A Novel Technique for Fluorescence Lifetime Tomography

Navid Ibtehaj Nizam,[1] Vikas Pandey,[1] Ismail Erbas,[1] Jason T. Smith,[1] and Xavier Intes[1]
Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

(*Electronic mail: nizamn@rpi.edu.)

(Dated: 19 September 2024)

Fluorescence lifetime has emerged as a unique imaging modality for quantitatively assessing *in vivo* the molecular environment of diseased tissues. Although fluorescence lifetime microscopy (in 2D) is a mature field, 3D imaging in deep tissues remains elusive and challenging owing to scattering. Herein, we report on a deep neural network (coined AUTO-FLI) that performs both 3D intensity and quantitative lifetime reconstructions in deep tissues. The proposed Deep Learning (DL)-based approach involves an *in silico* scheme to generate fluorescence lifetime data accurately. The developed DL model is validated both *in silico* and on experimental phantoms. Overall, AUTO-FLI provides accurate 3D quantitative estimates of both intensity and lifetime distributions in highly scattering media, demonstrating its unique potential for fluorescence lifetime-based molecular imaging at the mesoscopic and macroscopic scale.

I. INTRODUCTION

In recent years, molecular imaging has become a vital tool for investigating biological processes *in vivo* in both preclinical and clinical settings. Due to its high sensitivity, availability of a wide range of fluorophores, and multiplexing capabilities, fluorescence optical imaging has emerged as the most prominent molecular imaging technique in the last decade[1,2]. While intensity (brightness) remains the primary contrast mechanism in fluorescence optical imaging, fluorescence lifetime is increasingly leveraged for its unique advantages. Fluorescence Lifetime Imaging (FLI) can quantitatively sense various intracellular parameters, including metabolic state, viscosity, temperature, and pH[3–6]. Additionally, fluorescence lifetime allows for robust quantification of Förster Resonance Energy Transfer (FRET), enabling nanoscale assays *in vivo* and the study of drug-target engagement[4,7,8].

FLI has grown steadily over the past three decades, with increased adoption due to user-friendly FLI Microscopes (FLIM)[4]. Simultaneously, FLI has expanded into translational applications, from mesoscopic (mFLI)[9] to macroscopic regimes (MFLI)[10,11]. These larger-scale implementations are more challenging than microscopic ones because they require Near-Infrared (NIR) fluorophores for deeper tissue penetration. Red-shifted fluorophores typically have shorter lifetimes (nanosecond or sub-nanosecond) compared to a few nanoseconds in the visible range, and detectors have low quantum efficiency (a few percent)[12]. Additionally, 3D lifetime imaging in deep tissues is hindered by light scattering. Achieving 3D quantitative maps in these regimes requires solving a diffuse optical inverse problem, which considers the spatial and temporal characteristics of illumination light, its propagation, energy transformation, temporal delays generating fluorescence photons, and the propagation of these photons to the detection surface. This problem is computationally difficult to solve using classical methods, limiting mFLI and MFLI mainly to 2D imaging applications.

The challenge in 3D Fluorescence Lifetime Tomography (FLT) arises from it being a "double ill-posed problem" because FLT involves reconstructing a 3D volume from a scattering medium while simultaneously determining the lifetime for each voxel. Two main approaches have been proposed to address the FLT challenges: estimating the lifetime from 2D measurements before casting the inverse problem using this prior information, or linearizing the lifetime exponential model to stabilize the inverse problem. Pertaining to the first method, Kumar et al.[13] proposed an asymptotic approach in which 2D fluorescence data are preprocessed to obtain the relative intensities associated with individual fluorophores, and then the inverse problem is cast independently for each fluorophore using a quasi-Continuous Wave (CW) formulation. However, this asymptotic assumption is only valid for long lifetime values(t>0.5 ns). This approach was also combined with structural priors obtained by CT to further constrain the inverse problem[14]. However, the integration of CT data increases the complexity and cost of the imaging process, and the requirement for high-quality CT images may not always be feasible. Furthermore, Zhang et al.[15] used a mathematical technique called the fused LASSO method, which involves reconstructing the fluorescence yield map and object geometry as prior information, helping to mitigate the ill conditions associated with the nonlinear physical model of FLT. However, this linear approximation is computationally complex and highly dependent on regularization parameters. Another approach proposed by our group involves estimating lifetimes from 2D measurements and incorporating these values as priors in the 3D inverse problem, utilizing the rich information of time-resolved datasets[16]. Although this approach is technically more challenging due to the increased computational burden and the need for precise time-stamping, it leverages the unique strengths of early photons for improved resolution and late photons for quantification[17]. To summarize, the common limitations of these techniques are that they are time-consuming and heavily reliant on parameter optimization. Following recent trends, it is expected that Deep Learning (DL) methodologies can provide efficient frameworks to overcome these challenges. The seminal work by Smith et al.[18] introduced FLI-NET, a Convolutional Neural Network (CNN)-based architecture for rapid lifetime imaging. However, FLI-NET was designed for 2D FLI and was not extended to 3D FLT.

The use of Deep Neural Networks (DNNs) in Fluorescence

Molecular Tomography (for 3D intensity reconstructions) has flourished in recent years. Long et al.[19] proposed one of the first networks to achieve FMT reconstructions in the mesoscopic domain, employing a CNN. Other strategies by Zhang et al.[20] utilized CNNs with skip connections to achieve high-resolution FMT reconstructions. Previously, we introduced a novel architecture called ModAM[21], which combined a K-space illumination basis and a unique *in silico* data generation workflow for 3D FMT reconstructions[22]. However, despite the proliferation of DNNs specializing in FMT intensity reconstructions, little to no work in the literature has achieved 3D lifetime reconstructions with high fidelity. The lack of progress in this field is due to the high degree of difficulty in designing a network that accurately learns the lifetime from the raw fluorescence decays while simultaneously reconstructing a 3D inclusion. Furthermore, the reliance on experimental data for training is a massive bottleneck, as collecting a large experimental dataset for different physical and Optical Properties (OPs) involved with fluorescence lifetime estimation is extremely time-consuming and serves as a deterrent to designing a network that can achieve 3D FLT reconstructions.

