A Deep Learning Surrogate Model for Microwave Thermal Ablation: From Preoperative Planning to Intraoperative Re-planning

Ilias Nahmed[1], Francesco Dettori[1], Michel Duprez[1], Pablo Alvarez[1], and Stéphane Cotin[1]

Université de Strasbourg, CNRS, Inria, ICube, F-67000 Strasbourg, France
ilias.nahmed@inria.fr

Abstract. With the increasing interest in thermal ablation therapies, patient-specific predictive simulations have become crucial for improving both preoperative planning and intraoperative corrections. Planning strategies typically rely on the solution of an inverse problem, yet high-fidelity finite element simulations remain computationally prohibitive when embedded within an optimization loop under strict clinical time constraints. In this work, we introduce a real-time deep learning surrogate model for predicting liver microwave ablation outcomes, trained exclusively on synthetic data and personalized through 3D perfusion maps that encode patient anatomy and local perfusion rates. This surrogate model takes as input the antenna position, power and ablation duration, and outputs the predicted necrosis volume within a few milliseconds. Once integrated into an optimization loop, it becomes possible to propose antenna trajectories and ablation settings that maximize tumor coverage while minimizing damage to healthy tissue. We demonstrate this on clinical data via a physics-based optimization framework that computes optimal ablation parameters in under 2 minutes, critical for preoperative planning or intraoperative re-planning in microwave ablation. The surrogate model was validated on *in vivo* pre-clinical data where it achieved an average Dice coefficient of 0.82 for necrosis mask prediction.

Keywords: thermal ablation · simulation · planning · surrogate model

1 Introduction

Over the last decades, percutaneous thermal ablation has emerged as a primary minimally invasive alternative to surgical resection [11]. While Radiofrequency Ablation (RFA) has long been the dominant modality, Microwave Ablation (MWA) is increasingly preferred for large tumors due to its ability to generate higher temperatures and larger ablation zones within shorter treatment times [16]. Ablation outcomes in MWA depend on both intervention-dependent factors (*e.g.* microwave generator power, ablation duration and antenna geometry) and patient-specific variables (*e.g.* tumor proximity to vessels, dielectric and

thermal properties). Yet, in practice, clinicians plan MWA parameters following manufacturer-provided charts, which ignore patient-dependent factors and therefore often overestimate ablation sizes. This is particularly important when ablating highly-perfused organs such as the liver or kidney [18,6], where nearby vessels cause temperature loss via advection, *i.e.* the heat-sink effect.

To mitigate these limitations, computational models have been proposed to simulate ablation outcomes [14,2]. Although these models demonstrate superior accuracy over geometric approximations [5,10], their clinical utility is hindered by the prohibitive computational cost of the numerical schemes (*e.g.* finite element method (FEM)) used to solve the underlying partial differential equations (PDEs). This limitation gets further amplified in MWA planning or intraoperative guidance scenarios, where multiple simulations (potentially hundreds) are required for solving the associated inverse problems. Consequently, there is a need for efficient strategies that approximate these outcomes in the shortest possible amount of time.

Recently, surrogate models based on physics-guided neural networks have emerged as a computationally efficient solution, enabling lesion-extent prediction in milliseconds. Such developments include deep networks based on U-Net architectures [13], neural cellular automata [8], and DeepONet-style operator learning [9]. However, these existing frameworks focus on RFA applications, where power control is not possible, and optimization is essentially restricted to probe placement. In contrast, MWA outcomes are strongly sensitive to ablation power and duration [12], requiring these additional control variables to be optimized for improved conformity between the predicted ablation zone and the clinical target. Current methods also often neglect patient-specific vascularization, which influence has been identified as key in the heat-sink effect [4].

