Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

ACCEPTED MANUSCRIPT

Deep learning for dose assessment in radiotherapy by the super-localization of vaporized nanodroplets in high frame rate ultrasound imaging

To cite this article before publication: Brent van der Heyden *et al* 2022 *Phys. Med. Biol.* in press https://doi.org/10.1088/1361-6560/ac6cc3

MANUSCRIPT

Deep learning for dose assessment in radiotherapy by the super-localization of vaporized nanodroplets in high frame rate ultrasound imaging

Brent van der Heyden[1,2], Sophie V Heymans[3,1,4], Bram Carlier[2,5], Gonzalo Collado-Lara[4], Edmond Sterpin[2], Jan D'hooge[1,*]

(1) KU Leuven, Department of Cardiovascular Sciences, Leuven, Belgium

(2) KU Leuven, Department of Oncology, Laboratory of Experimental Radiotherapy, Leuven, Belgium

(3) KU Leuven Campus Kulak, Department of Physics, Kortrijk, Belgium

(4) Biomedical Engineering, Department of Cardiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands

(5) KU Leuven, Molecular Small Animal Imaging Center (MoSAIC), Leuven, Belgium

(*) **Corresponding author:**
Prof. Dr. Jan D'hooge
jan.dhooge@uzleuven.be

Keywords: proton therapy, range verification, dosimetry, ultrasound, nanodroplets, deep learning

Abstract (299/300 words)

Objective: External beam radiotherapy is aimed to precisely deliver a high radiation dose to malignancies, while optimally sparing surrounding healthy tissues. With the advent of increasingly complex treatment plans, the delivery should preferably be verified by Quality Assurance methods. Recently, online ultrasound imaging of vaporized radiosensitive nanodroplets was proposed as a promising tool for *in vivo* dosimetry in radiotherapy. Previously, the detection of sparse vaporization events was achieved by applying differential ultrasound (US) imaging followed by intensity thresholding using subjective parameter tuning, which is sensitive to image artifacts. *Approach:* A generalized deep learning solution (i.e., BubbleNet) is proposed to localize vaporized nanodroplets on differential US frames, while overcoming the aforementioned limitation. A 5-fold cross-validation was performed on a diversely composed 5747-frame training/validation dataset by manual segmentation. BubbleNet was then applied on a test dataset of 1536 differential US frames to evaluate dosimetric features. The intra-observer variability was determined by scoring the Dice Similarity Coefficient (DSC) on 150 frames segmented twice. Additionally, the BubbleNet generalization capability was tested on an external test dataset of 432 frames acquired by a phased array transducer at a much lower ultrasound frequency and reconstructed with unconventional pixel dimensions with respect to the training dataset. *Main Results:* The median DSC in the 5-fold cross validation was equal to ~0.88, which was in line with the intra-observer variability (=0.86). Next, BubbleNet was employed to detect vaporizations in differential US frames obtained during the irradiation of phantoms with a 154 MeV proton beam or a 6MV photon beam. BubbleNet improved the bubble-count statistics by ~30% compared to the earlier established intensity-weighted thresholding. The proton range was verified with a ~0.8 mm accuracy. *Significance*: BubbleNet is a flexible tool to localize individual vaporized nanodroplets on experimentally acquired US images, which improves the sensitivity compared to former thresholding-weighted methods.

1. Introduction

The utilization of innovative radiotherapy is one of the key elements for treating cancer (Baskar *et al.*, 2012). Several studies concluded that approximately 50% of all cancer patients worldwide could benefit from receiving radiotherapy at least once during their course of treatment (Barton *et al.*, 2014; Atun *et al.*, 2015; Borras *et al.*, 2015a; Borras *et al.*, 2015b). External beam radiotherapy is one of the common approaches used to deliver a precise high-dose distribution localized to the tumor volume, while optimally sparing the surrounding anatomical structures sensitive to radiation. Here, sharp tumor dose gradients can be achieved with increasingly complex radiotherapy treatment plans, of which the delivery should preferably be verified in quality assurance (QA) programs. Most clinics already introduced patient specific QA programs to evaluate the accuracy of complex modulated beams with film dosimeters or ionization chambers prior to actual patient treatment (Agazaryan *et al.*, 2003; Bellec *et al.*, 2017; Li and Hsi, 2017). Despite all efforts (Yeung *et al.*, 2005; Bissonnette *et al.*, 2012; Arjomandy *et al.*, 2019), it is still possible that treatment inaccuracies can occur during the irradiation of an individual patient. These inaccuracies can be introduced due to anatomical changes between the series of treatment fractions (Kwint *et al.*, 2014; Sonke *et al.*, 2019; Noble *et al.*, 2019), patient related motion or misalignments (Brandner *et al.*, 2017), dose calculation errors caused by uncertainties in the CT calibration (Wohlfahrt and Richter, 2020), or possibly sporadic machine failure. Therefore, *in vivo* dosimetry techniques have been suggested, and developed, to detect delivery

errors in radiotherapy treatments, assessing clinically relevant differences between planned and delivered dose on a patient-by-patient basis (Verhaegen *et al.*, 2020).

In photon therapy, electronic portal imaging (van Elmpt *et al.*, 2008; Mijnheer *et al.*, 2013) is one of the most promising dosimetry techniques to detect major treatment errors, but the method is mainly limited to the conventional linac systems, and not directly suited for robotic radiotherapy devices such as the CyberKnife® device, for instance. In hadron therapy, a series of technical solutions exist, or are under development, such as post-treatment PET-based range verification (Parodi *et al.*, 2007), prompt gamma spectroscopy imaging (Hueso-Gonzalez *et al.*, 2018), or ionoacoustic imaging (Kellnberger *et al.*, 2016). Despite these great advances, proton range verification solutions did not manage to disseminate widely in proton therapy centers beyond a few pioneering centers (Knopf *et al.*, 2011; Richter *et al.*, 2016; Hueso-Gonzalez *et al.*, 2018) due to their technological complexity, their extensive workflow, or their additional technological cost, which is already expensive upfront.

Alternatively, ultrasound (US) imaging of radiosensitive superheated nanodroplets was explored as a promising solution for on- or offline *in vivo* dosimetry purposes in radiotherapy with photons or hadrons (Carlier *et al.*, 2020; Heymans *et al.*, 2021; Toumia *et al.*, 2021; Collado-Lara *et al.*, 2022). These nanodroplets are intended to be injected intravenously, whereafter they diffuse in the patient's vasculature. Considering that the perfluorocarbon liquid core of the nanodroplets is operated at a superheated state (i.e., at a temperature above its boiling point) during irradiation, the entire nanodroplet can be converted into a microbubble US contrast agent if sufficient energy is deposited by the ionizing radiation over a certain length (i.e., the radiation exhibits sufficient LET).

Former studies (Toumia *et al.*, 2021; Carlier *et al.*, 2020; Heymans *et al.*, 2021; Collado-Lara *et al.*, 2022) focused on demonstrating the feasibility of nanodroplet-assisted radiation sensing using rather rudimentary US processing techniques. For example, Collado-Lara *et al.*, 2022 demonstrated that the distribution of vaporization events is closely related to the ion distribution produced by the proton therapy beam. In this study, the vaporized nanodroplets were detected by means of consecutive frame subtraction in the B-mode ultrasound images, which revealed the newly formed microbubbles in each differential frame as sparse bright spots. Each event with an intensity above a given threshold was counted as a vaporization. Importantly, the thresholding algorithms adopted to identify vaporization events suffer from exhaustive parameter tuning, especially for unseen US datasets. Here, new widely applicable methods are essential to localize nanodroplet vaporizations, and to evolve into a more matured technique for dose assessment in radiotherapy. Moreover, recent developments have revealed the capability of deep neural networks to segment and localize vaporized nanodroplets in super-resolution US images (Liu *et al.*, 2020; Park *et al.*, 2020; van Sloun *et al.*, 2021). In this work, the authors propose to extend the use of deep neural networks to detect the vaporizations of the radiosensitive nanodroplets induced by radiotherapeutic beams.