Here, we propose AUTO-FLI, a CNN that performs end-to-end 3D FMT reconstructions for both relative quantum yields and lifetimes. This is the first DL-based technique to achieve full 3D volumetric reconstruction of lifetime from raw 2D fluorescence decays. The two-stage workflow involves 3D quantum yield reconstruction using the ModAM network, which constrains the 3D lifetime output from a modified FLI-NET. Using a mono-exponential lifetime model (for simplicity and a first analysis), we deploy a K-space illumination basis[22] and generate training data with Monte-Carlo Xtreme (MCX), a MATLAB-based simulator[23]. This approach ensures accurate modeling of fluorescence imaging physics and mitigates the need for extensive experimental data. AUTO-FLI is validated on *in silico* and experimental data from tissue-mimicking phantoms, with results benchmarked against a state-of-the-art 2D fluorescence estimation software, AlliGator[24], and evaluated for Mean Squared Error (MSE) and Volume Error (VE).

II. METHODS

A. Data Generation Workflow

The data generation workflow is illustrated in Fig. 1. To generate the training dataset, we first design *in silico* phantoms in MCX (Fig. 1B). These phantoms contain one or two fluorescent inclusions, which are randomly picked from our previously proposed EEMINST dataset (Fig. 1A)[22]. This dataset is designed to have high spatial heterogeneity, enabling our network to learn to reconstruct complex structures (often associated with cancer biology). Moreover, the fluorescent embedding(s) (with thicknesses varying from $2-4$ mm) are placed at different depths($1-10$ mm) within the phantom. In the case of multiple fluorescent inclusions, the relative concentration of the species ranges from $1:1$ to $1:4$. Additionally, the background is given OPs relevant to soft tissues with an absorption coefficient (μ_a) of 0.004 mm^{-1} and a reduced scattering coefficient (μ_s') of 1 mm^{-1}. Next, we use wide-field, K-space illumination patterns (the same patterns used in Nizam et al.[22], Fig. 1D) acquired experimentally from our gated-Intensified CCD (ICCD) system (discussed in the experimental setup section) to illuminate the phantom (Fig. 1C). Through MC simulation, the flux at the position of the embeddings (ϕ) is gathered and then multiplied with the relative concentration matrix ($1:2$ as an example in Fig. 1E) to obtain ϕ_c (shown for a single K-space pattern in Fig. 1F). In the next step ϕ_c is convolved with an exponential decay profile corresponding to a mono-exponential lifetime ranging from $0.3-1.5$ ns (an example lifetime map is given in Fig. 1G), which results in a time-varying profile, ϕ_t (shown for a single K-space pattern in Fig. 1H).

This ϕ_t (obtained for each K-space pattern) is now used as a propagator from inside the medium (in the opposite direction to simulate fluorescence) and again through MC simulation (Fig. 1I), the flux at the surface of the phantom, ϕ_s (shown for a single K-space pattern in Fig. 1J), is collected. Finally, ϕ_s is convolved with the Instrument Response Function (IRF) (obtained experimentally using the gated ICCD system from a thin sheet of paper placed at the same height as the surface of the experimental phantom and shown for a single pixel in Fig.1K), to obtain the Temporal Point Spread Functions (TPSFs), ϕ_T (shown normalized for a single K-space pattern in Fig. 1L). Only the falling edge of the normalized TPSFs (corresponding to the decay) is preserved for training the network. Also, ϕ_T is summed in time, obtaining a CW map for each pattern, and using K-sparse acquisition (with the help of 64 virtual point detectors, as shown in Fig. 1M)[22]), we obtain a CW measurement vector for the phantom (shown in Fig. 1N). This CW measurement vector is used to train the ModAM branch of the network for 3D intensity reconstruction.

B. CNN Architecture and Training

The CNN architecture is shown in Fig. 2A. It consists of two branches, the ModAM branch for 3D intensity reconstruction and the modified FLI-NET branch for 3D lifetime reconstruction, the output of which is constrained by the 3D intensity mask. Both these branches are mainly a combination of 2D and 3D convolutional layers. The output of each layer is Batch Normalized (BN) and passed through ReLU activation. The rationale for the different layers in the CNN architecture has been explained previously in Smith et al.[18] and Nizam et al.[22]. However, for clarity, we briefly explain the purpose of the different layers.

The 3D convolution layers (Conv3D) in the modified FLI-NET branch enable the extraction of spatially independent features along each TPSF for each pattern. The Residual Block (ResBlock) allows even more temporal features to be extracted while resolving problems related to vanishing gradients (without increasing parameter count and, hence, computational complexity)[25]. Finally, the 2D convolutional layers (Conv2D) downsample the reconstructed volume to the size of our experimental phantom ($40\times40\times10mm^3$).