In this work, we propose a deep learning surrogate model for liver MWA outcome prediction. The model is conditioned on both intervention- and patient-specific parameters, and outputs 3D thermal damage maps and ablation zones in milliseconds. We demonstrate its potential for inverse problem-solving in MWA through two clinical applications: (i) an end-to-end physics-based MWA planning solution and (ii) an intraoperative MWA re-planning solution for optimizing ablation parameters after antenna misplacement. To ensure generalization without requiring vast clinical datasets, the network is trained exclusively on simulated MW ablations performed on synthetic anatomies, including tumors and vascular structures. To our knowledge, this is the first study to provide a real-time deep learning surrogate for MWA outcome prediction, enabling optimization-based personalization within clinical time constraints.

2 Methods

2.1 Computational model for microwave ablation simulation

Heat distribution and thermal evolution during MWA is typically modeled using Pennes' Bioheat equation [14]. Disregarding metabolic effects, it reads

$$\rho C \frac{\partial T}{\partial t} = \nabla \cdot (\kappa \nabla T) - \rho_b C_b \omega_b (T - T_b) + Q_{\text{ext}}, \tag{1}$$

where ρ, C, T, ρ_b, C_b and T_b denote the tissue density (kg/m^3), specific heat capacity (J/kg/K) and temperature (K) of the tissue, and for the blood entering the tissue respectively. κ represents the thermal conductivity (W/m/K), ω_b is the blood perfusion rate (1/s) that allows modeling the heat sink effect and Q_{ext} is the external heat source.

The external heat source Q_{ext} represents the thermal energy absorbed by the tissue during MWA. It is related to the electric field intensity $\boldsymbol{E}$ (V/m) of the propagating microwaves and the tissue's electric conductivity σ (S/m) by:

$$Q_{ext} = \frac{1}{2}\sigma \langle \boldsymbol{E}, \boldsymbol{E}^* \rangle. \tag{2}$$

The electric field intensity $\boldsymbol{E}$ is the solution to the microwave propagation problem described by the frequency domain Maxwell equations:

$$\nabla \times \mu_r^{-1} (\nabla \times \boldsymbol{E}) - k_0^2 (\varepsilon_r - j \frac{\sigma}{\omega \varepsilon_0}) \boldsymbol{E} = 0, \tag{3}$$

where ω is the angular frequency (rad/s), μ_r is the relative permeability, ε_r is the relative permittivity, ε_0 is the permittivity in vacuum (F/m), and k_0 is the propagation constant in vacuum (1/m).

We solved the coupled system of PDEs using a FE numerical scheme implemented on FEniCSx [1], with boundary conditions appropriate for MWA and a computational reduction strategy described in our previous work [10]. Finally, we use the Arrhenius model to compute tissue damage and necrosis maps.

2.2 Deep learning surrogate model

We introduce a supervised deep learning surrogate model to achieve clinically compatible computation times. Instead of adopting a physics-informed neural network (PINN) formulation where the governing PDEs are embedded directly into the loss, we rely on a fully supervised approach. Indeed, the coupled electromagnetic–thermal problem described above would make a PINN approach particularly intricate and computationally expensive. By contrast, once trained, our network can be deployed without per-case retraining.

Inputs/Outputs: The network processes both spatial and procedural inputs, consisting of a 3D volumetric perfusion map encoding both anatomy and voxel wise perfusion, along with an oriented 3D Gaussian heatmap representing the antenna tip and insertion axis. Since the Pennes bioheat equation is highly

