through several studies. All of the marker-based tracking approaches discussed achieved sub-millimetre accuracy (33, 44, 45). Mylonas et al. (33) and Liang et al. (44) use different DL approaches due to the differences in the image acquisition methods. The tracking window approach proposed by Mylonas et al. (33) is effective for continuous fluoroscopic imaging where the markers stay within the tracking window in subsequent image acquisitions. However, this approach is not suitable for the discrete x-ray image acquisition method used by CyberKnife. To overcome this, Liang et al. (44) utilised the U-Net architecture to perform semantic segmentation of the markers in full-size x-ray images. The method developed by Mylonas et al. (33) was successfully applied to both regular and arbitrarily shaped markers. The performance of these systems are approaching the upper limit for accuracy compared to the ground truth. Due to the previously discussed disadvantages of marker-based tracking, future work would more likely be focused on implementing markerless-based approaches.

To date, several studies have demonstrated the potential to eliminate the use of markers altogether through markerless tracking (35-37, 50-55, 57, 59-64). The implementation of markerless tracking is considered the ideal technique for the patient

564 and the healthcare system due to its non-invasiveness. Markerless-based approaches
565 for radiotherapy using AI have been developed for the pancreas, prostate, liver, lungs,
566 and vertebrae thus far. The ML approaches were only applied to lung tumour tracking
567 (50, 54, 55, 59, 62), while DL approaches were used for other sites. This may indicate
568 that ML is outperformed by DL at sites that have a lower contrast. Several studies
569 applied an R-CNN approach for tracking the target (35-37, 52). The R-CNN approach
570 helps to reduce the computation time by producing region proposals. This is ideal for
571 real-time applications since it is important to have minimal latency. The majority of
572 markerless tracking approaches in cine MRI images focused on lung tumour. Other low
573 contrast sites such as the pancreas and prostate stand to benefit from MRI-guided
574 radiotherapy due to improved tissue contrast. The markerless approaches do not
575 achieve the same level of accuracy in comparison to marker-based approaches.
576 Therefore, there is still the opportunity to further improve the performance of these
577 systems.

578 Another approach for IGRT is the monitoring of changes based on the full
579 surrounding anatomy. The majority of these methods do not produce a geometric
580 measurement that can allow for real-time motion compensation. These approaches are
581 ideal for performing gating during the treatment. Several studies used binary classifiers
582 to determine whether the target is within view for gating purposes (34, 67, 68). Novel
583 single-view reconstruction methods show potential for use during IGRT (66, 69). Single
584 projections acquired during the radiotherapy fraction could be used for 3D
585 reconstruction during the treatment. The reconstructed volumes could be compared to
586 the pre-treatment CBCT to determine whether intrafraction shifts have occurred.
587 However, the typical time for the 3D reconstruction method developed by Shen et al. is
588 around 0.5 s (66). Therefore, further refinement is required to improve the efficiency of
589 this method to allow for real time implementation. For ultrasound-guided radiotherapy,

a CNN can perform image registration of 3D ultrasound prostate images for motion management (71). Proton therapy could benefit from the implementation of the aforementioned algorithms. Clinical results of proton therapy to date have not met the original high expectations. A commonly cited reason for this discrepancy is the lack of effective image guidance for proton therapy (111-113). One prospective randomised trial for lung cancer estimated a three-fold reduction in lung toxicity with proton therapy compared with x-ray therapy from 15% to 5%. The actual results in practice were the opposite, with 11% lung toxicity (radiation pneumonitis grade ≥3) in the proton therapy arm and 7% in the x-ray arm (114). To this end, using AI to reconstruct real-time 3D volumes could be useful in both target localisation and 3D range uncertainty quantification during proton therapy treatments of internal targets.

For motion adaptation, a latency typically exists from the time the motion was observed and when the radiotherapy system can adapt it. The latency of the motion management system needs to be minimised to enable real-time adaption. Real-time is defined by the AAPM Task Group 264 as a system latency below 500 ms (115). To achieve real-time motion adaption, there is a prolific body of work that looks at motion prediction using ML and DL methods. AI-based predictors have demonstrated high accuracy in predicting the motion of tumours in the lung and abdomen. While there were mixed results in earlier studies comparing various ML algorithms, ANNs were close to or were the best performing methods. DL approaches have further improved the accuracy of motion prediction more recent studies.

