EMiTD: Enhanced Microwave Imaging for Breast Tumor Detection

Andrew Gigie, R Krishna Kanth, A Anil Kumar, Tapas Chakravarty, Anwesha Khasnobish, Arpan Pal
TCS Research, India
Email:{gigie.andrew, rokkamkrishna.kanth, achannaanil.kumar, tapas.chakravarty, anwesha.khasnobish, arpan.pal}@tcs.com

***Abstract*—In this paper, we propose an Enhanced Microwave imaging technique for efficient breast Tumor Detection referred as EMiTD, which mainly comprises of two key steps: an intelligent scanning approach to optimize the scan duration and an efficient model-based microwave imaging (MWI) technique. The proposed model-based MWI is framed as an inverse problem by building the forward model using a Point Spread Function(PSF), and is solved by imposing sparsity prior since tumor is concentrated to limited regions. Further, we optimize the entire scanning duration by viewing the problem as a sequential decision making process for a Deep Reinforcement Learning (DRL) agent. We benchmark the proposed EMiTD against other competing techniques using a publicly available dataset. Both visual and quantitative reconstruction results are provided which indicate that the proposed EMiTD as compared to other competing techniques provides significantly improved tumor localization with close to 2$\times$ improvement in Signal to Mean Ratio (SMR). Furthermore, negligible deterioration in visual results is observed with the proposed approach in spite of using only 33% of the total measurements.**

***Index Terms*—Microwave imaging, intelligent scanning, calibration model, tumor detection**

I. Introduction

Breast cancer is one of the most prevalent cancer in the world, predominantly occurring in women [1]. Detecting breast cancer early can significantly reduce the mortality rate, but this still remains a challenge owing to shortcomings in early screening and detection with existing modalities. Detection of breast cancer is typically done using screening methods such as X-ray mammography, Magnetic Resonance Imaging (MRI) and Ultrasound imaging (US), out of which X-ray mammography is considered as a standard detection method. But these methods have their own limitations such as compression discomfort, inherent health risks, expensive, and consumes more time and effort [2]. Recently, Microwave Imaging (MWI) based techniques which can overcome some of the above mentioned limitations have been explored in literature [2].

MWI relies on the change in electrical properties when excited with electromagnetic waves. It is observed that tumor cells have more water content as compared to normal cells and hence have higher dielectric properties of around 8-10% more than the normal cells [3]. MWI is based on the principle of radar that excites electromagnetic waves and reflections from the breast are captured at different predefined locations. Further, these collected measurements are processed using various algorithms to reconstruct the MWI of the breast. [4] uses Delay-And-Sum (DAS) which makes use of shifted time delay at different antenna positions. This is a fast and effective technique to reconstruct the image but results in significant clutter artifacts [2]. Therefore, various improvements are made on DAS resulting in different variants like Delay-Multiply-And-Sum (DMAS) [5], Improved Delay-And-Sum (IDAS) [6] etc. An evaluation of these algorithms on clinical patients can be found in [7], where only DAS and DMAS were consistent with clinical reports, with DMAS having significantly reduced clutter. Further, these techniques requires *dense radar measurements* to obtain a good quality MWI [4]–[9]. This makes the entire system more time-consuming and hence is not preferrable in the present case. On the other hand, more recently, [8], [9] proposed Quantitative Microwave Imaging (QMI) techniques by employing the point spread function (PSF). Direct inversion techniques such as in [8] is typically used to solve QMI technique but this is *computationally more complex* and also error prone. Further, to reduce computational complexity, [9] described 2D FFT based technique but this requires a 2D grid based scanning with very *dense radar measurements*. While QMI based approach looks promising, the above limitations have to be efficiently addressed to make it deployment friendly.