A deep convolutional neural network, named BubbleNet, was developed to localize individual vaporized radiosensitive nanodroplets on differential high-resolution US images. The neural network was trained on a set of manually segmented experimental data. The performance of BubbleNet was evaluated on an extensive and diverse measurement dataset, and the dose monitoring abilities were

demonstrated to measure the proton range and to report a relative measure of the photon dose distribution.

2. Materials and methods

2.1. Nanodroplet and phantom synthesis

A combination of retrospective datasets was used in this study to evaluate the performance of our proposed neural network architecture in a variety of irradiation and ultrasound imaging conditions. Briefly, the nanodroplets used in this retrospective study contain a perfluorobutane liquid core (C_4F_{10}, boiling point -2°C) encapsulated in a polyvinyl alcohol shell (PVA-PFB) and were synthesized following the protocols described elsewhere (Toumia *et al.*, 2021; Heymans *et al.*, 2021; Collado-Lara *et al.*, 2022). Nanodroplets (size range: 200 nm-8 µm) were all homogeneously dispersed and immobilized in an aqueous gel made of 0.1% Carbopol, prepared as described in (Collado-Lara *et al.*, 2022), and pre-heated to the desired temperature (50°C for proton irradiations, and 65°C for photon irradiations).

2.2. Irradiation conditions and imaging setup

The aqueous phantoms (Figure 1) were immersed in a water tank equipped with temperature control and heated to the desired temperature (50°C or 65°C). Phantom irradiations were conducted with (*i*) a 154 MeV monoenergetic proton pencil beam (Holland PTC, Delft, the Netherlands), using the same procedure as in (Collado-Lara *et al.*, 2022), or (*ii*) a 6 MV photon beam (TrueBeam®, Varian Medical Systems) at the clinical radiotherapy facility of UZ Leuven (Belgium). The specifications related to the adopted irradiation geometries and the conducted irradiations are summarily described in Subsections 2.2.1 and 2.2.2.

2.2.1. Proton irradiations

The proton beam irradiations were planned to pass through the center of every phantom's cross section, while the 80% distal fall-off was approximately located in the middle of the phantom length (Figure 1a and Figure1c). Online US imaging was performed using a linear US array operating at 8.9 MHz (ATL L12-5, 38 mm aperture) positioned at the center of the phantom length and connected to an US research system (Verasonics Vantage 256, Kirkland, USA). The phantom was imaged at a frame rate of 1kHz for a duration of 10 seconds, encompassing the irradiation time (5-7 seconds), using a plane wave imaging sequence (one angle of 0° and no compounding). Additionally, an external test dataset was acquired using a lower frequency US array (P4-2v, 19 mm aperture, 2.72 MHz transmit frequency) which transmitted three diverging waves (-18°, 0°, 18°) and imaged the phantom at a lower frame rate of 152 Hz. The US images were beamformed offline by the Verasonics beamformer.

2.2.2. Photon irradiations

The phantoms were irradiated with photon beams from their side with a square field of 10x10 cm^2 at isocenter. The source-to-skin distance was set to 90 cm, and the phantom center was located at 10 cm distance from the beam entrance point. The top left boundary of the radiation field, defined as the position where the radiation dose dropped to 50% of its maximum, was located within the phantom (Figure 1b). The delivered radiation doses ranged from 0.1 to 4.0 Gy, and the dose rates varied between 0.1 and 4.0 Gy/min.

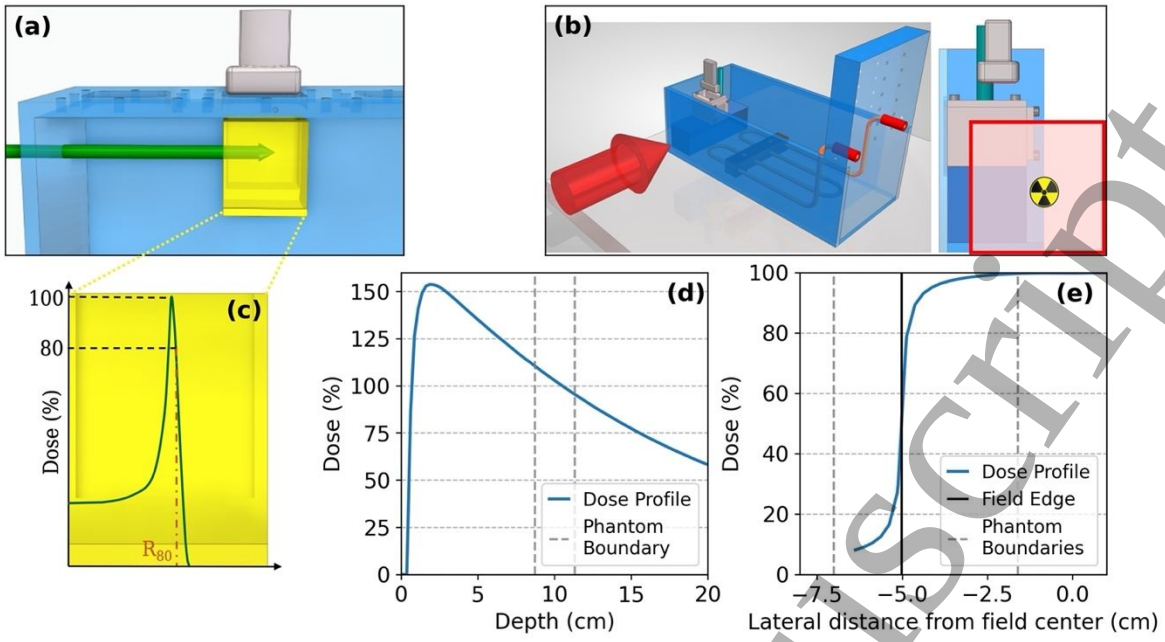

Figure 1. Online ultrasound imaging during irradiation of phantoms with PVA-PFB nanodroplets using a proton beam (green arrow, a) and a photon beam (red arrow, b). During the proton irradiation, the phantom was imaged parallel to the beam direction (c, as range verification), while for the photon irradiation (d), the ultrasound (US) imaging plane is transverse to the radiation beam (e, detection of lateral dose fall-off).

The phantoms were imaged during the photon irradiation by a 7.5 MHz linear array (L7-XTech, Vermon, France) and connected to a research US system (DiPhAs, Fraunhofer IBMT, Germany). The probe was positioned orthogonal to the direction of the radiation beam and was imaging a cross-section of the center of the phantom (Figure 1b). Plane wave imaging was employed (5 angles compounding) at pressures sufficiently low to avoid acoustic nanodroplet vaporization at 65°C. The US images were continuously acquired at a 10Hz frame rate (images were beamformed in real-time by DiPhAs). A detailed table with the irradiation conditions (particle type, dose, dose rate, concentration) and ultrasound imaging conditions (center frequency, frame rate, array type, image size) per phantom setup was added as Supplementary Material A.