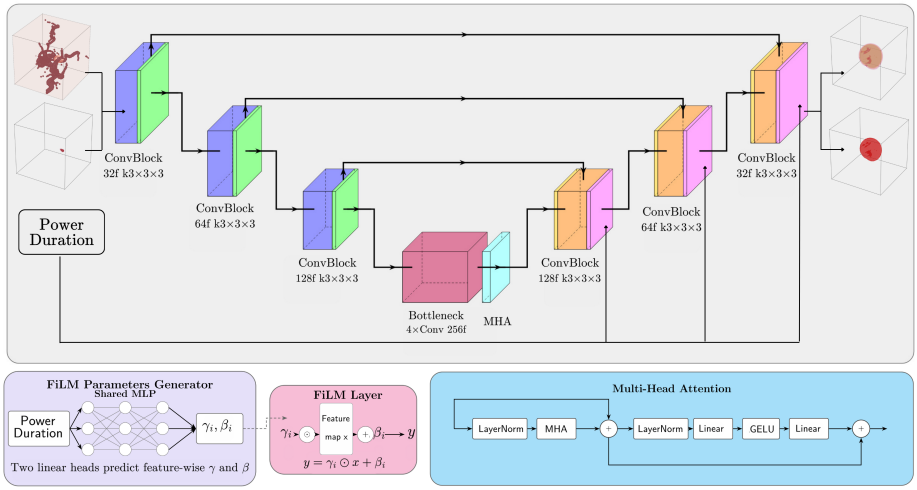

Fig. 1. Overview of the proposed 3D U-Net architecture with FiLM and MHA blocks.

sensitive to the perfusion parameter [4], explicitly incorporating it into the input is essential for accurate predictions. The network jointly predicts a 3D thermal damage map and a binary necrosis mask.

Architecture: The network is based on a 3D U-Net with feature-wise linear modulation (FiLM) layers and a multi-head self-attention layer (Fig. 1). The encoder consists of three downsampling stages (3^3 convolutions, group normalization and leaky ReLU activations) with channels progressively increasing from 32 to 256. The decoder mirrors this structure, using transposed convolutions for upsampling and skip connections. A bottleneck of four convolutional blocks (256 channels) is followed by a multi-head self-attention block to capture long-range spatial interactions. The procedural parameters (power and duration) are injected at each decoder level via small MLPs that map these inputs to scale and bias vectors modulating the FiLM layers. Finally, a 1^3 convolution layer produces the two output volumes.

Loss function: The network is supervised on both thermal injury and ablation zone maps through the following loss function:

$$\mathcal{L} = w_1 \mathrm{MSE}(D_{\text{target}}, D_{\text{pred}}) + w_2 \left(1 - \frac{2\,|M_{\text{target}} \cap M_{\text{pred}}|}{|M_{\text{target}} \cup M_{\text{pred}}|}\right) \tag{4}$$

where D_{target} and D_{pred} denote the target and predicted damage maps, M_{target} and M_{pred} denote the target and predicted ablation masks, and w_1, w_2 are empirically-tuned weighting factors. The supervision over the damage map complements the discrete necrosis map with smooth information of thermal injury. This is the rationale we used for keeping both quantities in the network's output.

2.3 Synthetic dataset generation

The network is trained only once on a large and diverse computer-generated dataset. By exposing the model to a sufficiently rich distribution of anatomical configurations, perfusion maps, and procedural parameters, we promote generalization across patients and clinical scenarios without requiring per-case retraining. This design enables fast inference at deployment while retaining the predictive fidelity of the underlying physics-based simulations.

The generated dataset consisted of 2400 samples derived from 1200 synthetic perfusion map volumes. For each sample, two randomly generated curved trajectories traversed the volume of interest to represent the portal vein and the vena cava. Additional branches were randomly generated from these primary trajectories to approximate hepatic vascular trees. Each vessel segment was dilated to radii sampled from a normal distribution based on physiological literature [3]. These vessels structures were assigned a perfusion value of 0.2 s^{-1} [4], while the remaining healthy tissue was assigned values ranging from 0.015 to 0.075 s^{-1}. Finally, for each volume, a tumor with irregular boundaries was generated and assigned a perfusion value 1.5 times of the healthy tissue perfusion [17].

MWA simulations were performed (see Sec. 2.1) with and without tumors, with antenna configurations sampled in the vicinity of the tumor center. Ablation parameters were varied to avoid the need for per-patient or per-procedure retraining, with power levels ranging from 30 to 100 W and durations from 180 to 500 s. The dataset was split into 85% for training and 15% for validation.

