




Article

Simple and Robust Deep Learning Approach for Fast Fluorescence Lifetime Imaging

Quan Wang ^{1,*}, Yahui Li ², Dong Xiao ¹ , Zhenya Zang ¹ , Zi'ao Jiao ¹, Yu Chen ³ and David Day Uei Li ¹ 

¹ Department of Biomedical Engineering, University of Strathclyde, Glasgow G4 0RU, UK

² Key Laboratory of Ultra-Fast Photoelectric Diagnostics Technology, Xi'an Institute of Optics and Precision Mechanics, Xi'an 710049, China

³ Department of Physics, University of Strathclyde, Glasgow G4 0NG, UK

* Correspondence: quan.wang.100@strath.ac.uk

Abstract: Fluorescence lifetime imaging (FLIM) is a powerful tool that provides unique quantitative information for biomedical research. In this study, we propose a multi-layer-perceptron-based mixer (MLP-Mixer) deep learning (DL) algorithm named FLIM-MLP-Mixer for fast and robust FLIM analysis. The FLIM-MLP-Mixer has a simple network architecture yet a powerful learning ability from data. Compared with the traditional fitting and previously reported DL methods, the FLIM-MLP-Mixer shows superior performance in terms of accuracy and calculation speed, which has been validated using both synthetic and experimental data. All results indicate that our proposed method is well suited for accurately estimating lifetime parameters from measured fluorescence histograms, and it has great potential in various real-time FLIM applications.

Keywords: fluorescence lifetime imaging (FLIM); deep learning; imaging analysis



Citation: Wang, Q.; Li, Y.; Xiao, D.; Zang, Z.; Jiao, Z.; Chen, Y.; Li, D.D.U. Simple and Robust Deep Learning Approach for Fast Fluorescence Lifetime Imaging. *Sensors* **2022**, *22*, 7293. <https://doi.org/10.3390/s22197293>

Academic Editors: Sidong Liu, Roneel V. Sharan and Hao Xiong

Received: 31 August 2022

Accepted: 22 September 2022

Published: 26 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Fluorescence lifetime imaging (FLIM) is a powerful technique in biomedical research [1,2], such as probing cellular microenvironments and physiological parameters, including pH, viscosity, temperature, and ion concentrations (e.g., Ca^{2+} , O_2) [3–7]. In combination with Förster resonance energy transfer (FRET), FLIM-FRET has been widely used as a “quantum ruler” to quantify protein–protein conformations and interactions [8,9]. Fluorescence lifetimes are local properties of fluorophores depending only on the physicochemical state of the local microenvironment (e.g., pH, ionic strength). They are free of artifacts due to fluctuations in laser power, optical path, and fluorophore concentrations. There are time and frequency domain approaches for measuring fluorescence lifetimes [10,11]. Among them, time-correlated single-photon counting (TCSPC) [12] has been widely used due to its superior photon efficiency, high signal-to-noise ratio, and superior temporal resolution [13]. TCSPC records photon arrival times after shining laser pulses and yields decay histograms, from which fluorescence lifetimes can be extracted. Theoretically, a fluorescence decay histogram can be expressed as [14]:

$$y(t) = \sum_{i=1}^n \alpha_i \exp\left(-\frac{t}{\tau_i}\right), \quad \sum_{i=1}^n \alpha_i = 1, \quad (1)$$

where n is the number of components, and α_i and τ_i are the fraction and the lifetime of the i -th component. Curve-fitting methods have been commonly used for FLIM analysis, including the nonlinear least-square method (NLSM) [15], maximum likelihood estimation (MLE) [16], Bayesian analysis [17], and Laguerre expansion methods [18]. However, they require many photons to deliver accurate results. They are time-consuming and limited to offline analysis. Moreover, they require prior knowledge to set proper initial parameters. Fitting-free methods have also been developed, such as the phasor approach [19] and the

center-of-mass method (CMM) [20]. The phasor approach offers a fast graphic interpretation [21,22]. CMM [20,23] offers fast analysis, but it can only provide intensity-weighted average lifetimes [14].