Finally, the use of AI in motion tracking is in its early days. At the time of this review, the number of AI papers for intrafraction motion monitoring is modest compared with other areas of radiotherapy such as automatic contouring. Yet, the change was dramatic and hinted at unprecedented potential. Solutions such as markerless tracking of the liver, pancreas or prostate, which were often cited as the reason for MRI-linac

use are no longer far-fetched ideas for standard linear accelerators. However, most of the featured studies are initial tests conducted with often off-the-shelf algorithms, adopted from other areas of medicine and computer vision and small datasets. Furthermore, there are a range of potential pitfalls. Similar to other areas of medicine where quality assurance is key to ensure patient safety, there is currently a lack of standardisation in datasets and reporting metrics. Many studies use DRRs due to a lack of large, annotated datasets. The development of publicly available datasets and the standardisation of clinically relevant evaluation metrics would enable the growth and benchmarking of AI solutions. A major concern is the robustness of algorithms since the performance of the algorithms may be correlated to the particular dataset used for training (116). For example, if an organisation upgraded the hardware or software in their imaging system, the imaging data may change in some form. This dataset shift could potentially affect the performance of the AI algorithm that was developed using data prior to the change. Large publicly available datasets have the potential to solve this and improve the robustness of the algorithms, reducing the effects of dataset shifts. Ethical and legal concerns are another key issue that may prevent clinical implementation. Data privacy is one challenge that organisations will have to deal with in order to create large, annotated datasets for the development of AI solutions. Patients may have concerns about who has access to their data, whether their data is being used for subsequent projects or being commercialised, and the risk of privacy breaches (117). There is also the potential for bias to occur if the models are trained and then deployed on different patient demographics, leading to specific patient populations receiving improved treatments over others (116).

While there are several concerns associated with AI that make clinical implementation difficult, ML and DL are becoming increasingly explored for motion tracking due to the apparent strengths of these approaches. With the surge of AI

publications in other research fields including medical imaging, there are many well established algorithms (109, 110). Therefore, it is relatively easy to choose and implement an already established algorithm depending on the specific application. Another advantage is that many of the discussed AI approaches have low latency during inference which meet requirements outlined by the AAPM Task Group 264 (115). As discussed above, the crucial next steps to translate AI solutions for intrafraction motion monitoring are (1) the development of large annotated database, and (2) standardisation of the reporting metrics. This will help to ensure that the AI algorithms are robust to different systems and populations, achieving similar impact of the ImageNet dataset on the development of deep learning (118). Furthermore, the development of consensus regulations and guidelines regarding the use of AI would be beneficial and help enable clinical implementation. At the time of writing, the authors believe that we are at the inflection point for real-time intrafraction motion monitoring and the use of AI for motion tracking will continue to develop and increase.

Conclusion

The growth of AI in the field of medicine and image processing has led to its implementation in motion monitoring during radiotherapy. AI approaches have shown great success in tracking markers with improved accuracy compared to traditional methods. Furthermore, AI methods have demonstrated the potential to eliminate the use of markers altogether through markerless tracking. The algorithms for monitoring intrafraction changes in the full anatomy demonstrate the potential for monitoring the full anatomy in the field of view. There is a latency associated with these AI approaches and other traditional methods due to the time for computation and motion compensation. However, AI has also been demonstrated to solve this problem by accurately predicting future motion. With the continued improvement of AI techniques,

668 we can expect to see further refinement of monitoring the target during radiotherapy
669 using a standard linear accelerator.

670

671 **Acknowledgments**

672 We thank the three anonymous reviewers whose comments helped improve this
673 manuscript. We would also like to thank Dr Mark Gardner and Marco Mueller for
674 reviewing this manuscript and providing helpful suggestions.

675

676 **References**

677 1. Gregory M.M. Videtic, Donington J, Giuliani M, Heinzerling J, Karas TZ, Kelsey
678 CR, et al. Stereotactic body radiation therapy for early-stage non-small cell lung
679 cancer: Executive Summary of an ASTRO Evidence-Based Guideline. Pract Radiat
680 Oncol. 2017;7(5).
681 2. Morgan SC, Hoffman K, Loblaw A, Buyyounouski MK, Patton C, Barocas D, et
682 al. ASTRO/ASCO/AUA Guideline on Hypofractionation for Localized Prostate Cancer.
683 Practical Radiation Oncology. 2018.
684 3. Aurbry JF, Beaulieu L, Girouard L-M, Aubin S, Tremblay D, Laverdière J, et al.
685 Measurements of intrafraction motion and interfraction and intrafraction rotation of
686 prostate by three-dimensional analysis of daily portal imaging with radiopaque markers.
687 Int J Radiat Oncol Biol Phys. 2004;60(1):30-9.
688 4. Kupelian P, Willoughby T, Mahadevan A, Djemil T, Weinstein G, Jani S, et al.
689 Multi-institutional clinical experience with the Calypso System in localization and
690 continuous, real-time monitoring of the prostate gland during external radiotherapy. Int
691 J Radiat Oncol Biol Phys. 2007;67(4):1088-98.