2.4 Clinically-motivated inverse problem-solving

In clinical practice, the antenna insertion direction is chosen carefully to maximize ablation outcome and avoid critical structures. Before MWA, this choice relies either on clinical expertise and/or geometric planning tools to minimize insertion risk given the patient's anatomy [7]. During MWA, because of human error, there is no guarantee that the insertion direction will match the planned path and antenna misplacement can occur. In both cases, it is necessary to find optimal ablation parameters appropriate for the tumor under treatment, and given the antenna insertion direction. Here, we propose solutions to these two clinical contexts by efficiently solving the following minimization problem:

$$\underset{P,T,\boldsymbol{t}}{\arg\min}\ \ \lambda_1 \left(1 - \frac{|M_{\text{tumor}} \cap A_{\text{pred}}|}{|M_{\text{tumor}}|}\right) + \lambda_2 \left(1 - \frac{2\,|M_{\text{tumor}} \cap A_{\text{pred}}|}{|M_{\text{tumor}} \cup A_{\text{pred}}|}\right), \qquad (5)$$

where $A_{\text{pred}}(P, T, \boldsymbol{t})$ is the ablation zone predicted by our surrogate model for a given power P, ablation duration T and antenna translation along its axis $\boldsymbol{t}$, M_{tumor} is the target tumor mask, λ_1, λ_2 are empirically-tuned weighting factors.

The first term maximizes tumor coverage ($|M_{\text{tumor}} \cap A_{\text{pred}}|/|M_{\text{tumor}}|$), while the second minimizes damage to nearby structures via the Dice Similarity Coefficient (DSC) ($2\,|M_{\text{tumor}} \cap A_{\text{pred}}|/|M_{\text{tumor}} \cup A_{\text{pred}}|$). We solved (5) using the CMA-ES genetic algorithm, with a population size of 15 and 500 maximum iterations.

End-to-end preoperative MWA planning solution: Following [7], the algorithm finds a feasible antenna trajectory satisfying geometric and clinical constraints, including rib avoidance and protection of adjacent organs. In contrast to [7] that relied on a custom genetic algorithm, we employ a standard NSGA-II implementation to reduce optimization overhead. Using the geometrically optimized insertion direction, the algorithm determines the optimal antenna tip position along with ablation power and duration.

Intraoperative MWA re-planning solution: We build upon our previous work where we compensated for antenna misplacement during MWA without antenna reinsertion by adjusting power, duration, and axial translation [10]. Here, thanks to our surrogate model, the algorithm can identify new ablation parameters after antenna misplacement within 30 s, which significantly enhances the clinical feasibility of such a re-planning solution.

3 Results and discussion

3.1 Evaluation of the surrogate model

We evaluated the surrogate model's performance in three scenarios: (i) FE simulations on synthetic anatomies, (ii) FE simulations on patient anatomies, and (iii) *in vivo* porcine ablation experiments, with the approval of the local ethics committee. In addition, an ablation study of key network components was also conducted, including a variant using only FiLM layers (NN_F) and a simplified baseline (NN_B) without FiLM layers where power and duration inputs are passed directly to the bottleneck as additional channels. These results are summarized in Table 1.

Table 1. Evaluation of ablation predictions for the surrogate model (NN) and its variants. Average runtimes and DSCs are reported for the three scenarios.

	FE & synth. anatomies	**FE & patient anatomies**	**Porcine *in vivo* experiments**			**Avg. Runtime**
			Abl. 1	**Abl. 2**	**Abl. 3**	
FE	-	-	0.83	0.83	0.80	$\sim$ 260 s
NN	0.94 ± 0.1	0.94 ± 0.1	0.81	0.85	0.78	$\sim$ 2.8 ms
NN_F	0.94 ± 0.1	0.94 ± 0.1	0.80	0.84	0.77	$\sim$ 2.8 ms
NN_B	0.93 ± 0.1	0.93 ± 0.1	0.80	0.84	0.73	$\sim$ 2.4 ms

