# End-to-end forecasting of needle trajectory in percutaneous ablation

Mark A. Pinnock, Yipeng Hu, Steve Bandula and Dean C. Barratt

## **1. DESCRIPTION OF PURPOSE**

Percutaneous cryoablation is an image-guided treatment option that uses extreme cold to cause cell death by the formation of an ice ball around needles inserted in and around a tumour.<sup>1</sup> Convolutional neural networks (CNNs) have been proposed to aid accurate guidance in a number of percutaneous techniques, but so far have required additional post-processing steps to do so.<sup>2–4</sup> Estimating uncertainty is of critical importance in medical applications that forecast future events, where the consequences of a neural network that gives an incorrect prediction on data that is noisy or not previously encountered may be catastrophic. Dropout<sup>5</sup> has previously been explored as a technique for modelling uncertainty in deep learning, by serving as a basis for Monte Carlo dropout<sup>6</sup> (MCDO).

Our aim is to develop and assess an end-to-end, CNN-based technique for forecasting, with the following contributions: 1) We utilise the 3D U-Net<sup>7</sup> to directly predict needle trajectory based on images acquired in interventional computed tomography (iCT) procedures. 2) We then show how needle forecasting is improved by an ensemble of models generated using MCDO to a greater degree than in other tasks such as segmentation. 3) Lastly, we estimate and visualise the uncertainty of the predictions.

### 2. METHODS

The 3D U-Net is adopted in this work for 3D iCT volumes. For experiments where dropout is used, the dropout layer was added after each convolutional and transpose convolutional layer, both for regularisation during training and for generating the ensemble required for forecasting and uncertainty visualisation.

MCDO was performed during inference with dropout probability  $p_d$  and number of forward passes T. This generated a series of predictions, with each prediction  $y_t^*$  representing voxel-wise class probabilities. The mean of these predictions was calculated by  $\hat{y}_{i,p} = \frac{1}{T} \sum_{t=1}^{T} y_t^*$  to yield the predicted trajectory map. For uncertainty visualisation, an uncertainty map  $\hat{h}$  was generated from the entropy of the class probability distribution at each voxel,<sup>8</sup> where for class c out of all classes C, the voxel-wise entropy is  $\hat{h} = -\sum_{c \in C} \hat{y}_c \ln[\hat{y}_c]$ .

The data consisted of renal iCT images with needle segmentation masks. Out of the 20 subjects, 6 were randomly chosen for testing purposes and the rest were used for training. Each subject had between 6 and 14 time-series volumes covering initial needle insertion to the time at which the needle tip is situated within the target tumour, for a total of 122 training examples and 41 testing examples. All images were normalised to lie on the range [0, 1] with 0 and 1 corresponding to the minimum and maximum intensities of the CT scanner. To train the U-Net to predict needle trajectory, only the final segmentation mask in each subject's time series was used as the label for all input volumes for a given subject.

The network was trained for 500 epochs, with minibatch size 4, data augmentation (rotation, flipping and scaling) and the Adam optimiser. The optimal learning rate was found to be  $10^{-3}$  after cross-validation. Dropout was employed during training along with batch normalisation.

The proposed method using MCDO is hereafter referred to as UNet-DI. For comparison, the same network was tested with dropout only in training (UNet-D) and with neither dropout in training nor in inference (UNet) using the following metrics. Overlap of the predicted segmentation mask with ground truth was assessed using SDC. The needle tip localisation error was defined as the 2D Euclidean distance between predicted needle tip and ground truth. To avoid bias during manual needle tip identification, the images were shuffled to ensure that the reviewer was blinded as to whether the mask was from UNet, UNet-D, UNet-DI or ground truth. Any unusable segmentations preventing unambiguous needle tip localisation or featuring gross distortions were marked as such. All predictions were thresholded at 0.5 prior to calculation. The predicted trajectories also underwent visual assessment with regards to accuracy and morphology.