692 5. Deutschmann H, Kametriser G, Steininger P, Scherer P, Schöller H, Gaisberger
693 C, et al. First clinical release of an online, adaptive, aperture-based image-guided
694 radiotherapy strategy in intensity-modulated radiotherapy to correct for inter- and
695 intrafractional rotations of the prostate. Int J Radiat Oncol Biol Phys. 2012;83(5):1624-
696 32.
697 6. Kotte ANTJ, Hofman P, Lagendijk JJW, Vulpen Mv, Heide UAvd. Intrafraction
698 motion of the prostate during external-beam radiation therapy: analysis of 427 patients
699 with implanted fiducial markers. Int J Radiat Oncol Biol Phys. 2007;69(2):419-25.
700 7. Li JS, Jin L, Pollack A, Horwitz EM, Buyyounouski MK, Price RA, et al. Gains
701 from real-time tracking of prostate motion during external beam radiation therapy. Int J Radiat
702 Oncol Biol Phys. 2009;75(5):1613-20.
703 8. Amro H, Hamstra D, Mcshan D, Sandler H, Vineberg K, Hadley S, et al. The
704 dosimetric impact of prostate rotations during electromagnetically guided external
705 beam radiation therapy. Int J Radiat Oncol Biol Phys. 2013;85(1):230-6.
706 9. Kerkhof EM, Put RWvd, Raaymakers BW, Heide UAvd, Jürgenliemk-Schulz IM,
707 Lagendijk JJW. Intrafraction motion in patients with cervical cancer: The benefit of soft
708 tissue registration using MRI. Radiother Oncol. 2009;93(1):115-21.
709 10. Bertholet J, Worm ES, Fledelius W, Høyer M, Poulsen PR. Time-Resolved
710 Intrafraction Target Translations and Rotations During Stereotactic Liver Radiation
711 Therapy: Implications for Marker-based Localization Accuracy. Int J Radiat Oncol Biol
712 Phys. 2016;95(2):802-9.
713 11. Worm ES, Bertholet J, Høyer M, Fledelius W, Hansen AT, Larsen LP, et al.
714 Fiducial marker guided stereotactic liver radiotherapy: Is a time delay between marker
715 implantation and planning CT needed? Radiother Oncol. 2016;121(2016):75-8.
716 12. Huang C-Y, Tehrani JN, Ng JA, Booth JT, Keall PJ. Six Degrees-of-Freedom
717 Prostate and Lung Tumor Motion Measurements Using Kilovoltage Intrafraction
718 Monitoring. Int J Radiat Oncol Biol Phys. 2015;91(2):368-75.

719 13. Steiner E, Shieh C, Caillet V, Booth J, O'Brien R, Briggs A, et al. Both four-
720 dimensional computed tomography and four-dimensional cone beam computed
721 tomography under-predict lung target motion during radiotherapy. Radiother Oncol.
722 2019;135:65-73.
723 14. Jones BL, Schefter T, Miften M. Adaptive motion mapping in pancreatic SBRT
724 patients using Fourier transforms. Radiother Oncol. 2015;115(2):217-22.
725 15. Guzhva L, Flampouri S, Mendenhall NP, Morris CG, Hoppe BS. Intrafractional
726 Displacement of Cardiac Substructures Among Patients With Mediastinal Lymphoma
727 or Lung Cancer. Advances in Radiation Oncology. 2019;4(3).
728 16. Suzanne Lydiard, Beau Pontré, Lowe BS, Ball H, Sasso G, Keall P. Cardiac
729 radioablation for atrial fibrillation: Target motion characterization and treatment delivery
730 considerations. Med Phys. 2021;48(3).
731 17. Booth JT, Caillet V, Hardcastle N, O'Brien R, Szymura K, Crasta C, et al. The
732 first patient treatment of electromagnetic-guided real time adaptive radiotherapy using
733 MLC tracking for lung SABR. Radiother Oncol. 2016;121(1):19-25.
734 18. Caillet V, Keall PJ, Colvill E, Hardcastle N, O'Brien R, Szymura K, et al. MLC
735 tracking for lung SABR reduces planning target volumes and dose to organs at risk.
736 Radiother Oncol. 2017;124(1):18-24.
737 19. Poulsen PR, Worm ES, Hansen R, Larsen LP, Grau C, Høyer M. Respiratory
738 gating based on internal electromagnetic motion monitoring during stereotactic liver
739 radiation therapy: First Results. Acta Oncologica. 2016;54(9):1445-52.
740 20. Colvill E, Poulsen PR, Booth JT, O'Brien RT, Ng JA, Keall PJ. DMLC tracking
741 and gating can improve dose coverage for prostate VMAT. Medical Physics.
742 2014;41(9).
743 21. Shirato H, Shimizu S, Kunieda T, Kitamura K, van Herk M, Kagei K, et al.
744 Physical aspects of a real-time tumor-tracking system for gated radiotherapy. Int J
745 Radiat Oncol Biol Phys. 2000;48(4):1187-95.