In this paper, we propose an implementation friendly microwave imaging based tumor detection approach referred to as "Enhanced Microwave imaging for efficient breast Tumor Detection (EMiTD)". EMiTD addresses the aforementioned MWI limitations by using a computationally efficient model-based reconstruction algorithm and an intelligent radar scanning mechanism to reduce the scan duration. Firstly, we formulate the model-based MWI as an inverse image reconstruction problem by building the forward model using the PSF (which may be obtained via calibration). Since the tumor content is sparse and localized to fewer regions, we solve the inverse problem efficiently by using sparsity as a prior. Next, to reduce the number of scanning measurements, we employ an intelligent scanning mechanism based on Deep Reinforcement Learning (DRL) approach to optimize the radar acquisition locations [10], [11] . We prefix the number of measurements (say based on maximum duration) in a given episode of tumor detection and a coarse uniform scan is firstly performed to obtain an initial MWI. This coarse level image helps the RL agent to optimally suggest the remaining acquisition points. The proposed EMiTD is benchmarked against the existing methods using an open dataset collected using 3D breast phantoms having tumor [12]. Both the visual

results and Signal to Mean Ratio (SMR) (employed in [12]) is provided to compare the performance of the proposed EMiTD with the other standard DAS and DMAS approaches. The results clearly show that proposed EMiTD provides improved tumor localized image with reduced clutter and shows up to 2 times SMR improvement over other existing techniques. Further, only a marginal visual deterioration is observed with the proposed EMiTD despite using only 33% of the total measurements.

II. Proposed System

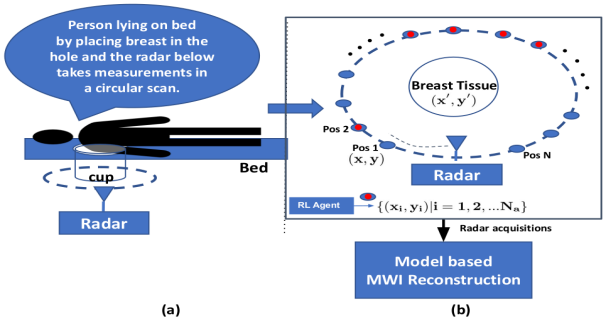

Fig. 1: (a) An illustration of a subject undergoing detection test, (b) A detailed view demonstrating the N_a optimized locations selected by the RL agent from a total of N possible radar locations around the breast tissue

Fig 1(a) illustrates the typical deployment of EMiTD similar to [12], where the subject simply lie on the bed by placing the breast inside a cup. The radar which is placed below will take measurements at different antenna positions in a circular manner. The simplified view of radar scan is shown in Fig 1(b), where the radar measurements taken at different locations are fed to a model-based MWI reconstruction algorithm to detect tumors. The RL agent intelligently chooses N_a locations from N ($N_a < N$) total measurements so that the overall scan duration is reduced, without compromising on the MWI image quality. The next subsection describes the model-based MWI reconstruction algorithm followed by the intelligent scanning using DRL.

A. Model-based MWI formulation

The monostatic radar time domain model of microwave scattering denoted by $m(x, y, t)$ captured at antenna position (x, y) and at time t can be expressed as [8]:

$$m(x,y,t) = \iint_{x',y'} \varepsilon(x',y') \underbrace{[h_{rx}^{inc} * \frac{\partial^2 E_{tx}^{inc}}{\partial t^2}]}_{Kernel\ =\ K(x,y,x',y',t)} dx'\, dy' \quad (1)$$

where $\varepsilon(x', y')$ denotes the *unknown contrast* of the breast tissue to be estimated at location (x', y'), h_{rx}^{inc} denote the impulse field response and E_{tx}^{inc} denote the total electric field by transmitter antenna. Please refer to [8] for more details. Notice from (1) that to compute unknown ε in (1), we require the knowledge of the Kernel K. K depends upon the system and we use the following approach to determine it.

1) Estimation of Kernel K: We make use of the calibration measurement to estimate K as suggested in [8]. Assume a reference point scatterer of dielectric value ε_0 and area A_r at position (x'_r, y'_r), then K from (1) can be approximated as [8]

$$K(x, y, x'_r, y'_r, t) \simeq \frac{m(x, y, t)}{\varepsilon_0 A_r} \quad (2)$$

This reference measurement corresponding to a point scatterer can be obtained either through simulations or by experiments. Although simulation enables us to precisely define the dielectric and environment properties, in practice as observed in [8] using a calibration measurement from actual experimental data is found to be more effective. Since, we are using the open dataset [12] for validation, we have used the time domain output pulse obtained from the VNA provided along with the dataset [12] for estimating K. Assuming that the background is uniform, we can use the Kernel $K(x, y, x'_r, y'_r, t)$ of the reference measurement to estimate the Kernel at any point (x', y') using the following