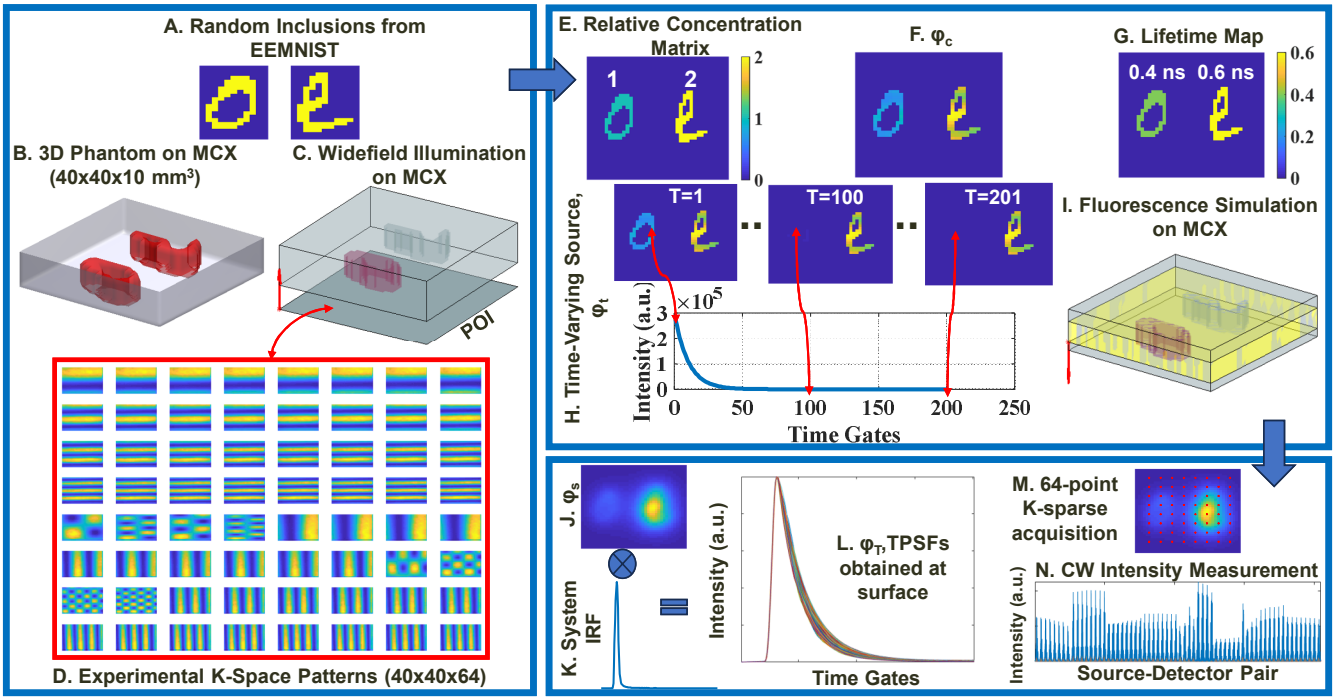


FIG. 1. The complete data generation workflow with an example *in silico* phantom in B containing two fluorescent inclusions randomly picked from the EEMNIST dataset shown in A. This phantom is illuminated using K-space illumination patterns (on the Plane of Illumination (POI) marked in C) acquired experimentally, as shown in D. The widefield illumination function in MCX shown in C allows the flux at the position of the embeddings to be collected, which is correlated with the concentration matrix in E to obtain ϕ_c in F. F is convolved with a mono-exponential lifetime profile associated with the lifetime map in G to obtain H as a time-varying source. H is used as a propagator from inside the medium on MCX, as shown in I, to obtain J, which is convolved with the system IRF in K to obtain the TPSFs in L and through 64-point K-space acquisition on K, as illustrated by the red dots in M, the CW measurement vector is obtained in N.

The ModAM branch starts with a Dense layer since Dense layers are widely associated with shape recognition and help capture the intricate variation in the CW measurement vector associated with a change in the morphology of the fluorescent inclusions. The Conv3D and Conv2D layers are used for upsampling and downsampling, respectively, the extracted features from the 1D measurement vector, and eventually, to reconstruct a 3D intensity volume.

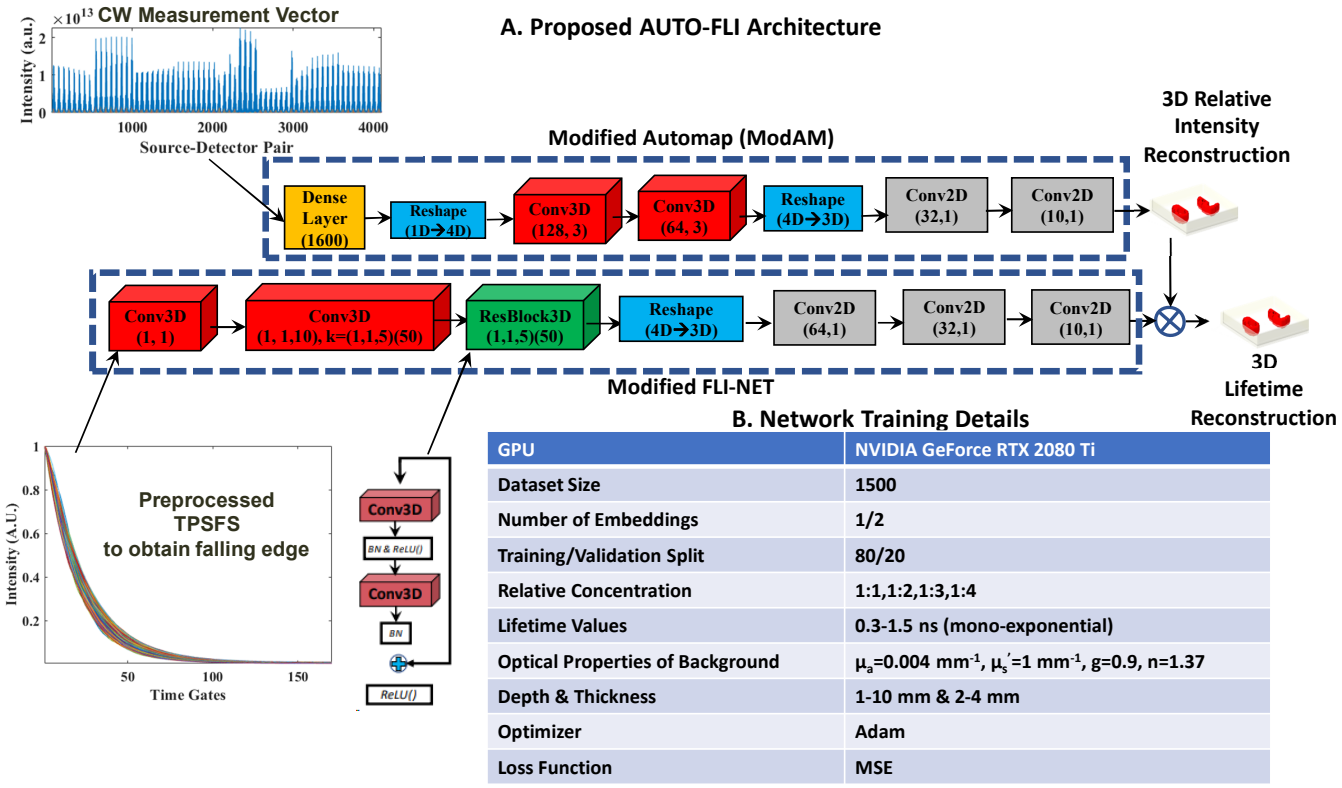
Additionally, we tabulate the details of network training in Fig. 2B. To summarize, we use a dataset size of 1500 samples, with an 80/20 training/validation split, while employing an Adam optimizer. As discussed previously, we use MSE as the loss function. Furthermore, the relative concentration, depth, thickness, and lifetime of the *in silico* phantoms are suitably varied over a range relevant to soft tissues to train the network.