746 22. King CR, Brooks JD, Gill H, Pawlicki T, Cotrutz C, Presti JC. Stereotactic body
747 radiotherapy for localized prostate cancer: Interim results of a prospective phase II
748 clinical trial. Int J Radiat Oncol Biol Phys. 2009;73(4):1043-8.
749 23. Kitamura K, Shirato H, Shimizu S, Shinohara N, Harabayashi T, Shimizu T, et
750 al. Registration accuracy and possible migration of internal fiducial gold marker
751 implanted in prostate and liver treated with real-time tumor-tracking radiation therapy
752 (RTRT). Int J Radiat Oncol Biol Phys. 2002;62(3):275 - 81.
753 24. Sazawa A, Shinohara N, Harabayashi T, Abe T, Shirato H, Nonomura K.
754 Alternative approach in the treatment of adrenal metastasis with a real-time tracking
755 radiotherapy in patients with hormone refractory prostate cancer. Int J Urol.
756 2009;16(4):410-2.
757 25. Shimizu S, Shirato H, Kitamura K, Shinohara N, Harabayashi T, Tsukamoto T,
758 et al. Use of an implanted marker and real-time tracking of the marker for the
759 positioning of prostate and bladder cancers. Int J Radiat Oncol Biol Phys.
760 2000;48(5):1591-7.
761 26. Shirato H, Harada T, Harabayashi T, Hida K, Endo H, Kitamura K, et al.
762 Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers
763 for precise setup and real-time tumor tracking in radiotherapy. Int J Radiat Oncol Biol
764 Phys. 2003;56(1):240-7.
765 27. Ozhasoglu C, Saw CB, Chen H, Burton S, Komanduri K, Yue NJ, et al.
766 Synchrony – Cyberknife Respiratory Compensation Technology. Medical Dosimetry.
767 2008;33(2):117-23.
768 28. Jin J-Y, Yin F-F, Tenn SE, Medin PM, Solberg TD. Use of the BrainLAB
769 ExacTrac X-Ray 6D System in Image-Guided Radiotherapy. Medical Dosimetry.
770 2008;33(2):124-34.

771 29. Castellanos E, Ericsson MH, Sorcini B, Green U, Nilsson S, Lennernäs B.
772 RayPilot – electromagnetic real-time positioning in radiotherapy of prostate cancer –
773 initial clinical results. Radiother Oncol. 2012;103(suppl 1):S433.
774 30. Ballhausen H, Li M, Hegemann N-S, Ganswindt U, Belka C. Intra-fraction
775 motion of the prostate is a random walk. Phys Med Biol. 2015;60(2).
776 31. Fallone BG, Murray B, Rathee S, Stanescu T, Steciw S, Vidakovic S, et al. First
777 MR images obtained during megavoltage photon irradiation from a prototype integrated
778 linac-MR system. Med Phys. 2009;36:2084-8.
779 32. Raaymakers BW, Lagendijk JJW, Overweg J, Kok JGM, Raaijmakers AJE,
780 Kerkhof EM, et al. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of
781 concept. Phys Med Biol. 2009;54:N229.
782 33. Mylonas A, Keall PJ, Booth JT, Shieh CC, Eade T, Poulsen PR, et al. A deep
783 learning framework for automatic detection of arbitrarily shaped fiducial markers in
784 intrafraction fluoroscopic images. Med Phys. 2019.
785 34. Lin T, Li R, Tang X, Dy JG, Jiang SB. Markerless gating for lung cancer
786 radiotherapy based on machine learning techniques. Physics in Medicine & Biology.
787 2009;54(6):1555.
788 35. Roggen T, Bobic M, Givehchi N, Scheib SG. Deep Learning model for
789 markerless tracking in spinal SBRT. Physica Medica. 2020;74:66-73.
790 36. Zhao W, Han B, Yang Y, Buyyounouski M, Hancock SL, Bagshaw H, et al.
791 Incorporating imaging information from deep neural network layers into image guided
792 radiation therapy (IGRT). Radiotherapy and Oncology. 2019;140:167-74.
793 37. Zhao W, Shen L, Han B, Yang Y, Cheng K, Toesca DA, et al. Markerless
794 pancreatic tumor target localization enabled by deep learning. International Journal of
795 Radiation Oncology* Biology* Physics. 2019;105(2):432-9.
796 38. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data
797 mining, inference, and prediction: Springer Science & Business Media; 2009.