Deep learning (DL) techniques have been proven promising for FLIM analysis [24,25]. It is to extract high dimensional features of input decay histograms and then map the input to the lifetime parameters. In contrast, the Bayesian estimation is to estimate the posterior distribution from the chosen prior distribution and likelihood function. Usually, the posterior distribution can be obtained analytically or approximately using Markov chain Monte Carlo (MCMC) methods. A general structure of a deep-learning-based predictive model is shown in Figure 1a, including the model training from training data in phase 1 and model testing from new data in phase 2. Various neural networks were proposed for multiexponential analysis, including multilayer perceptions (MLPs) [26], one- and high-dimensional convolutional neural networks (CNNs) [27], online training extreme learning machine [28], and generative adversarial networks (GANs) [29]. MLP is simple and can be easily implemented on different platforms, such as CPU, GPU, or other embedded systems. According to the general approximation theorem, MLP can approximate any continuous function. However, it shows poor feature extraction capacity. Therefore, MLP-based algorithms cannot resolve multiexponential decay models [26]. A feasible way to address this problem is adding feature engineering as network input. For example, phasor coordinates can be considered to improve the resolvability of multiexponential analysis. CNN is powerful for feature extraction, such as local parameter sharing, sparse interactions, and equivariant representations. High-dimensional CNNs have been successfully applied for multiexponential FLIM analysis with high precision [27]. Meanwhile, the light-weighted one-dimensional (1D) CNN has been proposed for FLIM analysis [30]. The 1D-CNN features high efficiency, fast training, and high inference speed. It is also hardware friendly to be applied in embedded systems. However, CNN heavily relies on convolution operators and requires hyperparameter optimization, and its performance is significantly affected by the size of convolutional kernels, because CNNs only have a local field-of-view with a convolutional kernel, which cannot simultaneously learn the complete decay information.

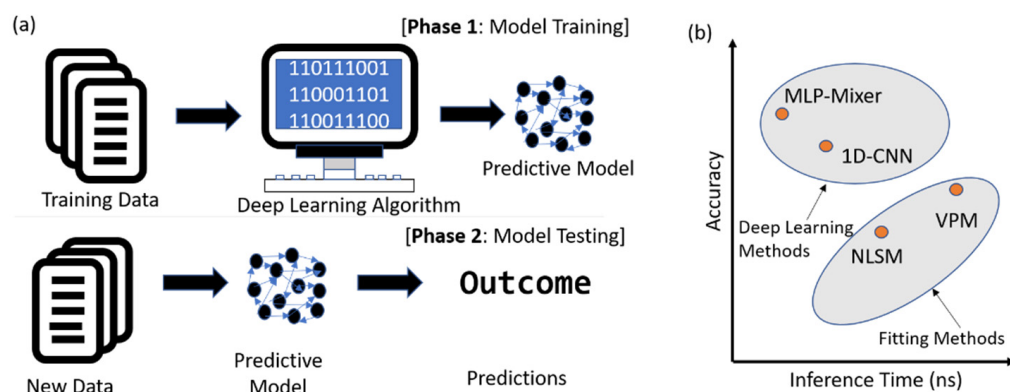


Figure 1. (a) Logic of the deep learning algorithm based on a predictive model considering training and testing phases. (b) Comparison of fluorescence lifetime prediction methods related to inference time and accuracy for one FLIM image.

We proposed a new MLP-based DL algorithm to address the above challenges. Our algorithm adopts the MLP-Mixer [31] to provide fast and accurate analysis even under a low signal-to-noise ratio (SNR). Inspired by CNNs and transformers, the MLP-Mixer has an isotropic design with only MLP layers, which is easy to implement. Thus, our algorithm is particularly suitable for embedded systems. Compared with CNNs, the MLP-Mixer has three distinct advantages. First, it has higher computational efficiency because only matrix multiplications are involved. Second, the MLP-Mixer is more suitable for analyzing sequence signals. It has a global field of view, processing the whole decay sequence at the same time. Third, the implementation and optimization of the MLP-Mixer are simple and

suitable for broad applications. We demonstrate that the MLP-Mixer is efficient, robustly solving simulated and experimental FLIM data. We compared the proposed method with 1D-CNN, NLSM, and VPM in terms of accuracy and inference time for one FLIM image; see Figure 1b for an intuitive summary. The results indicate that the MLP-Mixer outperforms traditional methods with less bias, especially when the fluorescent samples have small lifetime components. The inference time for one FLIM image is around 163-fold and 900-fold faster than NLSM and the variable projection method (VPM) [32], respectively.