### **3. RESULTS**

In terms of SDC, UNet-DI outperformed both UNet-D (0.48 vs 0.32, p = 0.0001) and UNet (0.48 vs 0.14, p < 0.0001) when compared with the Mann-Whitney U-Test. Only predictions considered adequate for use were used to calculate tip error (44% for UNet, 63% for UNet-D and 100% for UNet-DI). In terms of tip error, both UNet-DI (24 mm vs 82 mm, p < 0.0001) and UNet-D (25 mm vs 82 mm, p < 0.0001) outperformed UNet, but there was no statistically significant difference between UNet-DI and UNet-D.



Figure 1. Left and centre: UNet, UNet-D and UNet-DI unthresholded predictions for Subjects 1 and 2 respectively. Right: UNet-DI predicted trajectory before (top) and after (bottom) repositioning. White crosshairs indicate the final ground truth needle tip position.

In all cases, UNet segments only the current needle position instead of the trajectory (see Figure 1), and the mask is frequently distorted or ambiguous. In comparison, both UNet-D and UNet-DI are closer to the ground truth trajectory, more needle-like, and are also sensitive to the orientation of the needle. Figure 1 (Subject 3) shows how the trajectory predicted by UNet-DI shifts in a subsequent image in the series towards the target needle position after the needle is re-positioned. While both UNet-D and UNet-DI are capable of directly predicting the trajectory, the former suffers from more severely distorted needle morphology or incorrect segmentations as in Figure 2 (yellow circles). Likewise, misclassification and distortion by UNet-DI (red and yellow arrows) is far less severe than with UNet-D.

Figure 3 shows two examples of the trajectory and uncertainty maps generated by UNet-DI. The yellow and red circles demonstrate incorrect segmentations - the corresponding regions in the uncertainty map are bright, indicating high uncertainty. The use of MCDO thus allows us not only to perform forecasting but also to highlight uncertain and potentially erroneously segmented areas for review by an operator.

## 4. NEW OR BREAKTHROUGH WORK TO BE PRESENTATED

Dropout-enabled neural networks are proposed to forecast the trajectory of a needle in an end-to-end fashion, while also estimating the model's epistemic uncertainty to highlight potentially inaccurate predictions. Both dropout-trained networks outperformed the network without dropout. Needle tip localisation is not significantly improved by MCDO but, in cases where a distorted segmentation mask has been generated, ensembling the predictions does visually improve needle morphology, as confirmed by the improvement in SDC. The performance increase seen in this work is substantially greater than that seen in many works exploring MCDO in segmentation, which either found no improvement<sup>9, 10</sup> or incremental performance gains.<sup>11–14</sup> In addition, the generated uncertainty map highlights potentially incorrect or distorted predictions, allowing an operator to review predictions that may be of inadequate quality.

The promising results need to be interpreted within the context of several limitations. Only a relatively small number of the labelled data were available in this work, potentially reducing the model's generalisability. In addition, the predictive power of the model may have been limited by its relatively small size (approximately



Figure 2. UNet, UNet-D and UNet-DI unthresholded predictions for Subjects 4 and 6 respectively. Yellow circles indicate erroneous predictions seen in UNet-D but not UNet-DI, while arrows show erroneous or distorted predictions seen in UNet-DI that are less severe than UNet-D. White crosshairs indicate the final ground truth needle tip position.



Figure 3. Unthresholded predictions and uncertainty maps from UNet-DI. White crosshairs indicate the ground truth final needle position.

296,000 parameters). Lastly, MCDO estimates only epistemic uncertainty, whereas true variance of a model's predictions also comprises aleatoric uncertainty.

## 5. CONCLUSION

In summary, we have proposed a novel approach for training neural networks with dropout to allow an operator to visualise the predicted needle trajectory based on the current iCT volume containing needle orientation and location information. This is the first study to demonstrate the feasibility of end-to-end neural networks to predict future needle locations in interventional applications. The experiment results suggest that using an ensemble of predictions significantly improves the accuracy of the predicted trajectory, importantly, to the extent that it rectified a number of cases where the predictions without MCDO might be considered clinically uninformative, and also gives the operator an indication as to the uncertainty in the underlying forecast.