$$K(x, y, x', y', t) \simeq K(x, y, x'_r, y'_r, t - \delta t(x', y')) \quad (3)$$

where $\delta t(x', y') = \frac{\delta d}{c}$, where δd is defined as the Euclidean distance between the reference measurement (x'_r, y'_r) and point of interest (x', y'). It is important to notice that the kernel function K depends only on the system and is independent of ε. Thus, it needs to be evaluated *only once* for a given experimental setup.

2) Reconstruction: Discretizing the entire imaging plane into $N_x \times N_y$ and using (3), we can express (1) as follows

$$m(x, y, t) \simeq \sum_{x'} \sum_{y'} \varepsilon(x', y') K(x, y, x'_r, y'_r, t - \delta t(x', y')) \quad (4)$$

By stacking all the measurements after vectorization, we obtain the following inverse model formulation

$$\boldsymbol{m} = \boldsymbol{K}\boldsymbol{\varepsilon} + \eta \quad (5)$$

where $\boldsymbol{m} \in \mathbb{C}^{N_a N_t \times 1}$, $\boldsymbol{K} \in \mathbb{C}^{N_a N_t \times N_x N_y}$, $\boldsymbol{\varepsilon} \in \mathbb{C}^{N_x N_y \times 1}$ and N_t denotes the number of time instances taken at a particular antenna location (x, y). The dielectric constant of the tumor is large as compared to other cells [3]. In other words, only a few pixels of the reconstructed image $\boldsymbol{\varepsilon}$ corresponding to tumor will be significant whereas the other regions corresponding to normal cells can be neglected. Hence, we can assume $\boldsymbol{\varepsilon}$ to be sparse and (5) can be written as

$$\boldsymbol{\varepsilon} = \arg\min_{\boldsymbol{\varepsilon}} \|\boldsymbol{m} - \boldsymbol{K}\boldsymbol{\varepsilon}\|_2^2 + \lambda \|\boldsymbol{\varepsilon}\|_1 \quad (6)$$

where the ℓ_1 regularizer $\|\boldsymbol{\varepsilon}\|_1$ is introduced as it is well known to promote sparse solutions and λ is a hyper parameter which controls the amount of regularization. The reconstructed microwave image $\boldsymbol{\varepsilon}$ is obtained using soft thresholding and can be solved using popular ISTA algorithm, whose $(k+1)^{th}$ iterative update is shown below [13].

$$\boldsymbol{\varepsilon}^{(k+1)} = \text{soft}\left(\boldsymbol{\varepsilon}^{(k)} + \frac{1}{\alpha}\boldsymbol{K}^T(\boldsymbol{m} - \boldsymbol{K}\boldsymbol{\varepsilon}^{(k)}), \frac{\lambda}{2\alpha}\right) \quad (7)$$

Where α indicates the learning rate. On implementation, it was observed that for most instances, the above solution was

converging with less than $k = 5$ iterations. However, for efficient implementation one can also use the unfolded variant of (7) to make it faster [14].

As shall be shown later in Section III, that the proposed inverse modeling approach performs significantly better than the standard DAS and DMAS. Notice that taking more measurements of $m(x, y, t)$ not only increases the dimension of (6) (leading to increase in computational complexity), but as described in Section I it is also not preferable in practice as it leads to increase in total scan duration. Thus the following section describes an approach based on DRL to reduce this scan duration.

B. Intelligent Scanning using DRL

The radar acquisition setup is based on a DRL framework, where the task is to find optimized radar locations,(x, y), to ensure quick and accurate tumor detection. A typical DRL system has the following components [10]: (1) A set of states that defines observations received from the environment (2) policy denoted by π that enables to decide an action based on any given state (3) an environment that responds to the action taken by the agent to output the next state and (4) a reward for a given state action pair to indicate its performance.