C. Experimental Setup

The experimental setup consists of a gated-ICCD system. The optical imaging is performed in reflective geometry with a Field-of View of approximately 40×40 mm^2. A total of 201 gates are used with a gatewidth of 40 ps. The K-space illumination patterns are projected using a Digital Micromirror Device (D4110, Digital Light Innovations, TX). For imaging the experimental phantoms, we excite at a wavelength of 700 nm using a Mai Tai high-powered laser (Spectra-Physics, CA). Further details of the experimental setup can be found elsewhere[8].

D. Phantom Preparation

To prepare the phantom, we combine distilled water, 1% India Ink (Speedball Art Products, NC), 20% intralipid (Sigma–Aldrich, MO) of volumes of 157.05 ml, 1.05 ml, and 11.90 ml, respectively, with 1.7 g of agar to form a homogeneous phantom that has roughly the same background OPs as the *in silico* phantoms used in training. Two cylindrical cavities are formed in the phantom at depths of $2-5$ mm, with the centers approximately 20 mm apart, in which we contain AF700 at 50μM concentration in Dimethyl sulfoxide (DMSO) and Phosphate-buffered saline (PBS) buffers.



B. Network Training Details

GPU	NVIDIA GeForce RTX 2080 Ti
Dataset Size	1500
Number of Embeddings	1/2
Training/Validation Split	80/20
Relative Concentration	1:1,1:2,1:3,1:4
Lifetime Values	0.3-1.5 ns (mono-exponential)
Optical Properties of Background	μ_a=0.004 mm^{-1}, μ_s'=1 mm^{-1}, g=0.9, n=1.37
Depth & Thickness	1-10 mm & 2-4 mm
Optimizer	Adam
Loss Function	MSE

FIG. 2. The Proposed AUTO-FLI architecture, a combination of ModAM and modified FLI-NET is shown in A. The network training details is tabulated in B.

III. RESULTS

A. In Silico Results

The performance of our proposed AUTO-FLI network is first tested on *in silico* phantoms that were not part of the training stage. For demonstration, we used an *in silico* phantom containing two fluorescent inclusions, each of which has a different lifetime (0.9 ns for the left inclusion and 0.7 ns for the right inclusion) and different relative concentration (1 : 2). The inclusions are placed at depths of $2-4$ mm, as shown in Fig. 3A. The 3D lifetime reconstruction by the AUTO-FLI network, rendered by the Imaris software, is displayed in Fig. 3B. In Fig. 3C, we display the Ground Truth (GT) relative concentration map, and in Fig. 3D, we show the obtained 2D intensity maps corresponding to depths $1-4$ mm. Additionally, in Fig. 3E, we present the intensity profile along the red dotted line in Fig. 3C. It is evident that the ModAM branch of the AUTO-FLI network reconstructs the 3D intensity with a high degree of fidelity and manages to recover the relative concentration accurately. The accuracy is numerically validated by a low VE of 17.86% and an MSE of 0.0328. The 3D lifetime reconstruction results are shown in Figs.3F-H. The 3D reconstruction, in terms of structure, is the same as the output of the ModAM branch since the output of the 3D FLI-NET branch is constrained by the 3D intensity mask. In Fig. 3G, we display the 2D lifetime maps corresponding to depths $1-4$ mm. The GT lifetime map is illustrated for comparison in Fig. 3F. Additionally, we present the violin plot distribution of the reconstructed lifetime values for the two fluorescent species (along with the GT) in Fig. 3H. It can be deduced from these plots that the proposed AUTO-FLI framework accurately reconstructs the 3D lifetime (with an MSE value of 0.0088).