## 6. OTHER SUBMISSIONS

This work has not been submitted for publication elsewhere.

#### REFERENCES

- Uppot, R. N., Silverman, S. G., Zagoria, R. J., Tuncali, K., Childs, D. D., and Gervais, D. A., "Imaging-Guided Percutaneous Ablation of Renal Cell Carcinoma: A Primer of How We Do It," *American Journal of Roentgenology* 192, 1558–1570 (Jun 2009).
- [2] Jung, H., Gonzalez, Y., Shen, C., Klages, P., Albuquerque, K., and Jia, X., "Deep-learning-assisted automatic digitization of applicators in 3D CT image-based high-dose-rate brachytherapy of gynecological cancer," *Brachytherapy* 18, 841–851 (Nov 2019).
- [3] Lee, J. Y., Islam, M., Woh, J. R., Washeem, T. S., Ngoh, L. Y. C., Wong, W. K., and Ren, H., "Ultrasound needle segmentation and trajectory prediction using excitation network," *International Journal of Computer* Assisted Radiology and Surgery 15, 437–443 (Jan 2020).
- [4] Mwikirize, C., Nosher, J. L., and Hacihaliloglu, I., "Convolution neural networks for real-time needle detection and localization in 2D ultrasound," *International Journal of Computer Assisted Radiology and* Surgery 13, 647–657 (May 2018).
- [5] Srivastava, N., Hinton, G., Krizhevsky, A., and Salakhutdinov, R., "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research* 15, 1929–1958 (2014).
- [6] Gal, Y. and Ghahramani, Z., "Bayesian Convolutional Neural Networks with Bernoulli Approximate Variational Inference," arXiv preprint (Jun 2015).
- [7] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O., "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," in [International Conference on Medical Image Computing and Computer-Assisted Intervention], 424–432, Springer (2016).
- [8] Smith, L. and Gal, Y., "Understanding Measures of Uncertainty for Adversarial Example Detection," arXiv preprint (2018).
- [9] Jensen, M. H., Jørgensen, D. R., Jalaboi, R., Hansen, M. E., and Olsen, M. A., "Improving uncertainty estimation in convolutional neural networks using inter-rater agreement," in [International Conference on Medical Image Computing and Computer-Assisted Intervention], 540–548, Springer (2019).
- [10] Mehrtash, A., Wells, W. M., Tempany, C. M., Abolmaesumi, P., and Kapur, T., "Confidence Calibration and Predictive Uncertainty Estimation for Deep Medical Image Segmentation," *IEEE Transactions on Medical Imaging* (Nov 2020).
- [11] Sedai, S., Antony, B., Mahapatra, D., and Garnavi, R., "Joint Segmentation and Uncertainty Visualization of Retinal Layers in Optical Coherence Tomography Images Using Bayesian Deep Learning," in [Computational Pathology and Ophthalmic Medical Image Analysis], 219–227, Springer (2018).
- [12] Jungo, A. and Reyes, M., "Assessing Reliability and Challenges of Uncertainty Estimations for Medical Image Segmentation," in [International Conference on Medical Image Computing and Computer-Assisted Intervention], 48–56, Springer (2019).
- [13] Wickstrøm, K., Kampffmeyer, M., and Jenssen, R., "Uncertainty modeling and interpretability in convolutional neural networks for polyp segmentation," in [*IEEE International Workshop on Machine Learning for* Signal Processing (MLSP)], IEEE (2018).
- [14] Roy, A. G., Conjeti, S., Navab, N., and Wachinger, C., "Bayesian QuickNAT: Model uncertainty in deep whole-brain segmentation for structure-wise quality control," *NeuroImage* 195, 11–22 (Jul 2019).