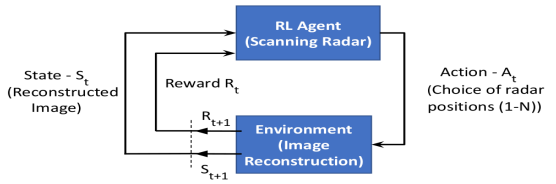

Fig. 2: DRL building blocks for the proposed EMiTD

1) DRL formulation: The DRL building blocks with respect to the proposed EMiTD system at any given time t is shown in Fig. 2.

Agent: The scanning radar acts like an RL agent.

State set: A set of reconstructed microwave image S_t reconstructed by the environment that is used by the agent to choose an action.

Action set: The set of all positions from which the radar can acquire measurements, i.e., $A_t \in \{1, 2, .., \text{N}\}$. Note that for a given episode an already observed action will be removed from the action space as taking measurements from same location will not help in improving the image reconstruction.

Environment: Based on the cumulative actions selected by the agent, a microwave image of scene is reconstructed. In our implementation, we use the DAS algorithm for generating the next state image. The DAS is chosen as it is fast, but it is important to note that one can employ any other algorithm like the one described in the previous subsection. This image is then fed as the next state to the RL agent for further acquisitions.

Reward function: For training the agent, we use the following reward function $R = 1 - MSE(S_t, GT)$, where *MSE* denotes the Mean Square Error between the state image S_t and ground truth image *GT* obtained using all measurements.

Discount factor: a hyper parameter γ that indicates the importance given to future rewards.

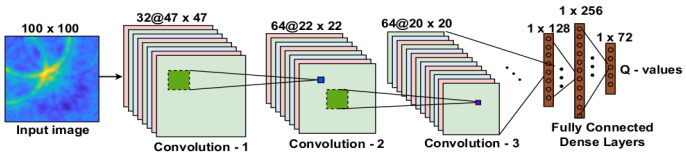

Fig. 3: DDQN value network architecture for training the RL agent

In a given episode of tumor detection, we restrict the number of radar acquisitions, $N_a < N$ based on the maximum scanning time. Initially, a coarse level scan having N_u, $N_u < N_a$ measurements is taken at uniform distances to obtain a coarse MWI S_0. This initial state S_0 helps the RL agent about the approximate tumor location and further select an optimal action from the action space A_t. The scanning radar then moves to that suggested location and collects data corresponding to that action. This data is then used by the DAS algorithm (environment) to reconstruct the next MWI state S_{t+1}. A reward is also computed by comparing the current state S_t with the ground truth DAS image GT obtained by using all the measurements and this is fed back to the RL agent to give feedback for a given state-action pair (only during training). The RL agent then uses S_{t+1} to choose the next action in the given episode. These steps are repeated until the given episode is completed. The RL agent must suitably be trained to choose the optimal radar positions for a given MWI state and this is explained in the next subsection.

2) Training RL agent: The RL agent tries to learn a policy $\pi(S_t; \theta) \rightarrow A_t$ that enables it to select an action A_t given a MWI state S_t. This policy can be learnt using many recent state-of-the-art DRL techniques [15], [16]. In our implementation we use Double Deep Q Networks (DDQN) [16] to train our RL agent. The value network architecture similar to [10] employed in our DDQN network is shown in Fig 3, which consists of three convolutional layers followed by a fully connected dense layers. The MWI image is fed as input to this network and the DDQN network outputs the value function corresponding to all the possible action space N. Further the policy can be obtained by greedily selecting an action based on this value function. An optimum $\pi(S_t; \theta)$ can be obtained by suitably training the network parameters θ. For training, we first randomly sample data from the *replay memory buffer* that consist of data tuple consisting of information about the state, action taken, reward obtained, and the next state. This sampled data from replay buffer is used to compute the loss based on the bellman equations and this is propagated to train the DDQN network parameters θ. A modified epsilon greedy exploration policy is used to fill the replay buffer [17]. For more details on training refer to [16]. The trained network can then be deployed for intelligent scanning which will give the optimal action indices required for the model-based MWI reconstruction.

C. Deployment

Based on the above discussions, here we summarize the deployment steps employed in proposed EMiTD.