2.3. US Image Processing and Neural network training

Sparse vaporization events (i.e., resulting microbubbles) were detected on a frame-by-frame basis by means of differential imaging, which consists in the subtraction of consecutively acquired US frames (Collado-Lara *et al.*, 2022). A post-processing step was added to denoise and smoothen the differential US images with a 2D Wiener filter and a 2D Gaussian filter. All image related filter settings are listed in Supplementary Material A. The differential images were normalized to [0, 1] between the minimum and the maximum image intensity values. A training- and cross-validation dataset of 5747 frames was composed manually from several phantom acquisitions, by segmenting all individual bubbles that appeared in a series of differential US frames. Differential US frames (n=672 out of 5747) without the presence of vaporized nanodroplets were also included in the training dataset. These bubbleless frames were acquired prior or after the irradiations. The training- and cross-validation dataset includes a variety of measurement data from 14 different phantoms setups and/or irradiation conditions (Supplementary Material A). The US pixel dimensions of the training dataset acquired at 7.5-8.9 MHz

varied between 71 µm-86 µm in the US radial dimension, and between 99 µm-250 µm in the US lateral dimension. Additionally, 1536 differential US frames from one phantom irradiation with 154 MeV protons were manually segmented to assess the BubbleNet performance on a never seen test dataset. Moreover, the generalization capability of BubbleNet was evaluated on an external test dataset of 432 manually segmented US frames, which were acquired with a different US frequency than the training dataset (2.7 MHz), and a correspondingly larger pixel size of 283 x 495 µm^2. Ultimately, the intra-observer variability was evaluated on 150 differential US frames, which were segmented twice by the same annotator within a sufficiently long waiting period of approximately 5 months.

2.4. BubbleNet - Network architecture

A nested U-Net architecture (i.e., U-Net++ L^2) (Zhou *et al.*, 2020) was built using Keras v2.4.2 (Chollet, 2015) to segment the vaporized nanodroplets in the differential US images. The U-Net++ L^2 network architecture, named BubbleNet, consists of an encoder-decoder pathway connected via skip connections and a nested convolutional block (Figure 2). The nested convolutional block bridges the semantic gap between the encoder and decoder feature maps. Every block of the encoder path applies successive 3 x 3 convolutional layers followed by a rectified linear unit (ReLU). Batch normalization operations were initiated to better stabilize the learning process. The spatial resolution in both dimensions is halved every block by applying a 2 x 2 max pooling layer. The number of feature maps after each block doubles along the encoder path from 32 to 128. The decoder path is symmetrical to the encoding path, with the exception that the convolution of each block is replaced by a transposed convolutional layer. Adam was employed as stochastic training optimizer with an initial learning rate of 1E-5 and was parametrized with β_1 = 0.900, β_2 = 0.999 (Kingma and Lei, 2015). The binary cross-entropy metric was adopted as probabilistic loss function to train BubbleNet, the fixed batch size was set to 18, and the number of epochs to 100. Data augmentation was applied every epoch in the form of random flipping, scaling [95%-105%] and translational [-3 px, 3 px] operations. Varying US image sizes are handled by the fully connected neural network architecture because the image width and image height are automatically padded to a value that is equal to a multiple of four, considering the two max pooling operations. A 5-fold cross validation according to an 80% train-validation (n=4697/5747) and 20% test (n=1050/5747) test scheme was employed in the model evaluation procedure to prevent the risk of unanticipated overfitting. All network calculations were performed on a single NVIDIA Volta V100 GPU using the Tier2 high-performance computer of the Flemish supercomputer center.

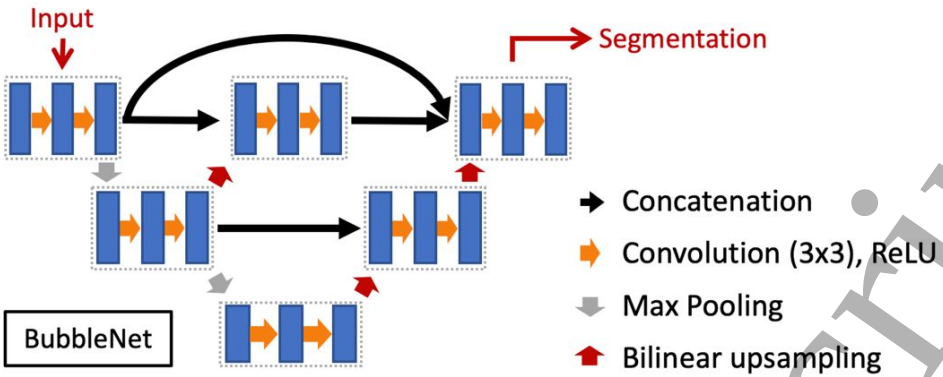

Figure 2. The BubbleNet network architecture (U-Net++) to segment and identify individual vaporized nanodroplets in differential Ultrasound (US) images.

2.5. BubbleNet – Network performance evaluation

The segmentation performance of BubbleNet was mainly evaluated with the Dice Similarity Coefficient (DSC). The DSC is a metric that represents the two-dimensional overlap between two segmentations made on one reference image. More specifically, a DSC value equal to 1 indicates a perfect overlap with identical segmentations, while a DSC value equal to 0 indicates no overlap. Then, the individual bubble coordinates (i.e., centers-of-mass) are determined by two-dimensional weighting of the bubble segmentation mask on top of the intensity values of the differential US image, regardless of the bubble segmentation technique. The bubble coordinates are then conventionally binned to obtain one-or two-dimensional histograms, and subsequently, to assess the radiation dose in one or more of the preferred dimensions.

For the proton irradiations, the one-dimensional histograms were compared to a verified model of the proton Bragg peak (Bortfeld, 1997). In the verification process of this model, the proton depth dose distribution was measured with a multi-layer ionization chamber (QubeNext, DE.TEC.TOR, Turin, Italy) and fitted to an analytical model of the Bragg peak, as explained in (Collado-Lara *et al.*, 2021; Collado-Lara *et al.*, 2022). To evaluate the proton range after bubble localization, Gaussian curves were fitted through the middle of the histogram bins for each of the bubble localization methods. A Gaussian distribution was fitted through the histogram at the end of the projected proton range where most of the primary protons reach their maximal LET value, and hence they can trigger droplet vaporizations (Collado-Lara *et al.*, 2022). The center position of the Gaussian fit is expected to coincide with the mean proton range at which 50% of the protons have stopped (Paganetti, 2012), i.e., the 80% distal dose fall-off (R_{80}). The R_{80} values obtained by BubbleNet, intensity-weighted thresholding, and manual segmentation will be compared with the analytical model, as a reference. For the photon irradiations, a comparison was made between the obtained two-dimensional histograms, and the isodose lines exported from the clinical treatment planning system, assuming a homogeneously distributed aqueous material.

3. Results

3.1. BubbleNet – Model accuracy and performance

The bubble segmentation performance between BubbleNet and a manual annotator is compared qualitatively on randomly picked frames out of our test dataset (n=1536) in Figure 3. A cropped US field-of-view is shown for optimal visualization purposes. Figure 3 depicts the bubble predictions on four differential US frames that were acquired at a high-frame rate of 1000 fps (a-d), which indirectly resulted in a low bubble count at the employed dose rates (i.e., 1-3 counts per frame in these examples). The individual DSC values were equal to 0.79, 0.92, 0.90, and 0.86 for the frames presented in Figure 3a-d respectively.