B. Experimental Phantom Results

To further validate the proposed AUTO-FLI network, we tested its performance on an experimental tissue-mimicking phantom. The same network trained for the *in silico* phantoms is used to reconstruct the experimental phantom. A schematic of the phantom is shown in Fig. 4A. The experimental phantom consists of two cylindrical fluorescent inclusions AF700 in DMSO (left) and PBS (right) embedded in agar (background OPs set using a mixture of ink and intralipid), the center of which is placed at a depth of 3 mm from the surface of the phantom. The inclusions are approximately 4 mm thick. A schematic of the phantom is shown in Fig. 4A. The two different buffers ensure that the two inclusions have slightly different lifetimes while having approxi-

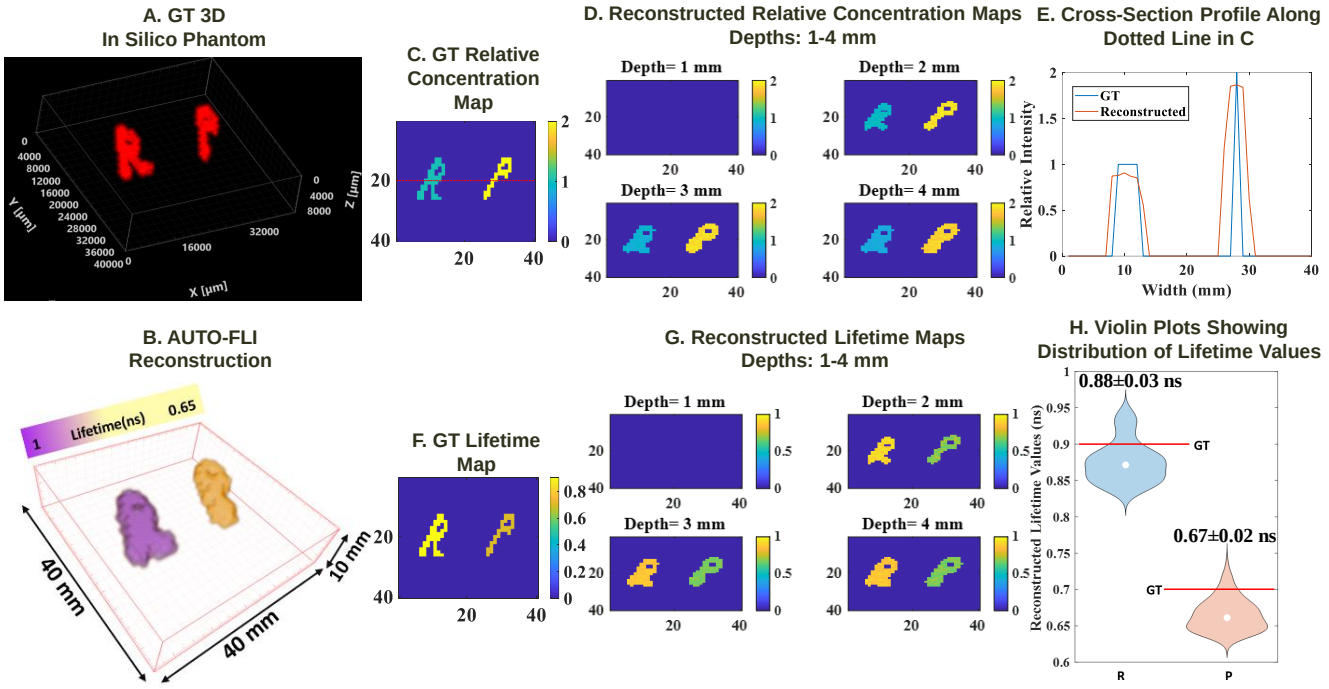


FIG. 3. The GT 3D *in silico* phantom, not part of the training stage of the AUTO-FLI network, is shown in A with the GT intensity and lifetime maps in C and E, respectively. The 3D reconstruction, rendered by Imaris, by the AUTO-FLI framework is displayed in B, and the reconstructed 2D intensity maps and 2D lifetime maps are in D and F, respectively. The intensity profile along the red dotted line along with the GT is plotted in E, while the violin plots showing the reconstructed lifetime distribution of the two fluorescent species (with red lines representing the GT) are presented in H. The 3D reconstruction results are shown at 50% of the maximum isovolume.

mately the same concentration. However, it should be noted that DMSO gives a slightly brighter signal compared to PBS. The results of the 3D volumetric reconstruction, again rendered by Imaris, are presented in Fig. 4B, and the 2D normalized intensity maps corresponding to depths $1-6$ mm are illustrated in Fig. 4C. Moreover, the violin plots showing the distribution of the relative intensity values are plotted in Fig. 4D. It is observed that the ModAM branch of the network reconstructs the relative concentration close to what is expected, with a slightly higher concentration of AF700 in Dimethyl sulfoxide (DMSO, left) compared to AF700 in Phosphate-buffered saline (PBS, right). The relative concentration (average) is obtained as 1.08 : 1 (DMSO:PBS). The 2D lifetime maps reconstructed (corresponding to depth $1-6$ mm) by the AUTO-FLI framework are shown in Fig. 4E, while the distribution of lifetime values are presented in Fig. 4F. The reconstructed lifetime values agree closely with the lifetime values for AF700 in DMSO and PBS quoted in the literature[26]. Furthermore, for benchmarking, we obtain the lifetime values estimated from the experimental phantom using the state-of-the-art AlliGator software. The AlliGator software estimates the lifetime of a 2D mask placed on the 2D widefield image of the phantom. The mean ($\pm$ standard deviation) of the left (DMSO) and right (PBS) inclusion is calculated by AlliGator as $1.14(\pm 0.02)$ ns and $1.03(\pm 0.02)$ ns, respectively. Thus, the lifetime values obtained from AlliGator are comparable to those produced by AUTO-FLI. However, it is to be noted that the AUTO-FLI network can carry out a single 3D lifetime reconstruction in approximately 1 s (NVIDIA GeForce RTX 2080 Ti). Conversely, on the same GPU, it takes almost 10 minutes for AlliGator to calculate the lifetime on 10 pixels. Hence, the experimental results further justify the suitability of the proposed workflow for experimental phantoms.