Step 1: Based on the maximum scanning time, we prefix the maximum number of acquisitions N_a for an episode.

Step 2: N_u scans are taken at uniform locations to obtain the

initial coarse microwave image. The trained RL agent then uses this coarse image to sequentially suggest the remaining $(N_a - N_u)$ locations.

Step 3: The radar data at these acquired locations is then fed to the model-based MWI reconstruction for improving MWI.

III. Results and Discussions

In order to validate the proposed EMiTD system we use an openly available dataset [12].

A. Dataset description

The open dataset [12] contains 3D MRI derived breast phantoms from 9 patients having breast cancer. The phantoms were made using different tissue layers such as adipose shell, fibroglandular shell and is filled with liquids to mimic the breast composition. Tumor dielectric properties were modeled using spherical glass tubes that were immersed into these phantoms. The entire breast phantoms composition along with the tumor were placed in the center and the radar collects data at N =72 antenna positions in a circular motion. The microwave dataset comprises of different permutation of adipose shells (A1- A3), fibroglandular shells (F1- F5) and different size of tumor ranging from (1cm - 3cm) in radius. The ground truth size and location of tumor is given in the dataset to compare the MWI reconstruction performance. The RL training is done on a set of 120 instances from this dataset each containing 72 radar measurements and the performance of the trained RL agent is evaluated for 20 test instances.

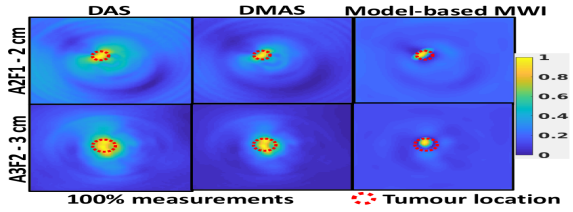

Fig. 4: Visual comparison of the proposed model-based MWI approach against DAS and DMAS. Note that we have used all the $N = 72$ measurements for reconstruction

B. Model-based MWI reconstruction

In this section, we compare the performance of the proposed model-based MWI reconstruction against the standard DAS [4], DMAS [5]. Notice, for this comparison the intelligent scanning is not considered and all $N = 72$ measurements are used. The 2-D FFT based QMI approach mentioned in [9] requires a very dense 2D scanning system for MWI which is not available with this dataset and hence this result is not provided. Since the dataset doesn't provide any calibration data, as mentioned in Section II-A, we have approximated it by the time domain output pulse from the VNA which is given in [12]. The value of α and λ after tuning is set to 0.1 and 50 respectively for the optimization framework.

Fig 4. shows the reconstructed image using different algorithms. As can be clearly seen, the proposed model-based MWI shows better tumor localization by suppressing the unwanted clutter. Here for the sake of illustration, we have only provided for two instances, however one can expect a similar performance in all other instances. This improvement with the proposed model-based MWI approach can be attributed for taking into account the kernel unlike the other techniques which uses only time delay.

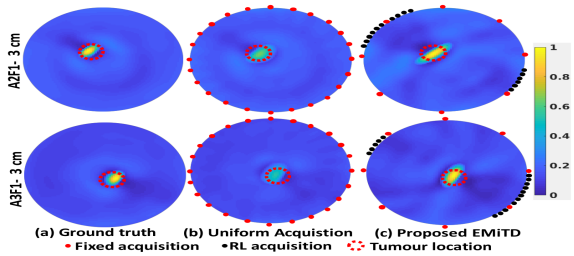

Fig. 5: Visual comparison of the proposed model-based MWI approach against uniform and RL based acquisitions. Note that we have used only $N_a = 24$ measurements for reconstruction

C. Intelligent Scanning Vs. Uniform Scanning

In this section, we provide results to demonstrate the advantage of intelligent scanning against a uniform set of measurements using the proposed model-based MWI algorithm.