In the first scenario, synthetic anatomies similar to the training set were used for evaluation. The surrogate model demonstrated high accuracy with an average DSC of 0.94 on the predicted ablation zones. In the second scenario, real patient anatomies extracted from the 3D-IRCADb-01 database [15] were used to evaluate the model. Using FE simulations with varying ablation parameters, one ablation was performed for each volume, resulting in a dataset of 123 volumes of size $100 \times 100 \times 100$. The model maintained robust performance with

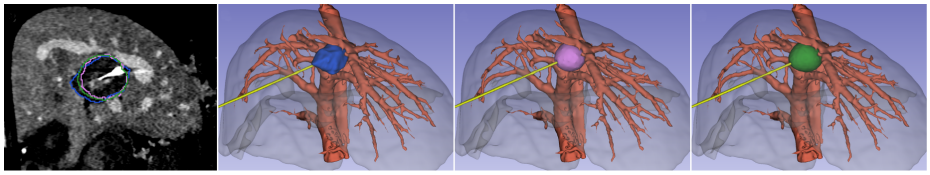

Fig. 2. Ablation zone predictions for one *in vivo* porcine experiment. Ablation zones corresponding to the ground truth (blue), FE prediction (magenta) and surrogate model prediction (green) are shown on a CT slice and 3D reconstructions of the liver.

an average DSC of 0.94, indicating strong generalization to real-patient anatomical structures despite being trained exclusively on synthetic data. Finally, in the third scenario consisting of three *in vivo* porcine ablations, we compared both the FE simulation and the surrogate model predictions against manually segmented lesions. The predicted ablation zones showed close agreement with experimental observations in shape and volume, yielding average DSCs of 0.82 across experiments, consistent with FE simulations. This is illustrated in Fig. 2 via 2D and 3D comparisons of the ground truth ablation zone against both FE and surrogate model outcomes.

In terms of computational efficiency, the average runtime for a single FE simulation was approximately 260 s, whereas the surrogate model required only 2.8 ms per inference. While performance was comparable across network variants, our ablation study indicates that FiLM and self-attention layers enhance generalization, a benefit most evident in the *in vivo* evaluation. Overall, these results indicate that the proposed surrogate model achieves near-FE accuracy while reducing computational cost by nearly five orders of magnitude.

3.2 Evaluation of the MWA planning and re-planning solutions

The planning solution proposed in Sec. 2.4 was evaluated using real patient tumors. Only tumors treatable with a single ablation (below 40 mL including 5 mm safety margin) and lying inside the liver were considered, resulting in a total of 112 tumors. Because no clinical ground truth exists for these cases, we evaluated our method by comparing it against a standard clinical baseline. First, we performed geometric planning to determine antenna placement and used manufacturer-provided charts to select the corresponding power and duration. The surrogate model was then employed to compare the ablation outcomes of this baseline approach against our proposed optimization solution.

On the 112 tumors, the average DSC was improved from 0.62 ± 0.12 to 0.80 ± 0.07 when using optimal parameters, which substantially increases healthy tissue preservation. Likewise, the coverage was improved from an average of 0.58 ± 0.11 to 0.90 ± 0.09. Suboptimal coverage occurred primarily in tumors with elongated shapes or near the 40 mL limit, which would likely require more than one ablation. Nevertheless, even in these challenging cases, the proposed physics-informed planning strategy significantly outperformed a geometry-based

Table 2. Evaluation of MWA re-planning solution. Mean relative retrieval errors for power and antenna translation for re-planning solutions based on FE and the surrogate model (NN), for three porcine *in vivo* experiments and 100 distinct initial conditions.