798 39. Bertholet J, Wan H, Toftegaard J, Schmidt M, Chotard F, Parikh P, et al. Fully
799 automatic segmentation of arbitrarily shaped fiducial markers in cone-beam CT
800 projections. Physics in Medicine & Biology. 2017;62(4):1327.
801 40. Fledelius W, Worm E, Elstrøm UV, Petersen JB, Grau C, Høyer M, et al.
802 Robust automatic segmentation of multiple implanted cylindrical gold fiducial markers
803 in cone-beam CT projections. Medical physics. 2011;38(12):6351-61.
804 41. Fledelius W, Worm E, Høyer M, Grau C, Poulsen P. Real-time segmentation of
805 multiple implanted cylindrical liver markers in kilovoltage and megavoltage x-ray
806 images. Physics in Medicine & Biology. 2014;59(11):2787.
807 42. Poulsen PR, Fledelius W, Keall PJ, Weiss E, Lu J, Brackbill E, et al. A method
808 for robust segmentation of arbitrarily shaped radiopaque structures in cone-beam CT
809 projections. Medical physics. 2011;38(4):2151-6.
810 43. Regmi R, Lovelock DM, Hunt M, Zhang P, Pham H, Xiong J, et al. Automatic
811 tracking of arbitrarily shaped implanted markers in kilovoltage projection images: a
812 feasibility study. Medical physics. 2014;41(7):071906.
813 44. Liang Z, Zhou Q, Yang J, Zhang L, Liu D, Tu B, et al. Artificial intelligence-
814 based framework in evaluating intrafraction motion for liver cancer robotic stereotactic
815 body radiation therapy with fiducial tracking. Medical Physics. 2020;47(11):5482-9.
816 45. Lin WY, Lin SF, Yang SC, Liou SC, Nath R, Liu W. Real-time automatic fiducial
817 marker tracking in low contrast cine-MV images. Medical physics. 2013;40(1):011715.
818 46. Kothary N, Heit JJ, Louie JD, Kuo WT, Loo Jr BW, Koong A, et al. Safety and
819 efficacy of percutaneous fiducial marker implantation for image-guided radiation
820 therapy. Journal of vascular and interventional radiology. 2009;20(2):235-9.
821 47. Park WG, Yan BM, Schellenberg D, Kim J, Chang DT, Koong A, et al. EUS-
822 guided gold fiducial insertion for image-guided radiation therapy of pancreatic cancer:

823 50 successful cases without fluoroscopy. Gastrointestinal endoscopy. 2010;71(3):513-
824 8.
825 48. Bhagat N, Fidelman N, Durack JC, Collins J, Gordon RL, LaBerge JM, et al.
826 Complications associated with the percutaneous insertion of fiducial markers in the
827 thorax. Cardiovascular and interventional radiology. 2010;33(6):1186-91.
828 49. Lim YK, Kwak J, Kim DW, Shin D, Yoon M, Park S, et al. Microscopic gold
829 particle-based fiducial markers for proton therapy of prostate cancer. International
830 Journal of Radiation Oncology* Biology* Physics. 2009;74(5):1609-16.
831 50. Lin T, Cervino LI, Tang X, Vasconcelos N, Jiang SB. Fluoroscopic tumor
832 tracking for image-guided lung cancer radiotherapy. Physics in Medicine & Biology.
833 2009;54(4):981.
834 51. Terunuma T, Tokui A, Sakae T. Novel real-time tumor-contouring method using
835 deep learning to prevent mistracking in X-ray fluoroscopy. Radiological physics and
836 technology. 2018;11(1):43-53.
837 52. Edmunds D, Sharp G, Winey B. Automatic diaphragm segmentation for real-
838 time lung tumor tracking on cone-beam CT projections: a convolutional neural network
839 approach. Biomedical Physics & Engineering Express. 2019;5(3):035005.
840 53. Hirai R, Sakata Y, Tanizawa A, Mori S. Real-time tumor tracking using
841 fluoroscopic imaging with deep neural network analysis. Physica Medica. 2019;59:22-
842 9.
843 54. Mori S, Sakata Y, Hirai R, Furuichi W, Shimabukuro K, Kohno R, et al.
844 Commissioning of a fluoroscopic-based real-time markerless tumor tracking system in
845 a superconducting rotating gantry for carbon-ion pencil beam scanning treatment.
846 Medical physics. 2019;46(4):1561-74.
847 55. Sakata Y, Hirai R, Kobuna K, Tanizawa A, Mori S. A machine learning-based
848 real-time tumor tracking system for fluoroscopic gating of lung radiotherapy. Physics in
849 Medicine & Biology. 2020;65(8):085014.