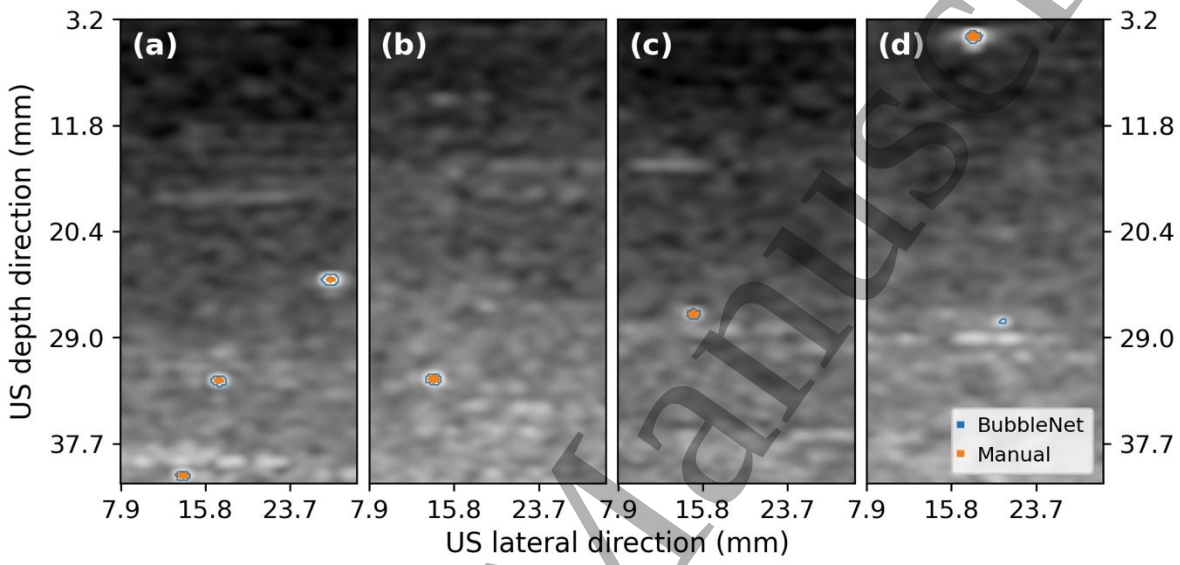

Figure 3. A comparison in the bubble identification performance between BubbleNet and a manual annotator on four (a-d) randomly picked differential Ultrasound frames with a low number of vaporizations per frame. BubbleNet produces a segmentation map, but for visualization purposes, only the outline of the segmentation is displayed.

In Figure 4, the bubble segmentation predictions are shown on four randomly picked differential US frames out of our test dataset (*n*=1536) that were acquired with an irradiation-imaging combination that resulted in substantially higher bubble counts (i.e., 18-40 counts per frame in these examples) than the results presented in Figure 3. The individual DSC values were equal to 0.71, 0.77, 0.73, and 0.79 for the frames presented in Figure 4a-d respectively.

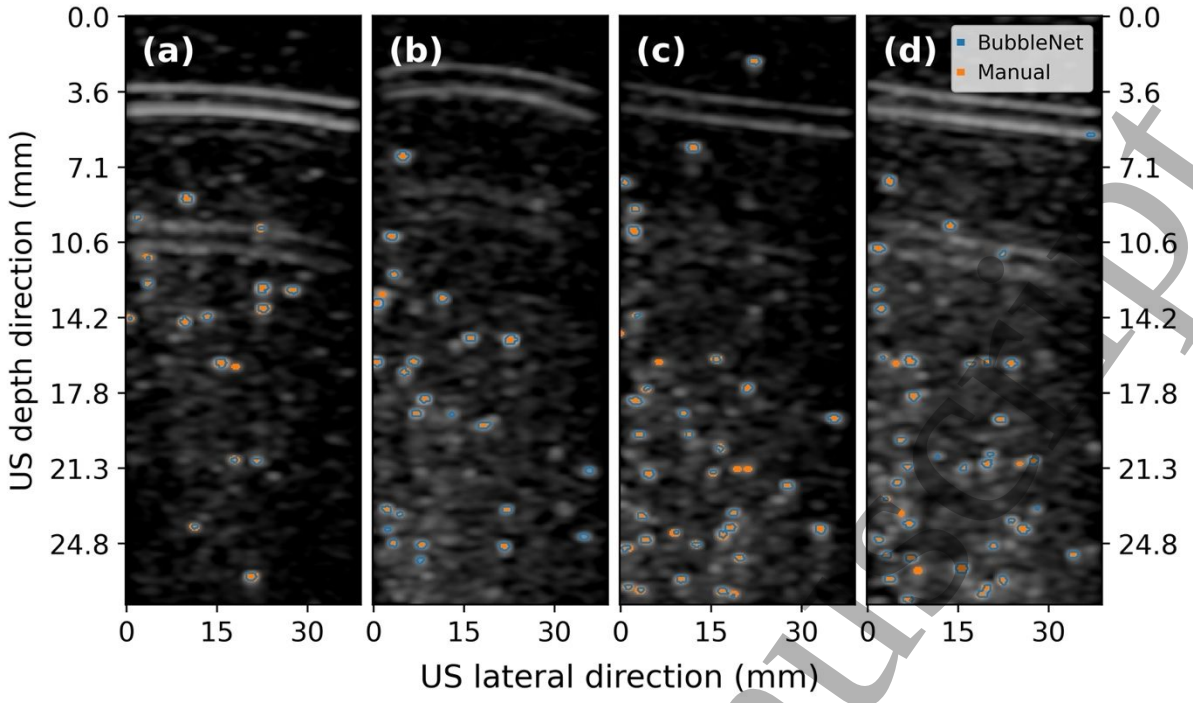

Figure 4. A comparison in the bubble identification performance between BubbleNet and a manual annotator on four (a-d) randomly picked differential Ultrasound frames with a high number of vaporizations per frame.

A largely different external test dataset (*n*=432) in terms of imaging frequency (2.72 MHz instead of 7.5-8.9 MHz), field-of-view (phased array instead of linear array), and pixel size (283 x 495 μm^2 instead of e.g., 71 x 250 μm^2, 86 x 99 μm^2, etc.) was included in the evaluation procedure of BubbleNet to evaluate the generalization performance of the deep convolutional neural network. An illustration is found in Figure 5, for which the segmentation performance of BubbleNet is qualitatively assessed with respect to a manual annotator. Although these imaging conditions were different with respect to the training dataset, relatively good DSCs were found; 0.79, 0.89, 0.85, and 0.88 for Figure 5a-d respectively.

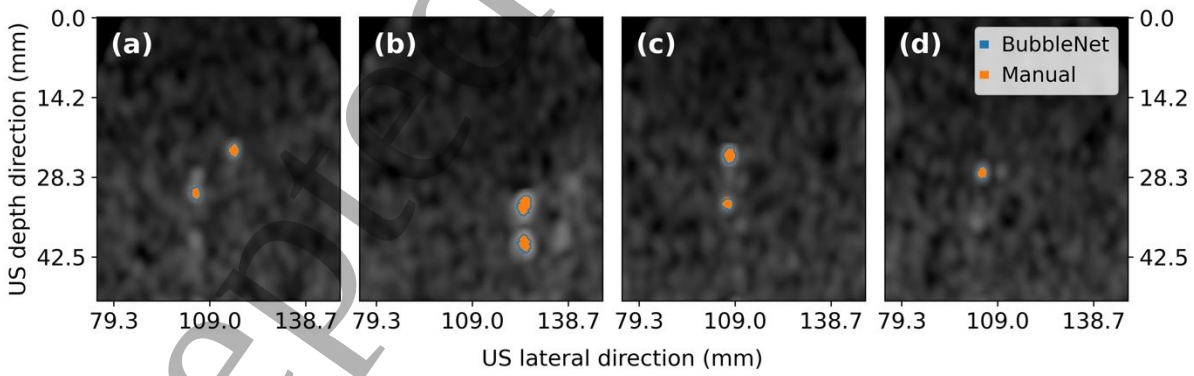

Figure 5. A comparison in segmentation performance on four (a-d) randomly picked differential Ultrasound frames between BubbleNet (blue solid line) and a manual annotator (orange segmentation map) for an unseen dataset with different acoustic imaging settings.