Experiment	**Power (FE)**	**Power (NN)**	**Trans. (FE)**	**Trans. (NN)**
Ablation 1	2.3% ±0.0%	2.6% ±0.2%	0.5% ±0.0%	0.8% ±0.0%
Ablation 2	2.3% ±0.0%	3.3% ±0.4%	2.1% ±0.0%	0.1% ±0.0%
Ablation 3	4.7% ±0.0%	4.1% ±0.4%	1.1% ±0.0%	0.1% ±0.0%

planning approach. The algorithm found optimal solutions in less than 2 minutes, including at least 30 s for the geometric planning optimization. Regarding the intraoperative re-planning solution, we followed the evaluation protocol from our previous work [10]. For an *in vivo* ablation experiment with known ablation parameters, we artificially shift the antenna tip along its axis, and use the *in vivo* ablation as tumor target to solve (5). The relative errors of optimal to known ablation parameters are reported in Table 2. The proposed solution achieved results consistent to our previously proposed FE-based re-planning solution [10], but at a fraction of the time: the total optimization time was reduced from more than 50 min to approximately 30 s.

4 Conclusion

This work presents a deep learning surrogate model for predicting MWA outcomes. By leveraging a 3D U-Net architecture enhanced with FiLM conditioning and self-attention, the model effectively encodes 3D perfusion maps and procedural ablation parameters to provide patient-specific predictions in real-time. Although trained exclusively on synthetic data, the model demonstrated remarkable generalization to both *in vivo* porcine experiments and human anatomies, achieving an average DSC of 0.82 on pre-clinical data. Due to its training on a wide range of power settings and physiological conditions, the model adapts well to patient-specific variables, requiring retraining only if new antenna hardware is introduced. In the context of MWA planning, by identifying optimal, tumor-covering ablation parameters in under 2 minutes, our work demonstrates significant potential for clinical translation. And while our work has focused on liver MWA, our approach remains generic and extensible to other organs. Further, by priorizing classical optimization over black-box approaches, we promote interpretability: in the last iterations of the optimization loop, the surrogate model can be replaced with the FE solution to compensate for model errors. This strategy enhances reliability, a key feature for clinical applications.

A primary limitation of this work lies in its focus on single-probe configurations, whereas large or irregular tumors may require the simultaneous or sequential use of multiple antennas. Future research will extend the framework to account for cumulative thermal effects and overlapping necrosis zones. Such advancements will likely require architectural modifications to the network to handle the non-linear interactions between multiple electromagnetic sources.

Acknowledgments. This work was funded by the MEDITWIN Bpifrance i-Demo project as part of the France 2030 program. The authors would like to thank Dr. Juan Verde and the *Institut Hospitalo-Universitaire (IHU)* of Strasbourg for their support.

Disclosure of Interests. The authors have no conflict of interest to declare.