IV. DISCUSSION

This work proposes a framework for simultaneous 3D intensity and lifetime reconstructions using a two-stage DNN architecture. Using a single-stage network for simultaneous 3D intensity and lifetime reconstructions from raw TPSFs presents significant challenges due to the inherent differences in these parameters. Intensity reconstruction focuses on spatial distribution, while lifetime reconstruction deals with temporal decay characteristics. Additionally, the need for accurate time-stamping in lifetime imaging adds another layer of complexity. This complexity often results in sub-optimal performance for either or both tasks. Previous studies, for instance, the concept of multi-task learning[27], suggests that dividing complex tasks into simpler, more focused tasks can lead to more accurate outcomes in DL models. Our results align with these findings, as the two-stage AUTO-FLI network showed improved stability and accuracy compared to a single-stage network.To show the impact of removing the ModAM

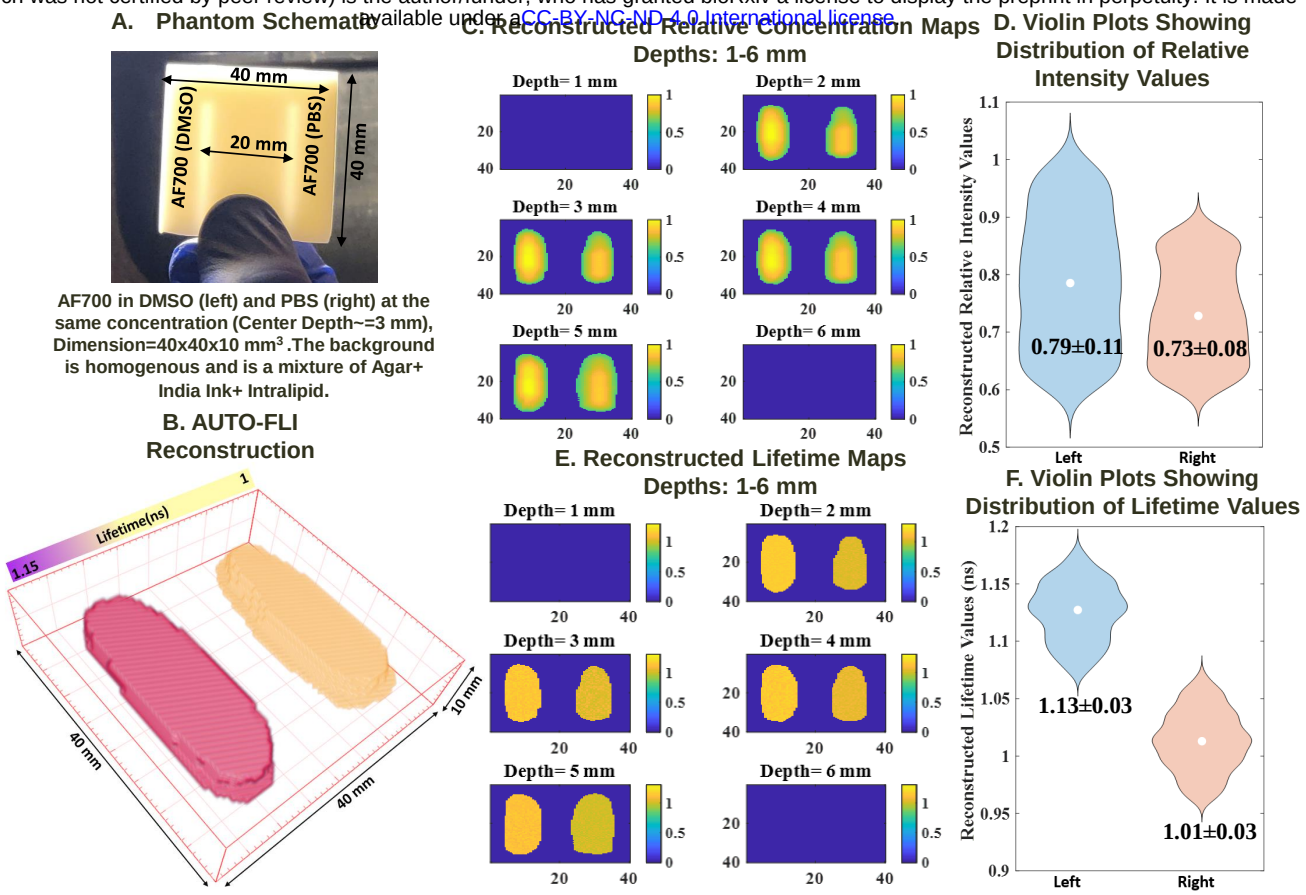


FIG. 4. The schematic of the experimental phantom is shown in A. The 3D reconstruction (rendered by Imaris) is presented in B, along with the obtained 2D relative concentration and lifetime maps in C and E, respectively. The distribution of the relative intensity and lifetime values, along with the mean and the standard deviation, for both fluorescent inclusions are presented in terms of violin plots in D and F, respectively. The 3D reconstruction results are shown at 50% of the maximum isovolume.

stage, we generate the 3D lifetime reconstruction using only the 3D FLI-Net stage on the *in silico* phantom shown previously. The results are shown in Fig. 5. It is seen that although the network reconstructs the shapes to an extent, it is not as accurate as the two-stage AUTO-FLI framework, as exhibited by an increased VE and an increased MSE of of 44.26%, and 0.023, respectively. The results further demonstrate the need to split the learning process into two stages. Furthermore, it is essential to test the stability of the proposed AUTO-FLI network. We exhibit the stability of the network by repeating the training and validation phase 10 times and plot the resulting training and validation curves in Fig.6 (with the MSE as the loss function). All the curves show smooth convergence with a low standard deviation (± 0.005), which reveals the stability of the proposed DNN.