Quantitatively, the boxplots presented in Figure 6a indicate the DSC range of every batch in the 5-fold cross validation (n=5747), of which the median DSC is reported inside the box. The frame percentage identified as a statistical outlier was reported at the bottom left of the boxplot panel. An outlier suggests a serious segmentation mismatch between the manual segmentation and the segmentation

predicted by BubbleNet. E.g., a false negative or a false positive bubble localization. Here, the outlier percentage ranged between 3.5% and 6.5%. In Figure 6b-c, the average DSC (±1 SD) learning curves for training (Figure 6b) and validation (Figure 6c) are plotted against the number of epochs. For the representation of the five validation learning curves in Figure 6c, a decreased standard deviation and a lower degree of fluctuation is observed near the end of training (i.e., epoch 100), indicating a stabilized neural network for each of the individual batches. The number of manually segmented frames per bubble category is visualized in a horizontal bar chart (Figure 6d) and represents a total of 15050 localized bubbles on 5747 differential US frames. It should be noted that this bar chart does not include the 1536 manually segmented differential US frames to assess the proton range estimation capabilities, and the 432 manually segmented differential US frames to assess the performance on an unseen external dataset with challenging imaging parameters. A bar chart representation on the total number of manually segmented frames (7715 frames, =17989 bubbles) was added as Supplementary Material B. Additionally, in the intra-observer variability comparison, the median DSC (Quartile 1-Quartile 3) was determined to be 0.86 (0.79-0.95) on a duplicate segmentation dataset of 150 differential US frames. In this comparison, the minimum DSC was equal to 0.43 and the maximum DSC was equal to 0.99. Furthermore, the average computing time (± 1SD) required to train BubbleNet for 100 epochs was 201 ± 47 minutes, and the average BubbleNet prediction time (± 1SD) was 27.9 ± 0.6 milliseconds per US difference frame.

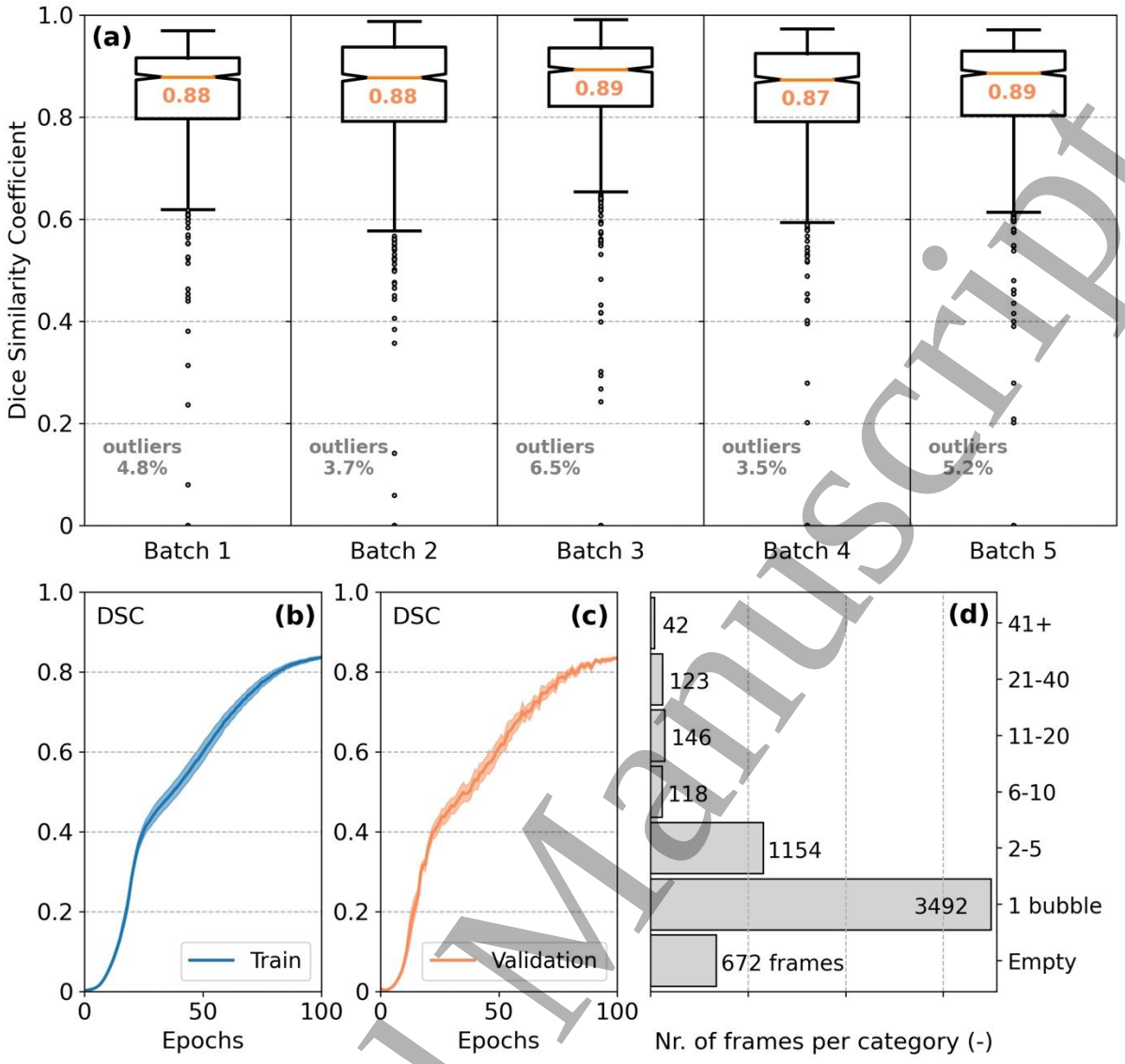

Figure 6. The boxplots in (a) show the Dice Similarity Coefficient (DSC) values from every batch covered by the 5-fold cross validation. The average DSC learning curves of BubbleNet (± 1 SD, standard deviation) are plotted against the number of epochs for training (b) and validation (c) data. The number of manually segmented frames in the training and cross-validation dataset is visualized in a horizontal bar chart per bubble category (d).

3.2. Range verification in proton therapy

The range of a 154 MeV proton beam was measured using three different bubble localization methods with a 0.8 mm bin width (Figure 7): BubbleNet, manual bubble segmentation, and intensity-weighted thresholding (Collado-Lara *et al.*, 2022). The performance for each of these methods was directly compared to a reference dose distribution that was obtained via a verified analytical model of the proton Bragg peak (Bortfeld, 1997) in our experimental conditions. The 95% confidence interval extracted from the analytical Bragg peak model (± 0.23 mm) was superior to the reported interpolated resolution of the QubeNext multi-layer ionization chamber in water of ± 0.50 mm. A similar accuracy offset (Δx) of -0.8 mm was found for all three methods with respect to R_{80}. Nevertheless, better bubble count statistics were observed in the BubbleNet-based proton range (n=1966 bubbles) compared to the previously adopted intensity thresholding-based localization

method (n=1520 bubbles), while manually, 1841 bubbles were segmented. The range determination method was repeated for two other irradiation setups for which the manual segmentations were lacking, and with a lower original nanodroplet concentration (20µM instead of 31 µM). Here, the BubbleNet accuracy offset compared to the analytical model was -0.8 mm and -0.9 mm, while the accuracy of the intensity-weighted thresholding method was -0.8 mm and -1.0 mm.