850 56. Takahashi W, Oshikawa S, Mori S. Real-time markerless tumour tracking with
851 patient-specific deep learning using a personalised data generation strategy: proof of
852 concept by phantom study. The British journal of radiology. 2020;93(1109):20190420.
853 57. Mori S, Hirai R, Sakata Y. Simulated four-dimensional CT for markerless tumor
854 tracking using a deep learning network with multi-task learning. Physica Medica.
855 2020;80:151-8.
856 58. Fürweger C, Drexler C, Kufeld M, Muacevic A, Wowra B, Schlaefer A. Patient
857 motion and targeting accuracy in robotic spinal radiosurgery: 260 single-fraction
858 fiducial-free cases. International Journal of Radiation Oncology* Biology* Physics.
859 2010;78(3):937-45.
860 59. Cervino LI, Du J, Jiang SB. MRI-guided tumor tracking in lung cancer
861 radiotherapy. Physics in Medicine & Biology. 2011;56(13):3773.
862 60. Yun J, Yip E, Gabos Z, Wachowicz K, Rathee S, Fallone B. Neural-network
863 based autocontouring algorithm for intrafractional lung-tumor tracking using Linac-MR.
864 Medical physics. 2015;42(5):2296-310.
865 61. Yun J, Yip E, Gabos Z, Wachowicz K, Rathee S, Fallone B. Improved lung
866 tumor autocontouring algorithm for intrafractional tumor tracking using 0.5 T linac-MR.
867 Biomedical Physics & Engineering Express. 2016;2(6):067004.
868 62. Bourque AE, Carrier J-F, Filion É, Bedwani S. A particle filter motion prediction
869 algorithm based on an autoregressive model for real-time MRI-guided radiotherapy of
870 lung cancer. Biomedical Physics & Engineering Express. 2017;3(3):035001.
871 63. Fast MF, Eiben B, Menten MJ, Wetscherek A, Hawkes DJ, McClelland JR, et al.
872 Tumour auto-contouring on 2d cine MRI for locally advanced lung cancer: A
873 comparative study. Radiotherapy and Oncology. 2017;125(3):485-91.
874 64. Friedrich F, Hörner-Rieber J, Renkamp CK, Klüter S, Bachert P, Ladd ME, et al.
875 Stability of conventional and machine learning-based tumor auto-segmentation

876 techniques using undersampled dynamic radial bSSFP acquisitions on a 0.35 T hybrid
877 MR-linac system. Medical Physics. 2021;48(2):587-96.
878 65. Cusumano D, Boldrini L, Dhont J, Fiorino C, Green O, Güngör G, et al. Artificial
879 Intelligence in magnetic Resonance guided Radiotherapy: Medical and physical
880 considerations on state of art and future perspectives. Physica Medica. 2021;85:175-
881 91.
882 66. Shen L, Zhao W, Xing L. Patient-specific reconstruction of volumetric computed
883 tomography images from a single projection view via deep learning. Nature biomedical
884 engineering. 2019;3(11):880-8.
885 67. Tang X, Lin T, Jiang S. A feasibility study of treatment verification using EPID
886 cine images for hypofractionated lung radiotherapy. Physics in Medicine & Biology.
887 2009;54(18):S1.
888 68. Cui Y, Dy JG, Alexander B, Jiang SB. Fluoroscopic gating without implanted
889 fiducial markers for lung cancer radiotherapy based on support vector machines.
890 Physics in Medicine & Biology. 2008;53(16):N315.
891 69. Xu Y, Yan H, Ouyang L, Wang J, Zhou L, Cervino L, et al. A method for
892 volumetric imaging in radiotherapy using single x-ray projection. Medical physics.
893 2015;42(5):2498-509.
894 70. Terpstra ML, Maspero M, d'Agata F, Stemkens B, Intven MP, Lagendijk JJ, et
895 al. Deep learning-based image reconstruction and motion estimation from
896 undersampled radial k-space for real-time MRI-guided radiotherapy. Physics in
897 Medicine & Biology. 2020;65(15):155015.
898 71. Zhu N, Najafi M, Han B, Hancock S, Hristov D. Feasibility of Image Registration
899 for Ultrasound-Guided Prostate Radiotherapy Based on Similarity Measurement by a
900 Convolutional Neural Network. Technology in cancer research & treatment.
901 2019;18:1533033818821964.

902 72. Balasubramanian A, Shamsuddin R, Prabhakaran B, Sawant A. Predictive
903 modeling of respiratory tumor motion for real-time prediction of baseline shifts. Physics
904 in Medicine & Biology. 2017;62(5):1791.
905 73. Bukhari W, Hong S. Real-time prediction and gating of respiratory motion using
906 an extended Kalman filter and Gaussian process regression. Physics in Medicine &
907 Biology. 2014;60(1):233.
908 74. Bukovsky I, Homma N, Ichiji K, Cejnek M, Slama M, Benes PM, et al. A fast
909 neural network approach to predict lung tumor motion during respiration for radiation
910 therapy applications. BioMed research international. 2015;2015.
911 75. Dürichen R, Wissel T, Ernst F, Schweikard A, editors. Respiratory motion
912 compensation with relevance vector machines. International Conference on Medical
913 Image Computing and Computer-Assisted Intervention; 2013: Springer.
914 76. Goodband JH, Haas OC, Mills J. A comparison of neural network approaches
915 for on-line prediction in IGRT. Medical physics. 2008;35(3):1113-22.
916 77. Hong S, Bukhari W. Real-time prediction of respiratory motion using a cascade
917 structure of an extended Kalman filter and support vector regression. Physics in
918 Medicine & Biology. 2014;59(13):3555.
919 78. Isaksson M, Jalden J, Murphy MJ. On using an adaptive neural network to
920 predict lung tumor motion during respiration for radiotherapy applications. Medical
921 physics. 2005;32(12):3801-9.
922 79. Jiang K, Fujii F, Shiinoki T. Prediction of lung tumor motion using nonlinear
923 autoregressive model with exogenous input. Physics in Medicine & Biology.
924 2019;64(21):21NT02.
925 80. Jurkovic I-A, Stathakis S, Papanikolaou N, Mavroidis P. Prediction of lung tumor
926 motion extent through artificial neural network (ANN) using tumor size and location
927 data. Biomedical Physics & Engineering Express. 2016;2(2):025012.