Additionally, there has been significant interest recently in using Physics-Informed Neural Networks to solve problems in diffuse optical imaging[28]. Such networks consider the mathematical equations for the underlying physics that govern, for instance, the diffusion equation for light propagation in a turbid medium. Although our network follows a more traditional CNN architecture, the accurate *in silico* modeling of fluorescence emission and detection on MCX considers the medium's physical and optical properties. This accurate modeling of the fluorescence process using the state-of-the-art MCX software is one of the critical reasons for the success AUTO-FLI for FLT reconstructions, while similar efforts in the past have not been as effective.

Moreover, there are some limitations to the proposed workflow. We have considered a simplified mono-exponential model for lifetime estimation, whereas the original FLI-NET model used a bi-exponential model suitable for FRET studies. Also, we have not included *in vivo* results in this work since this is the first study in simultaneous 3D intensity and lifetime reconstructions. Future studies will look to develop the workflow for higher-order exponential models and *in vivo* small animal models. *In vivo* imaging may require the development of the data generation workflow to incorporate mesh-based models by using Mesh-based Monte Carlo (MMC) instead of MCX because complicated animal models require a higher degree of accuracy in the data generator, which is provided by mesh-based representations in the MC rather than a voxel-based one (as provided by MCX).

V. CONCLUSION

In this study, we introduced AUTO-FLI, an innovative DL framework for 3D FLT that offers significant advancements in accuracy and computational efficiency. By employing a two-

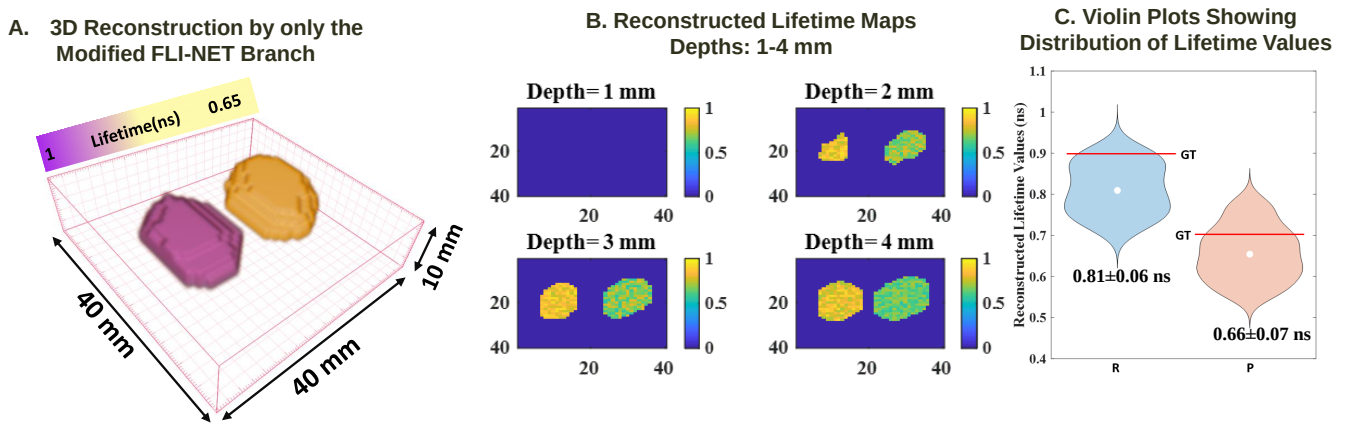


FIG. 5. The 3D reconstruction obtained using only the Modified FLI-NET stage (by eliminating ModAM) is shown in A for the same *in silico* phantom used in Fig. 3 along with the obtained 2D lifetime maps in B, and the violin plot distribution of the reconstructed lifetime values in C. Again, the 3D reconstruction results are shown at 50% of the maximum isovolume.

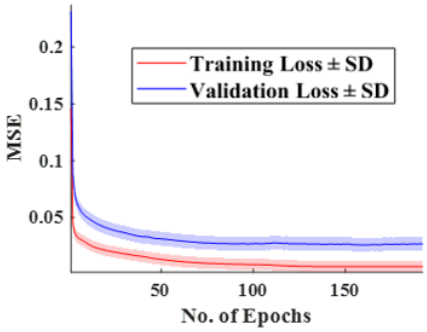


FIG. 6. An ensemble of training and validation curves obtained by training the AUTO-FLI network 10 times.

stage CNN architecture, AUTO-FLI achieves high-fidelity reconstructions of 3D intensity and lifetime distributions from raw 2D fluorescence decays, outperforming traditional methods. The integration of MCX for data generation reduces training time and resources, enhancing the model's robustness and versatility across various biological and medical imaging scenarios. This technique holds promising potential for both preclinical research and clinical diagnostics.

ACKNOWLEDGMENTS

The authors would like to acknowledge the generous funding received from the National Institutes of Health under grants R01-CA271371, R01-CA237267, R01-CA207725, and R01-CA250636. The authors would also like to thank Mr. Harrison Yee for his help with Imaris.

DATA AVAILABILITY STATEMENT

The codes and data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

DISCLOSURES

The authors declare no conflicts of interest.

[1]B. W. Pogue, "Optics in the molecular imaging race," Opt. Photonics News **26**, 24–31 (2015).
[2]S. H. Yun and S. J. Kwok, "Light in diagnosis, therapy and surgery," Nature biomedical engineering **1**, 0008 (2017).
[3]M. C. Skala, K. M. Riching, A. Gendron-Fitzpatrick, J. Eickhoff, K. W. Eliceiri, J. G. White, and N. Ramanujam, "In vivo multiphoton microscopy of nadh and fad redox states, fluorescence lifetimes, and cellular morphology in precancerous epithelia," Proceedings of the National Academy of Sciences **104**, 19494–19499 (2007).
[4]S. Rajoria, L. Zhao, X. Intes, and M. Barroso, "Flim-fret for cancer applications," Current molecular imaging (discontinued) **3**, 144–161 (2014).
[5]B. Das, L. Shi, Y. Budansky, A. Rodriguez-Contreras, and R. Alfano, "Alzheimer mouse brain tissue measured by time resolved fluorescence spectroscopy using single-and multi-photon excitation of label free native molecules," Journal of biophotonics **11**, e201600318 (2018).
[6]L. Sauer, K. M. Andersen, C. Dysli, M. S. Zinkernagel, P. S. Bernstein, and M. Hammer, "Review of clinical approaches in fluorescence lifetime imaging ophthalmoscopy," Journal of biomedical optics **23**, 091415–091415 (2018).