2. Methods and Theory

2.1. Architecture Design

Figure 2a shows the topological structure of the proposed the FLIM-MLP-Mixer. The input histogram is first divided into nonoverlapping patches. The patch size is set to 10 for capturing more general features of decay curves while maintaining a moderate temporal resolution. All patches are fed into a mixer block simultaneously. The core module of the network is the mixer, which consists of layers that mix features (i) at a given spatial location and (ii) between different spatial locations. For the mixer design, there are two MLP blocks. One is the token-mixing MLP, in which an input (batches \times patches \times channels) is accepted by a norm layer and then transposed to a matrix (batches \times channels \times patches). The data are then passed through the MLP1 block, which consists of two fully connected layers with a nonlinear activate function (GELU) [33]. The output of the MLP1 block is transposed to a matrix (batches \times patches \times channels) for the next MLP block, which is the channel-mixing MLP block. This MLP block (MLP2) is responsible for extracting patch features. In either token-mixing or channel-mixing MLP block, skip connections [34] are employed. After the mixer, there are three dense layers for reconstructing average lifetimes. The exhaustive searching method is used to optimize the hyperparameters of the neural networks.

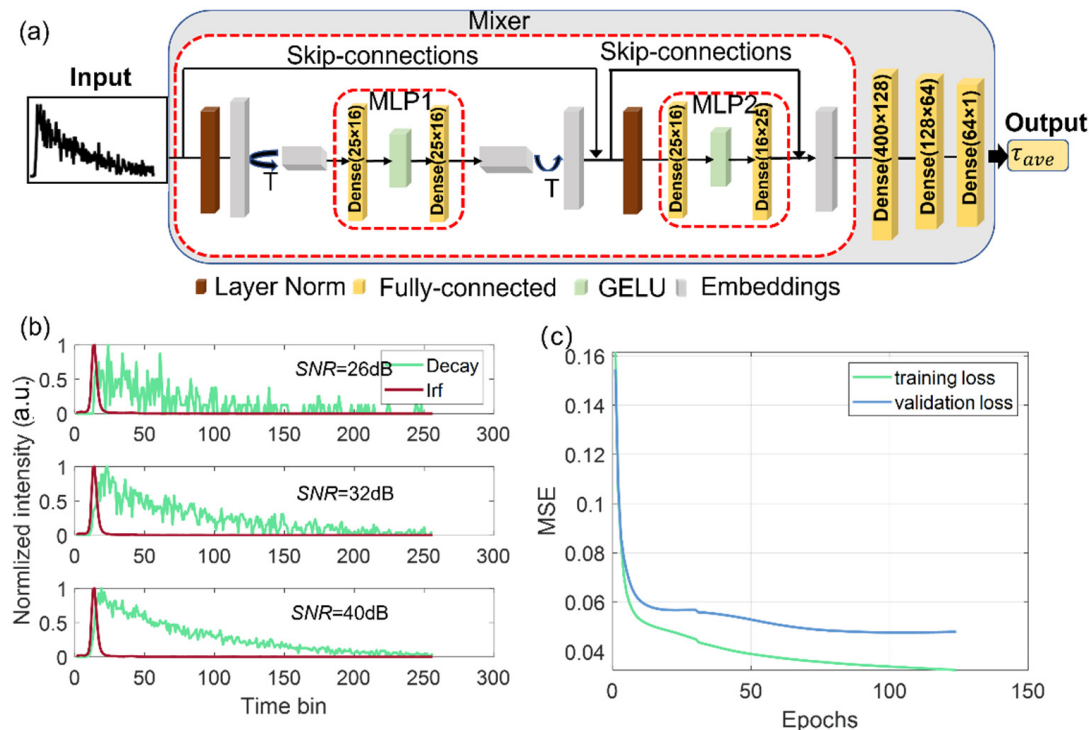


Figure 2. (a) MLP-Mixer architecture. The MLP-Mixer consists of per-patch linear embeddings, mixer layers, and a regression layer. The mixer block contains one token-mixing MLP (MLP1) and one channel-mixing MLP (MLP2), each consisting of two fully connected layers and a GELU nonlinearity. Other components include skip connections, dropout, and layer norm on channels. (b) Decays under different SNRs, the green curve stands for decay, and the red curve means IRF. (c) Mean square error (MSE) plots of the training and validation loss for τ_{ave} .

The idea behind the mixer is to separate the per-location (channel-mixing) operations and cross-location (token-mixing) operations. The skip connections can avoid gradient vanishing problems, leading to faster convergence and better performance. Our network has advantages with a light-weighted structure. The mixer relies on basic matrix multiplication routines, which require fewer hardware resources, including smaller memory size and less power consumption. It is much faster without sophisticated matrix operations. We can quickly implement hardware-friendly MLP neural networks in field-programmable gate array (FPGA) devices and integrate them into a wide-field SPAD FLIM system. In the training process, we use one block for the MLP-Mixer, in which the dimension of the token and channels are 16, and the patch size is 10.