928 81. Mafi M, Moghadam SM. Real-time prediction of tumor motion using a dynamic
929 neural network. Medical & biological engineering & computing. 2020;58(3):529-39.
930 82. Murphy MJ, Pokhrel D. Optimization of an adaptive neural network to predict
931 breathing. Medical physics. 2009;36(1):40-7.
932 83. Riaz N, Shanker P, Wiersma R, Gudmundsson O, Mao W, Widrow B, et al.
933 Predicting respiratory tumor motion with multi-dimensional adaptive filters and support
934 vector regression. Physics in Medicine & Biology. 2009;54(19):5735.
935 84. Ruan D, Keall P. Online prediction of respiratory motion: multidimensional
936 processing with low-dimensional feature learning. Physics in Medicine & Biology.
937 2010;55(11):3011.
938 85. Sun W, Jiang M, Ren L, Dang J, You T, Yin F. Respiratory signal prediction
939 based on adaptive boosting and multi-layer perceptron neural network. Physics in
940 Medicine & Biology. 2017;62(17):6822.
941 86. Sun W, Wei Q, Ren L, Dang J, Yin F-F. Adaptive respiratory signal prediction
942 using dual multi-layer perceptron neural networks. Physics in Medicine & Biology.
943 2020;65(18):185005.
944 87. Teo TP, Ahmed SB, Kawalec P, Alayoubi N, Bruce N, Lyn E, et al. Feasibility of
945 predicting tumor motion using online data acquired during treatment and a generalized
946 neural network optimized with offline patient tumor trajectories. Medical physics.
947 2018;45(2):830-45.
948 88. Torshabi AE, Pella A, Riboldi M, Baroni G. Targeting accuracy in real-time
949 tumor tracking via external surrogates: a comparative study. Technology in cancer
950 research & treatment. 2010;9(6):551-61.
951 89. Yan H, Yin F-F, Zhu G-P, Ajlouni M, Kim JH. Adaptive prediction of internal
952 target motion using external marker motion: a technical study. Physics in Medicine &
953 Biology. 2005;51(1):31.

954 90. Dhont J, Vandemeulebroucke J, Cusumano D, Boldrini L, Cellini F, Valentini V,
955 et al. Multi-object tracking in MRI-guided radiotherapy using the tracking-learning-
956 detection framework. Radiotherapy and Oncology. 2019;138:25-9.
957 91. Mirzapour SA, Mazur T, Sharp G, Salari E. Intra-fraction motion prediction in
958 MRI-guided radiation therapy using Markov processes. Physics in Medicine & Biology.
959 2019;64(19):195006.
960 92. Yun J, Mackenzie M, Rathee S, Robinson D, Fallone B. An artificial neural
961 network (ANN)-based lung-tumor motion predictor for intrafractional MR tumor tracking.
962 Medical physics. 2012;39(7Part1):4423-33.
963 93. Prall M, Kaderka R, Saito N, Graeff C, Bert C, Durante M, et al. Ion beam
964 tracking using ultrasound motion detection. Medical physics. 2014;41(4):041708.
965 94. Sharp GC, Jiang SB, Shimizu S, Shirato H. Prediction of respiratory tumour
966 motion for real-time image-guided radiotherapy. Physics in Medicine & Biology.
967 2004;49(3):425.
968 95. Murphy MJ, Dieterich S. Comparative performance of linear and nonlinear
969 neural networks to predict irregular breathing. Physics in Medicine & Biology.
970 2006;51(22):5903.
971 96. Ruan D, Fessler JA, Balter J. Real-time prediction of respiratory motion based
972 on local regression methods. Physics in Medicine & Biology. 2007;52(23):7137.
973 97. Krauss A, Nill S, Oelfke U. The comparative performance of four respiratory
974 motion predictors for real-time tumour tracking. Physics in Medicine & Biology.
975 2011;56(16):5303.
976 98. Dick D, Wu X, Hatoum GF, Zhao W. A fiducial-less tracking method for
977 radiation therapy of liver tumors by diaphragm disparity analysis part 2: validation study
978 by using clinical data. Journal of Radiation Oncology. 2018;7(4):345-56.
979 99. Dick D, Wu X, Hatoum GF, Zhao W. A fiducial-less tracking method for
980 radiation therapy of liver tumors by diaphragm disparity analysis part 1: simulation