The RL agent is trained with the data sampled from the replay buffer using the DDQN architecture shown in Fig 3. Reward for the RL agent at any given state S_t is computed by using the MSE of the reconstructed DAS microwave image (S_t) with the ground truth DAS image (GT) obtained using all the 72 measurements. Since, RL based scanning requires to compute MWI (state) of the scene for each action, we employ DAS in our implementation and only at the end of episode use the proposed model-based MWI. Instead of DAS, one can also use the proposed model-based MWI for state determination. During experiments, we have found that the tumor location plays a key role in optimal action determination (see Fig. 5(c)) and the impact is found to be minimal between using model-based MWI and DAS. Since DAS is faster compared to other algorithms, we have used DAS only for state determination. DDQN network uses a replay buffer size of 20000 similar to [10] and discount factor γ is set to 0.99. The DDQN network is trained for 2.3M transition steps to learn the optimum policy for any given state.

As described in Section II-B, we have fixed the maximum number of acquisitions $N_a = 24$, out of which the initial $N_u = 8$ scans are done uniformly. The RL agent then intelligently picks the antenna locations for the remaining $N_a - N_u = 16$ locations. Fig 5(a), 5(b) and 5(c) shows the visual MWI obtained using all the 72 measurements, using only 24 uniform measurements and the proposed EMiTD with $N_a = 24$ respectively. Observe from the Fig. 5(b) using only uniform measurements results in poorer tumor localization. However, with the proposed EMiTD, it is clearly evident that the tumor localization has improved significantly. Fig. 5(c) also shows the measurement locations suggested by the RL agent from which it is important to notice that more measurements are intelligently taken around the tumor which resulted in this improvement.

D. Overall System

Next, we provide the quantitative and qualitative comparison of the proposed EMiTD system against other techniques. The intelligently scanned radar measurements from $N_a = 24$

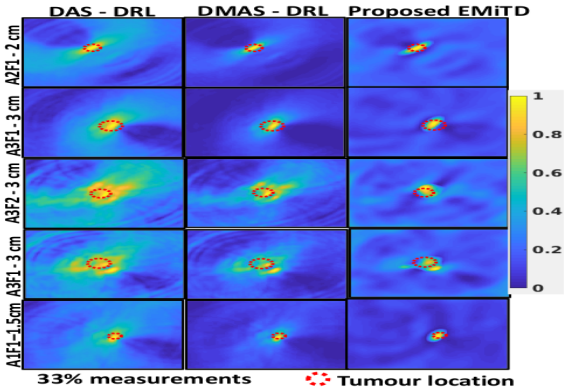

Fig. 6: Visual comparison of the proposed EMiTD approach against DAS-DRL and DMAS-DRL. Note that we have used only 24 measurements for reconstruction

Scanning	SMR Metric (dB)		
	DAS	DMAS	Model-based MWI
Uniform	7.6	14.7	28.2
RL	9.2	17.1	**31.0**

TABLE I: Comparison of SMR Metric for DAS, DMAS and Model-based MWI for both uniform and RL based scanning.

locations are fed to different algorithms such as DAS, DMAS and model-based MWI to reconstruct the corresponding MWI indicated as DAS-RL, DMAS-RL and the proposed EMiTD respectively. For visual comparison we provide 5 test cases as shown in Fig 6. As observed in all the cases EMiTD shows improved tumor localization with better clutter suppression. Further, Signal to Mean Ratio (SMR) used in [12] is used as a metric to evaluate MWI reconstruction performance. SMR is defined as

$$SMR = 20\log_{10}\frac{S_{max}}{C_{mean}} \tag{8}$$

where S_{max} is defined as maximum intensity in the tumor region (area within the red circle) and C_{mean} is defined as the mean of the intensity in the clutter region (area outside the red circle). Average SMR values for 20 test cases is shown in Table I, where The first row shows the different algorithms such as DAS, DMAS and model-based MWI using uniform scanning and the second row shows the same using RL based scanning. As observed RL based scanning has shown improved SMR metric for all the MWI algorithms. Further, DMAS performs better than DAS in terms of clutter suppression and a performance improvement of close to $2\times$ can be observed with the proposed model-based MWI.

The results thus shows that the proposed EMiTD which comprises of the efficient model-based MWI reconstruction and the intelligent RL based scanning system outperforms the existing techniques by a factor of $2\times$. Furthermore, only a minimal deterioration is observed despite using only 33% of the total measurements as compared to using full measurements.