[7]E. A. Jares-Erijman and T. M. Jovin, "FRET imaging," Nature Biotechnology **21**, 1387–1395 (2003).

[8]M. Ochoa, A. Rudkouskaya, J. T. Smith, X. Intes, and M. Barroso, "Macroscopic fluorescence lifetime imaging for monitoring of drug–target engagement," in *Biomedical Engineering Technologies: Volume 2* (Springer, 2022) pp. 837–856.

[9]S. Gao, M. Li, J. T. Smith, and X. Intes, "Design and characterization of a time-domain optical tomography platform for mesoscopic lifetime imaging," Biomedical Optics Express **13**, 4637–4651 (2022).

[10]V. Venugopal, J. Chen, and X. Intes, "Development of an optical imaging platform for functional imaging of small animals using wide-field excitation," Biomedical optics express **1**, 143–156 (2010).

[11]R. Pal, T. M. Lwin, M. Krishnamoorthy, H. R. Collins, C. D. Chan, A. Prilutskiy, M. P. Nasrallah, T. H. Dijkhuis, S. Shukla, A. L. Kendall, *et al.*, "Fluorescence lifetime of injected indocyanine green as a universal marker of solid tumours in patients," Nature Biomedical Engineering **7**, 1649–1666 (2023).

[12]L. Chavez, S. Gao, and X. Intes, "Characterization of fluorescence lifetime of organic fluorophores for molecular imaging in the shortwave infrared window," Journal of Biomedical Optics **28**, 094806–094806 (2023).

[13]A. T. Kumar, S. B. Raymond, G. Boverman, D. A. Boas, and B. J. Bacskai, "Time resolved fluorescence tomography of turbid media based on lifetime contrast," Optics express **14**, 12255–12270 (2006).

[14]W. L. Rice, D. M. Shcherbakova, V. V. Verkhusha, and A. T. Kumar, "In vivo tomographic imaging of deep-seated cancer using fluorescence lifetime contrast," Cancer research **75**, 1236–1243 (2015).

[15]P. Zhang, J. Liu, H. Hui, Y. An, K. Wang, X. Yang, and J. Tian, "Linear scheme for the direct reconstruction of noncontact time-domain fluorescence molecular lifetime tomography," Applied Optics **59**, 7961–7967 (2020).

[16]V. Venugopal, J. Chen, M. Barroso, and X. Intes, "Quantitative tomographic imaging of intermolecular fret in small animals," Biomedical optics express **3**, 3161–3175 (2012).

[17]L. Zhao, H. Yang, W. Cong, G. Wang, and X. Intes, "L p regularization for early gate fluorescence molecular tomography," Optics letters **39**, 4156–4159 (2014).

[18]J. T. Smith, R. Yao, N. Sinsuebphon, A. Rudkouskaya, N. Un, J. Mazurkiewicz, M. Barroso, P. Yan, and X. Intes, "Fast fit-free analysis of fluorescence lifetime imaging via deep learning," Proceedings of the National Academy of Sciences **116**, 24019–24030 (2019).

[19]F. Long, "Deep learning-based mesoscopic fluorescence molecular tomography: an in silico study," Journal of Medical Imaging **5**, 036001–036001 (2018).

[20]P. Zhang, G. Fan, T. Xing, F. Song, and G. Zhang, "Uhr-deepfmt: ultra-high spatial resolution reconstruction of fluorescence molecular tomography based on 3-d fusion dual-sampling deep neural network," IEEE Transactions on Medical Imaging **40**, 3217–3228 (2021).

[21]B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," Nature **555**, 487–492 (2018).

[22]N. I. Nizam, M. Ochoa, J. T. Smith, and X. Intes, "3d k-space reflectance fluorescence tomography via deep learning," Optics letters **47**, 1533–1536 (2022).

[23]Q. Fang and D. A. Boas, "Monte Carlo Simulation of Photon Migration in 3D Turbid Media Accelerated by Graphics Processing Units," Opt. Express **17**, 20178–20190 (2009).

[24]S.-J. Chen, N. Sinsuebphon, A. Rudkouskaya, M. Barroso, X. Intes, and X. Michalet, "In vitro and in vivo phasor analysis of stoichiometry and pharmacokinetics using short-lifetime near-infrared dyes and time-gated imaging," Journal of biophotonics **12**, e201800185 (2019).

[25]K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016) pp. 770–778.

[26]J. T. Smith, A. Rudkouskaya, S. Gao, J. M. Gupta, A. Ulku, C. Bruschini, E. Charbon, S. Weiss, M. Barroso, X. Intes, *et al.*, "In vitro and in vivo nir fluorescence lifetime imaging with a time-gated spad camera," Optica **9**, 532–544 (2022).

[27]S. Ruder, "An overview of multi-task learning in deep neural networks," arXiv preprint arXiv:1706.05098 (2017).

[28]F. Wang, S. H. Kim, Y. Zhao, A. Raghuram, A. Veeraraghavan, J. Robinson, and A. H. Hielscher, "High-speed time-domain diffuse optical tomography with a sensitivity equation-based neural network," IEEE Transactions on Computational Imaging (2023).