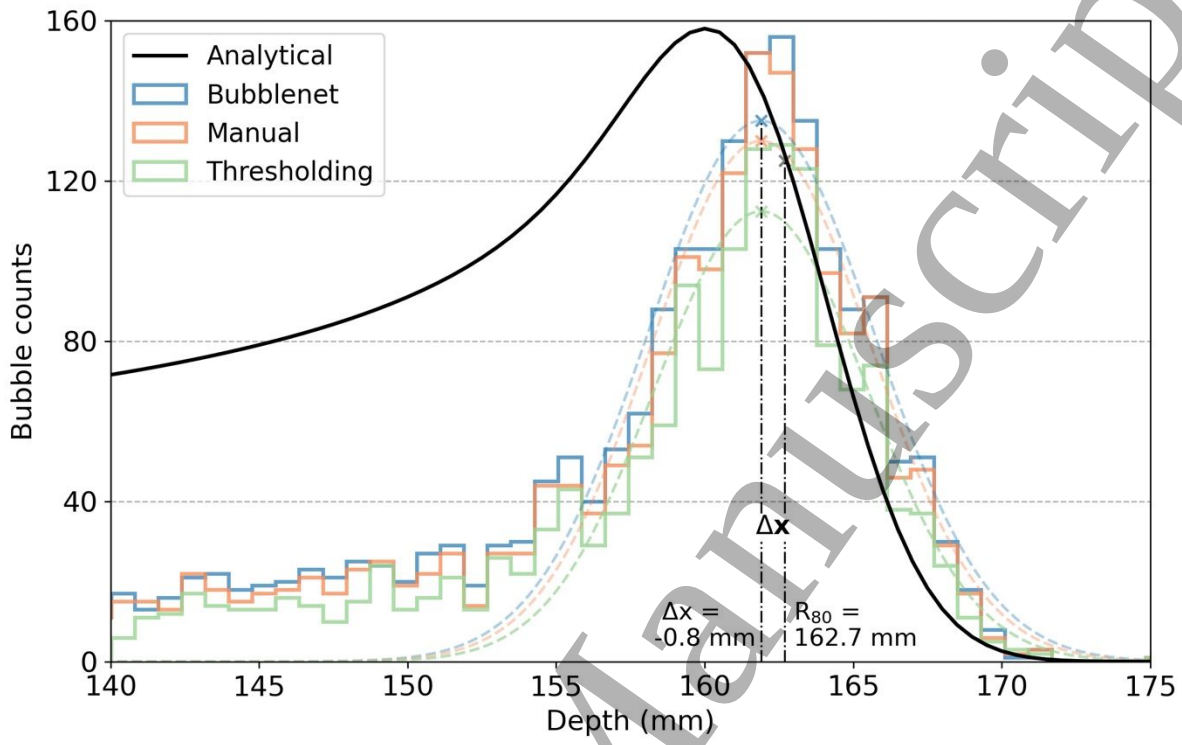

Figure 7. The 154 MeV proton range determination with BubbleNet (blue), a manual segmentation of the individual frames (orange), and by simple dataset-specific bubble intensity thresholding (green).

3.3. Lateral dose profile assessment in photon therapy

The two-dimensional intensity distribution (0.8 x 0.8 mm^2 bins) of vaporized nanodroplets was evaluated for a 6MV photon irradiation in a Carbopol phantom, whose dimension is indicated by the dashed black box in Figure 8a-b. Identical window levelling parameters are applied in Figure 8a-b after normalization to unity, and in Figure 8c-d, the normalized cumulative red isodose lines in the lateral and depth dimension (extracted from the treatment planning system) serve as reference in the comparison between BubbleNet and the intensity-weighted thresholding method (Figure 8c-d). BubbleNet was able to detect 27% more bubbles (n=15685) that were kept unnoticed by the intensity-parameter sensitive thresholding algorithm (n=12339). The elevated bubble count gives rise to smoother cumulative dose histograms in both directions, although the biggest advances by BubbleNet are noticeable in the US depth dimension (Figure 8d).

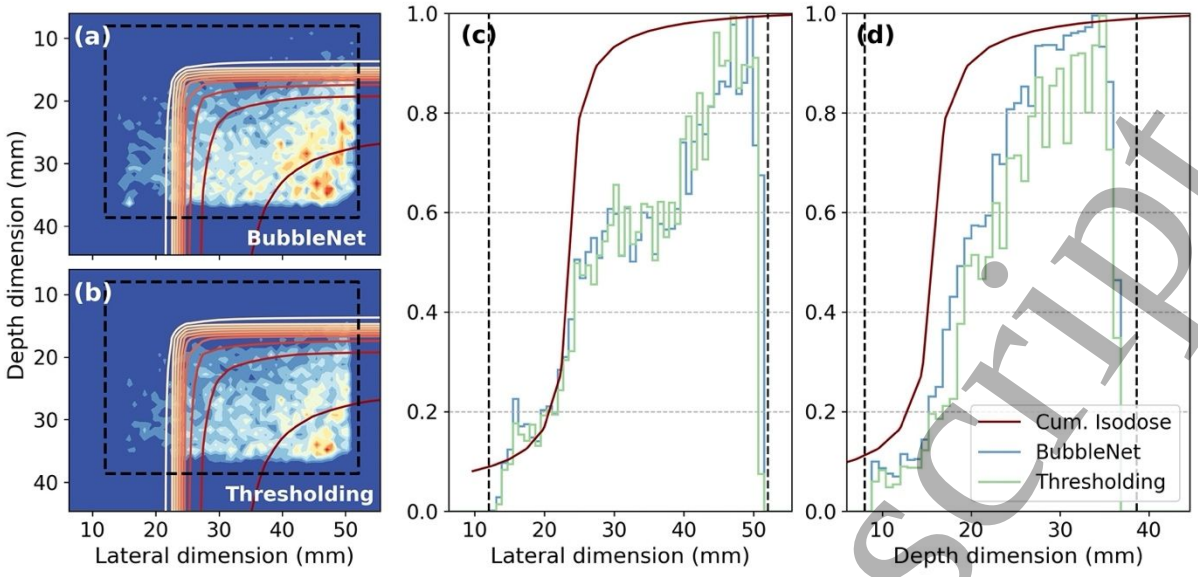

Figure 8. The estimated dose profiles, including the cumulative isodose lines (dark red), after unity normalization by BubbleNet (blue) and by intensity-weighted bubble thresholding (green) in a Carbopol phantom that was irradiated with 6MV photons.

4. Discussion

A widely applicable deep convolutional neural network (BubbleNet, Figure 2) was developed to detect vaporized nanodroplets on differential US frames, eliminating the need of time-consuming manual bubble identification methods or subjective dataset-specific intensity-weighted bubble thresholding. A 5-fold cross validation (Figure 6) was completed to determine the model performance in the form of an overlap metric (i.e., DSC). An overall good agreement was found between the segmentation performance of the left-out batches (median DSC = 0.88) which showed a similar performance with a slightly larger degree of deviation, compared to the intra-observer variability measured on 150 differential US frames (median DSC = 0.86). The complete dataset was composed and evaluated on differential US frames with a low- and high number of bubbles per frame. Furthermore, the generalization abilities of BubbleNet were evaluated on an unseen test dataset acquired with a lower frequency ultrasound array and with larger pixel dimensions (283 x 495 μm^2) with respect to the training dataset (71 x 250 μm^2; 85 x 250 μm^2; 86 x 99 μm^2) (Supplementary Material A).

Next, the segmentation capabilities of BubbleNet were explored to assess the 80% distal dose fall off in a proton treatment beam (Figure 7), and to qualitatively visualize the dose distribution of a 6MV photon beam in our phantom (Figure 8). Like the other evaluated methods, BubbleNet was able to measure the proton range of a 154 MeV beam with a -0.8 mm accuracy. Nevertheless, 29% more bubbles were detected by BubbleNet, compared to the more conventional intensity-weighted thresholding method (Figure 8). BubbleNet's increased detection efficiency was also confirmed for the 7.5 MHz US image dataset acquired of the Carbopol phantom (27%), which was irradiated with 6MV photons (Figure 8). An increased bubble count with respect to other available methods is relevant because a raised number of identified vaporizations in the irradiated US imaging volume indirectly leads to a better precision in the localization of the dose distribution. If a lower bubble count was previously sufficient to derive solid and scientific conclusions, BubbleNet could possibly trigger a shift

towards the injection of lower nanodroplet concentrations. Moreover, BubbleNet can ignore US image regions subject to high-frequency artifacts (e.g., Figure 4).