IV. Conclusion and Future Scope

An enhanced MWI system (EMiTD) which comprises of an efficient model-based reconstruction algorithm and an RL based intelligent scanning mechanism is described in this paper. The reconstruction algorithm is framed as an inverse problem which is solved by imposing sparsity prior. A DRL based intelligent scanning is proposed to optimally choose the radar locations for a quicker scan. The reconstruction performance of EMiTD is compared with other methods using an open dataset and close to $2\times$ SMR improvement is obtained against other competing techniques. Furthermore, the visual results also demonstrates a negligible deterioration for the proposed EMiTD method in spite of using only 33% of total measurements compared to using full set of measurements.

Thus, the proposed EMiTD overcomes the limitations of the existing MWI approaches and provides improved tumor reconstruction with reduced scan duration. This, makes it an attractive breast tumor screening system that can help to extend the detection process from *clinics to anywhere*.

References

[1] WHO, "Most common cancer type," https://www.who.int/news-room/fact-sheets/detail/cancer, 2022.

[2] N. AlSawaftah, S. El-Abed, S. Dhou, and A. Zakaria, "Microwave imaging for early breast cancer detection: Current state, challenges, and future directions," *Journal of Imaging*, vol. 8, no. 5, p. 123, 2022.

[3] T. Sugitani, S.-i. Kubota, S.-i. Kuroki, K. Sogo, K. Arihiro, M. Okada, T. Kadoya, M. Hide, M. Oda, and T. Kikkawa, "Complex permittivities of breast tumor tissues obtained from cancer surgeries," *Applied Physics Letters*, vol. 104, no. 25, p. 253702, 2014.

[4] X. Li and S. Hagness, "A confocal microwave imaging algorithm for breast cancer detection," *IEEE Microwave and Wireless Components Letters*, vol. 11, no. 3, pp. 130–132, 2001.

[5] H. B. Lim, N. T. T. Nhung, E.-P. Li, and N. D. Thang, "Confocal microwave imaging for breast cancer detection: Delay-multiply-and-sum image reconstruction algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 6, pp. 1697–1704, 2008.

[6] M. Klemm, I. Craddock, J. Leendertz, A. Preece, and R. Benjamin, "Improved delay-and-sum beamforming algorithm for breast cancer detection," *International Journal of Antennas and Propagation*, vol. 2008, 2008.

[7] M. A. Elahi, D. O'Loughlin, B. R. Lavoie, M. Glavin, E. Jones, E. C. Fear, and M. O'Halloran, "Evaluation of image reconstruction algorithms for confocal microwave imaging: Application to patient data," *Sensors*, vol. 18, no. 6, p. 1678, 2018.

[8] N. K. Nikolova, *Introduction to microwave imaging*. Cambridge University Press, 2017.

[9] D. Tajik, F. Foroutan, D. S. Shumakov, A. D. Pitcher, and N. K. Nikolova, "Real-time microwave imaging of a compressed breast phantom with planar scanning," *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, vol. 2, no. 3, pp. 154–162, 2018.

[10] L. Pineda, S. Basu, A. Romero, R. Calandra, and M. Drozdzal, "Active mr k-space sampling with reinforcement learning," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23*. Springer, 2020, pp. 23–33.

[11] Z. Shen, Y. Wang, D. Wu, X. Yang, and B. Dong, "Learning to scan: A deep reinforcement learning approach for personalized scanning in ct imaging," *arXiv preprint arXiv:2006.02420*, 2020.

[12] T. Reimer, J. Krenkevich, and S. Pistorius, "An open-access experimental dataset for breast microwave imaging," in *2020 14th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2020, pp. 1–5.

[13] G. C. Calafiore and L. El Ghaoui, *Optimization models*. Cambridge university press, 2014.

[14] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 18–44, 2021.

[15] M. Hessel, J. Modayil, H. Van Hasselt, T. Schaul, G. Ostrovski, W. Dabney, D. Horgan, B. Piot, M. Azar, and D. Silver, "Rainbow: Combining improvements in deep reinforcement learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, 2018.

[16] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 30, no. 1, 2016.

[17] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.