References

1. Baratta, I.A., Dean, J.P., Dokken, J.S., Habera, M., Hale, J.S., Richardson, C.N., Rognes, M.E., Scroggs, M.W., Sime, N., Wells, G.N.: DOLFINx: the next generation FEniCS problem solving environment. preprint (2023). https://doi.org/10.5281/zenodo.10447666
2. Bošković, N., Radmilović-Radjenović, M., Radjenović, B.: Finite Element Analysis of Microwave Tumor Ablation Based on Open-Source Software Components. Mathematics **11**(12), 2654 (Jun 2023). https://doi.org/10.3390/math11122654
3. Debbaut, C., Segers, P., Cornillie, P., Casteleyn, C., Dierick, M., Laleman, W., Monbaliu, D.: Analyzing the human liver vascular architecture by combining vascular corrosion casting and micro-CT scanning: a feasibility study. Journal of Anatomy **224**(4), 509–517 (2014). https://doi.org/10.1111/joa.12156
4. Dettori, F., Nahmed, I., Duprez, M., Verde, J., Alvarez, P., Cotin, S.: Personalization of liver microwave ablation simulation. International Journal of Computer Assisted Radiology and Surgery (May 2026). https://doi.org/10.1007/s11548-026-03712-1
5. Heshmat, A., et al.: Using Patient-Specific 3D Modeling and Simulations to Optimize Microwave Ablation Therapy for Liver Cancer. Cancers **16**(11), 2095 (May 2024). https://doi.org/10.3390/cancers16112095
6. Ku, M.C., Fernández-Seara, M.A., Kober, F., Niendorf, T.: Noninvasive renal perfusion measurement using arterial spin labeling (ASL) MRI: basic concept (2021). https://doi.org/0.1007/978-1-0716-0978-1_13
7. Lukes, A., Bale, R., Freysinger, W.: Automatic trajectory planning for stereotactic radiofrequency ablation in non-discrete search space . https://doi.org/10.1007/s11548-025-03386-1
8. Mehtali, J., Verde, J.M., Essert, C.: C-NCA: Chained Neural Cellular Automata for Fast and Accurate Thermal Ablation Estimation . In: Proceedings of Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. vol. LNCS 15963. Springer Nature Switzerland (September 2025)
9. Meister, F., Audigier, C., Passerini, T., Lluch, E., Mihalef, V., Maier, A., Mansi, T.: Fast automatic liver tumor radiofrequency ablation planning via learned physics model. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2022. pp. 167–176. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-16449-1_17
10. Nahmed, I., Dettori, F., Verde, J., Duprez, M., Alvarez, P., Cotin, S.: A digital twin for microwave liver treatment replanning. International Journal of Computer Assisted Radiology and Surgery (Apr 2026). https://doi.org/10.1007/s11548-026-03641-z
11. Poulou, L.S.: Percutaneous microwave ablation *vs* radiofrequency ablation in the treatment of hepatocellular carcinoma. World Journal of Hepatology **7**(8), 1054 (2015). https://doi.org/10.4254/wjh.v7.i8.1054

12. Qiao, H., Sohail, A., Kim, P., Hübner, F., Vogl, T.J.: Intelligent ensemble learning-enhanced finite element modeling for precision thermal ablation in cancer therapy. Computer Methods and Programs in Biomedicine **274**, 109157 (2026). https://doi.org/10.1016/j.cmpb.2025.109157
13. Shin, M., Seo, M., Cho, S., Park, J., Kwon, J.H., Lee, D., Yoon, K.: PhysRFANet: Physics-guided neural network for real-time prediction of thermal effect during radiofrequency ablation treatment **138**, 109349. https://doi.org/10.1016/j.engappai.2024.109349
14. Singh, S., Melnik, R.: Thermal ablation of biological tissues in disease treatment: A review of computational models and future directions. Electromagnetic Biology and Medicine **39**(2), 49–88 (Apr 2020). https://doi.org/10.1080/15368378.2020.1741383
15. Soler, L., Hostettler, A., Agnus, V., Charnoz, A., Fasquel, J.B., Moreau, J., Osswald, A.B., Bouhadjar, M., Marescaux, J.: 3D Image Reconstruction for Comparison of Algorithm Database: A Patient-Specific Anatomical and Medical Image Database. Tech. rep., IRCAD, Strasbourg, France (2010), https://www.ircad.fr/research/data-sets/liver-segmentation-3d-ircadb-01/
16. Sugimoto, K., et al.: Microwave ablation vs. single-needle radiofrequency ablation for the treatment of HCC up to 4 cm: A randomized-controlled trial. JHEP Reports **7**(1), 101269 (Jan 2025). https://doi.org/10.1016/j.jhepr.2024.101269
17. Wang, Y., Hobbs, B.P., Ng, C.S.: CT Perfusion Characteristics Identify Metastatic Sites in Liver. BioMed Research International (2015). https://doi.org/10.1155/2015/120749
18. Winokur, R.S., et al.: Characterization of In Vivo Ablation Zones Following Percutaneous Microwave Ablation of the Liver with Two Commercially Available Devices: Are Manufacturer Published Reference Values Useful? Journal of Vascular and Interventional Radiology **25**(12), 1939–1946.e1 (2014). https://doi.org/10.1016/j.jvir.2014.08.014