Additionally, it is relevant to note that the value picked for intensity-weighted thresholding is a very sensitive parameter to be tuned, and moreover alternating with respect to several imaging parameters. A threshold value picked too high fails to detect vaporized nanodroplets (false negative) that show up as a low intensity US signal, and a value picked too low is likely to identify high-frequency image artifacts as a false positive vaporized nanodroplet. To illustrate the sensitivity of the intensity-weighted thresholding method in one dataset (e.g., Phantom 8, Supplementary Material A); a 5% lower threshold value gave a 55.5% increase in bubble counts, while a 5% higher threshold value gave a 43.6% decrease in bubble counts, which substantially complicates the future standardization of nanodroplet based dose assessment in radiotherapy. Additionally, the US beam is attenuated in the US depth dimension, which directly complicates the manual procedure of optimizing one single thresholding value. It should also be noted that the manual threshold value parameter tuning requires staff time, which would be avoided by BubbleNet.

Furthermore, it is important to highlight that the conversion from bubble counts to radiation dose is not a simple one-to-one conversion. In proton irradiations, a LET threshold must be exceeded before droplet evaporation could be triggered. At 50°C, the LET threshold was reported to be 60 keV/µm by (Carlier *et al.*, 2020). This threshold is thereby typically reached by recoil nuclei, alpha particles, and by the primary protons at the very end of their particle track. To illustrate, in Figure 7, the plateau region in vaporization events before 155 mm depth could typically be originating from rare high-LET nuclear interactions and the generation of high-LET recoil-nuclei, while at the end of the proton range (between 155 mm and 170 mm depth), a vaporization peak is observed because the LET of the primary protons exceeds the vaporization threshold.

BubbleNet is composed of a U-Net ++ L^2 network architecture to detect microbubbles on differential US-image. Previously, in (Park *et al.*, 2020), a U-Net based architecture was integrated to detect microbubbles on high-resolution US images to achieve diagnostic imaging of the microvasculature system, and additionally, the vasculature blood flow. (Liu *et al.*, 2020) employed a multi-layer neural network to localize microbubbles in the vasculature with US microscopy, and (van Sloun *et al.*, 2021) presented a patch-based fully connected convolutional neural network to localize individual microbubbles in vascular architectures without incorporating structual priors. In all referred works, the neural network was trained on synthetic data, and evaluated in variable US imaging conditions and bubble concentrations. Although the training of convolutional neural networks on synthetic data implies an advantageous ground truth of the bubble positioning, it ignores the indirect uncertainties that come with experimental US due to imaging artifacts such as reverberation or speckle. Moreover, recently published research is mainly focused on the super-resolution imaging of the vasculature with US contrast agents, while this work aims to contribute towards the technological development of an *in vivo* dosimeter for radiotherapy with radiosensitive nanodroplets and high frame rate US imaging.

The development of a widely deployable online *in vivo* dosimeter is still a persistent challenge in radiotherapy. Although some promising full-scale prototypes for proton range verification exist such as post-treatment PET imaging or prompt-gamma imaging, these methods regrettably fail to be settled in the conventional hadron therapy workflow due to the extra operational and equipment costs. Alternatively, US imaging of radiation-induced vaporization events of nanodroplets seems a plausible and cheaper dosimetry solution. However, some technological limitations remain to be resolved. One of these limitations was the detection of vaporized nanodroplets on a diversely composed dataset. US imaging combined with super-resolution techniques enable to obtain high-resolution images (e.g., 100 x 100 μm^2), even for deep seated target volumes, and although the active imaging field-of-view will be spatially limited, such resolution is substantially better than the more recent treatment verification solutions such as the reconstructed 3D dose-grid by portal dosimetry (e.g. 3x3 mm^3) in photon therapy (van Elmpt *et al.*, 2008), or the post-treatment PET (e.g. 2 mm) (Knopf *et al.*, 2011) or prompt-gamma imaging dosimetry solutions in proton therapy (e.g. 1-3 mm) (Hueso-Gonzalez *et al.*, 2018). This also gives room to speculate that *in vivo* dosimetry with radiosensitive nanodroplets may enter an application field to narrow down the existing measurement uncertainties currently obtained in small-field dosimetry (Palmans *et al.*, 2018; Andreo, 2018). Also, the resolution of dose verification techniques is sensitive to motion, especially over the full duration of a radiotherapy treatment fraction. In motion-sensitive anatomical treatment volumes, 2D deformable image registration could be required to calculate artifact-free differential US images, however, US imaging of the clinical target volume could also indirectly enable the possibility to track the target motion or even underlying soft tissue deformations in real-time. One of the other major advantages of *in vivo* dosimetry with nanodroplets are the ultra-high frame rates that can be achieved in US imaging. For instance, with the growing number of publications indicating the advantage of ultra-high dose rate radiotherapy (FLASH) for sparing healthy tissues (Esplen *et al.*, 2020), the high spatial (<100µm) and temporal (>1000 fps) resolution of US imaging might be able to monitor the deposited treatment dose, free of detector saturation, with ultra-high dose rates going up to 40 Gy per second.

5. Conclusion

A deep convolutional neural network (BubbleNet) was developed and evaluated in this work to localize vaporized nanodroplets on differential US images for dose assessments in radiotherapy applications. The trained BubbleNet network showed a good segmentation performance (median DSC=0.88) on a collection of differential US images that were acquired with a varying degree of imaging and bubble parameters. Approximately 30% better bubble count statistics were obtained with BubbleNet compared to the conventional intensity-weighted thresholding for nanodroplet localization on differential US images after proton and photon irradiation. Furthermore, BubbleNet would avoid the time-consuming process of adequately tuning the intensity threshold value for each US imaging dataset.

Acknowledgements

This work has been supported by the European Union's Horizon 2020 research and innovation program under grant agreement n°766456 ("AMPHORA"). Co-author B.C. received a PhD fellowship fundamental research from the Research Foundation Flanders (n°11A9520N). Dr. Laurence Delombaerde and Dr. Marta Rovituso are acknowledged for their help during the experimental work.

Conflict of Interest

The authors have no relevant conflict of interest to disclose.

References

Agazaryan N, Solberg T D and DeMarco J J 2003 Patient specific quality assurance for the delivery of intensity modulated radiotherapy *J Appl Clin Med Phys* **4** 40-50

Andreo P 2018 The physics of small megavoltage photon beam dosimetry *Radiother Oncol* **126** 205-13

Arjomandy B, Taylor P, Ainsley C, Safai S, Sahoo N, Pankuch M, Farr J B, Yong Park S*, et al.* 2019 AAPM task group 224: Comprehensive proton therapy machine quality assurance *Med Phys* **46** e678-e705

Atun R, Jaffray D A, Barton M B, Bray F, Baumann M, Vikram B, Hanna T P, Knaul F M*, et al.* 2015 Expanding global access to radiotherapy *Lancet Oncol* **16** 1153-86

Barton M B, Jacob S, Shafiq J, Wong K, Thompson S R, Hanna T P and Delaney G P 2014 Estimating the demand for radiotherapy from the evidence: a review of changes from 2003 to 2012 *Radiother Oncol* **112** 140-4

Baskar R, Lee K A, Yeo R and Yeoh K W 2012 Cancer and radiation therapy: current advances and future directions *Int J Med Sci* **9** 193-9

Bellec J, Delaby N, Jouyaux F, Perdrieux M, Bouvier J, Sorel S, Henry O and Lafond C 2017 Plan delivery quality assurance for CyberKnife: Statistical process control analysis of 350 film-based patient-specific QAs *Phys Med* **39** 50-8