981 study using machine learning through artificial neural network. Journal of Radiation
982 Oncology. 2018;7(3):275-84.
983 100. Lee SJ, Motai Y, Weiss E, Sun SS. Customized prediction of respiratory motion
984 with clustering from multiple patient interaction. ACM Transactions on Intelligent
985 Systems and Technology (TIST). 2013;4(4):1-17.
986 101. Lin H, Shi C, Wang B, Chan MF, Tang X, Ji W. Towards real-time respiratory
987 motion prediction based on long short-term memory neural networks. Physics in
988 Medicine & Biology. 2019;64(8):085010.
989 102. Park S, Lee SJ, Weiss E, Motai Y. Intra-and inter-fractional variation prediction
990 of lung tumors using fuzzy deep learning. IEEE journal of translational engineering in
991 health and medicine. 2016;4:1-12.
992 103. Romaguera LV, Plantefève R, Romero FP, Hébert F, Carrier J-F, Kadoury S.
993 Prediction of in-plane organ deformation during free-breathing radiotherapy via
994 discriminative spatial transformer networks. Medical image analysis. 2020;64:101754.
995 104. Torshabi AE. Investigation of the robustness of adaptive neuro-fuzzy inference
996 system for tracking moving tumors in external radiotherapy. Australasian physical &
997 engineering sciences in medicine. 2014;37(4):771-8.
998 105. Wang G, Li Z, Li G, Dai G, Xiao Q, Bai L, et al. Real-time liver tracking
999 algorithm based on LSTM and SVR networks for use in surface-guided radiation
1000 therapy. Radiation Oncology. 2021;16(1):1-12.
1001 106. Kim KH, Park K, Kim H, Jo B, Ahn SH, Kim C, et al. Facial expression
1002 monitoring system for predicting patient’s sudden movement during radiotherapy using
1003 deep learning. Journal of Applied Clinical Medical Physics. 2020;21(8):191-9.
1004 107. Ma Y, Wang Z, Yang H, Yang L. Artificial intelligence applications in the
1005 development of autonomous vehicles: a survey. IEEE/CAA Journal of Automatica
1006 Sinica. 2020;7(2):315-29.

1007 108. Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, et al.
1008 Mastering the game of go without human knowledge. nature. 2017;550(7676):354-9.
1009 109. Savadjiev P, Chong J, Dohan A, Vakalopoulou M, Reinhold C, Paragios N, et
1010 al. Demystification of AI-driven medical image interpretation: past, present and future.
1011 European radiology. 2019;29(3):1616-24.
1012 110. Ker J, Wang L, Rao J, Lim T. Deep learning applications in medical image
1013 analysis. Ieee Access. 2017;6:9375-89.
1014 111. Oborn BM, Dowdell S, Metcalfe PE, Crozier S, Mohan R, Keall PJ. Future of
1015 medical physics: real-time MRI-guided proton therapy. Medical physics.
1016 2017;44(8):e77-e90.
1017 112. Liao Z, Mohan R. Future of Protons Depends on Precision. Journal of clinical
1018 oncology: official journal of the American Society of Clinical Oncology.
1019 2018;36(19):2002-.
1020 113. Beddok A, Vela A, Calugaru V, Tessonnier T, Kubes J, Dutheil P, et al. Proton
1021 therapy for head and neck squamous cell carcinomas: A review of the physical and
1022 clinical challenges. Radiotherapy and Oncology. 2020;147:30-9.
1023 114. Liao Z, Lee JJ, Komaki R, Gomez DR, O'Reilly MS, Fossella FV, et al.
1024 Bayesian adaptive randomization trial of passive scattering proton therapy and
1025 intensity-modulated photon radiotherapy for locally advanced non–small-cell lung
1026 cancer. Journal of Clinical Oncology. 2018;36(18):1813.
1027 115. Keall PJ, Sawant A, Berbeco RI, Booth JT, Cho B, Cerviño LI, et al. AAPM Task
1028 Group 264: The safe clinical implementation of MLC tracking in radiotherapy. Medical
1029 physics. 2020.
1030 116. Subbaswamy A, Saria S. From development to deployment: dataset shift,
1031 causality, and shift-stable models in health AI. Biostatistics. 2020;21(2):345-52.

1032 117. Ahmad OF, Stoyanov D, Lovat LB. Barriers and pitfalls for artificial intelligence
1033 in gastroenterology: ethical and regulatory issues. Techniques and Innovations in
1034 Gastrointestinal Endoscopy. 2020;22(2):80-4.
1035 118. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L, editors. Imagenet: A large-
1036 scale hierarchical image database. 2009 IEEE conference on computer vision and
1037 pattern recognition; 2009: Ieee.

1038