Bissonnette J P, Balter P A, Dong L, Langen K M, Lovelock D M, Miften M, Moseley D J, Pouliot J*, et al.* 2012 Quality assurance for image-guided radiation therapy utilizing CT-based technologies: a report of the AAPM TG-179 *Med Phys* **39** 1946-63

Borras J M, Barton M, Grau C, Corral J, Verhoeven R, Lemmens V, van Eycken L, Henau K*, et al.* 2015a The impact of cancer incidence and stage on optimal utilization of radiotherapy: Methodology of a population based analysis by the ESTRO-HERO project *Radiother Oncol* **116** 45-50

Borras J M, Lievens Y, Dunscombe P, Coffey M, Malicki J, Corral J, Gasparotto C, Defourny N*, et al.* 2015b The optimal utilization proportion of external beam radiotherapy in European countries: An ESTRO-HERO analysis *Radiother Oncol* **116** 38-44

Bortfeld T 1997 An analytical approximation of the Bragg curve for therapeutic proton beams *Med Phys* **24** 2024-33

Brandner E D, Chetty I J, Giaddui T G, Xiao Y and Huq M S 2017 Motion management strategies and technical issues associated with stereotactic

body radiotherapy of thoracic and upper abdominal tumors: A review from NRG oncology *Med Phys* **44** 2595-612

Carlier B, Heymans S V, Nooijens S, Toumia Y, Ingram M, Paradossi G, D'Agostino E, Himmelreich U*, et al.* 2020 Proton range verification with ultrasound imaging using injectable radiation sensitive nanodroplets: a feasibility study *Phys Med Biol* **65** 065013

Chollet F 2015 Keras. GitHub repository)

Collado-Lara G, Heymans S V, Godart J, D'Agostino E, D'Hooge J, Van Den Abeele K, Vos H J and de Jong N 2021 Effect of a Radiotherapeutic Megavoltage Beam on Ultrasound Contrast Agents *Ultrasound Med Biol* **47** 1857-67

Collado-Lara G, Heymans S V, Rovituso M, Carlier B, Toumia Y, Verweij M, Paradossi G, Sterpin E*, et al.* 2022 Spatiotemporal Distribution of Nanodroplet Vaporization in a Proton Beam Using Real-Time Ultrasound Imaging for Range Verification *Ultrasound Med Biol* **48** 149-56

Esplen N, Mendonca M S and Bazalova-Carter M 2020 Physics and biology of ultrahigh dose-rate (FLASH) radiotherapy: a topical review *Phys Med Biol* **65** 23TR03

Heymans S V, Carlier B, Toumia Y, Nooijens S, Ingram M, Giammanco A, d'Agostino E, Crijns W*, et al.* 2021 Modulating ultrasound contrast generation from injectable nanodroplets for proton range verification by varying the degree of superheat *Med Phys* **48** 1983-95

Hueso-Gonzalez F, Rabe M, Ruggieri T A, Bortfeld T and Verburg J M 2018 A full-scale clinical prototype for proton range verification using prompt gamma-ray spectroscopy *Phys Med Biol* **63** 185019

Kellnberger S, Assmann W, Lehrack S, Reinhardt S, Thirolf P, Queiros D, Sergiadis G, Dollinger G*, et al.* 2016 Ionoacoustic tomography of the proton Bragg peak in combination with ultrasound and optoacoustic imaging *Sci Rep* **6** 29305

Kingma D and Lei J 2015 Adam: A method for stochastic optimization *arXiv*

Knopf A C, Parodi K, Paganetti H, Bortfeld T, Daartz J, Engelsman M, Liebsch N and Shih H 2011 Accuracy of proton beam range verification using post-treatment positron emission tomography/computed tomography as function of treatment site *Int J Radiat Oncol Biol Phys* **79** 297-304

Kwint M, Conijn S, Schaake E, Knegjens J, Rossi M, Remeijer P, Sonke J J and Belderbos J 2014 Intra thoracic anatomical changes in lung cancer patients during the course of radiotherapy *Radiother Oncol* **113** 392-7

Li Y and Hsi W C 2017 Analysis of measurement deviations for the patient-specific quality assurance using intensity-modulated spot-scanning particle beams *Phys Med Biol* **62** 2675-93

Liu X, Zhou T, Lu M, Yang Y, He Q and Luo J 2020 Deep Learning for Ultrasound Localization Microscopy *IEEE Trans Med Imaging* **39** 3064-78

Mijnheer B, Beddar S, Izewska J and Reft C 2013 In vivo dosimetry in external beam radiotherapy *Med Phys* **40** 070903

Noble D J, Yeap P L, Seah S Y K, Harrison K, Shelley L E A, Romanchikova M, Bates A M, Zheng Y*, et al.* 2019 Anatomical change during radiotherapy for head and neck cancer, and its effect on delivered dose to the spinal cord *Radiother Oncol* **130** 32-8

Paganetti H 2012 Range uncertainties in proton therapy and the role of Monte Carlo simulations *Phys Med Biol* **57** R99-117

Palmans H, Andreo P, Huq M S, Seuntjens J, Christaki K E and Meghzifene A 2018 Dosimetry of small static fields used in external photon beam radiotherapy: Summary of TRS-483, the IAEA-AAPM international Code of Practice for reference and relative dose determination *Med Phys* **45** e1123-e45

Park J H, Choi W, Yoon G Y and Lee S J 2020 Deep Learning-Based Super-resolution Ultrasound Speckle Tracking Velocimetry *Ultrasound Med Biol* **46** 598-609

Parodi K, Paganetti H, Shih H A, Michaud S, Loeffler J S, DeLaney T F, Liebsch N J, Munzenrider J E*, et al.* 2007 Patient study of in vivo verification of beam delivery and range, using positron emission tomography and computed tomography imaging after proton therapy *Int J Radiat Oncol Biol Phys* **68** 920-34

Richter C, Pausch G, Barczyk S, Priegnitz M, Keitz I, Thiele J, Smeets J, Stappen F V*, et al.* 2016 First clinical application of a prompt gamma based in vivo proton range verification system *Radiother Oncol* **118** 232-7

Sonke J J, Aznar M and Rasch C 2019 Adaptive Radiotherapy for Anatomical Changes *Semin Radiat Oncol* **29** 245-57

Toumia Y, Miceli R, Domenici F, Heymans S V, Carlier B, Cociorb M, Oddo L, Rossi P*, et al.* 2021 Ultrasound-assisted investigation of photon triggered vaporization of poly(vinylalcohol) phase-change nanodroplets: A preliminary concept study with dosimetry perspective *Phys Med* **89** 232-42

van Elmpt W, McDermott L, Nijsten S, Wendling M, Lambin P and Mijnheer B 2008 A literature review of electronic portal imaging for radiotherapy dosimetry *Radiother Oncol* **88** 289-309

van Sloun R J G, Solomon O, Bruce M, Khaing Z Z, Wijkstra H, Eldar Y C and Mischi M 2021 Super-Resolution Ultrasound Localization Microscopy Through Deep Learning *IEEE Trans Med Imaging* **40** 829-39

Verhaegen F, Fonseca G P, Johansen J G, Beaulieu L, Beddar S, Greer P, Jornet N, Kertzscher G*, et al.* 2020 c *Phys Imaging Radiat Oncol* **16** 18-9

Wohlfahrt P and Richter C 2020 Status and innovations in pre-treatment CT imaging for proton therapy *Br J Radiol* **93** 20190590

Yeung T K, Bortolotto K, Cosby S, Hoar M and Lederer E 2005 Quality assurance in radiotherapy: evaluation of errors and incidents recorded over a 10 year period *Radiother Oncol* **74** 283-91

Zhou Z, Siddiquee M M R, Tajbakhsh N and Liang J 2020 UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation *IEEE Trans Med Imaging* **39** 1856-67