2.2. Training Dataset Preparation

Training data are critical for neural networks. FLIM training data can be easily obtained using synthetic data since mathematical FLIM decay and noise models are well developed. A theoretical fluorescence signal measured by a FLIM system can be described as [30]:

$$y(t) = IRF(t) * A \sum_{i=1}^n \alpha_i e^{-\frac{t}{\tau_i}} + \varepsilon(t), \quad \sum_{i=1}^n \alpha_i = 1, \quad (2)$$

where A is the amplitude of the underlying fluorescence decay, α_i is the i -th fraction ratio, τ_i is the i -th lifetime component, and $\varepsilon(t)$ is the Poisson noise, which is dominant in a TCSPC system. $IRF(\cdot)$ is the FLIM instrument response function, approximated by a Gaussian function with FWHM = 167 ps. The asterisk (*) refers to a convolution operator. The total photon counts $Y = \int y(t)dt$ of synthetic signals range from 100 to 1×10^4 . The SNR of a signal is defined as:

$$SNR = 20 * \log \sqrt{Y}, \quad (3)$$

In most FLIM applications, signals' underlying decay model is unknown; it is often valuable for determining average lifetimes rather than lifetime components, which are suspicious for biological interpretation. Amplitude-weighted average lifetimes can estimate FRET efficiency and access dynamic quenching behaviors, which are of great interest [23]. Therefore, the FLIM-MLP-Mixer is designed for evaluating amplitude-weighted average lifetimes, defined as [14]:

$$\tau_A = \sum_{i=1}^n \alpha_i \tau_i, \quad (4)$$

2.3. Training and Evaluation

The proposed neural network was deployed in PyTorch 1.0 [35]. The loss function is defined as:

$$L(\Theta) = \frac{1}{N} \sum_{i=1}^N \|F(X^i, \Theta) - Y^i\|^2, \quad (5)$$

where X is the input and Y is the corresponding label of the signal. F is the mapping function with the network parameters Θ , and N is the batch size of the training dataset. The optimizer uses the Adam algorithm with a learning rate of 1×10^{-4} in the standard back-propagation [36]. The training dataset contains 40,000 different signals and their corresponding lifetimes. Both mono- and biexponential decays were included in training datasets. For monoexponential signals, the lifetime range is [0.1, 5] ns, and for biexponential signals, τ_1 and τ_2 are set in [0.1, 1], [1, 5] ns. The trained lifetimes cover a wide range of lifetimes of commonly used fluorophores for biomedical applications. Figure 2b shows the signals under different SNRs, the green curve stands for decay, and the red curve means IRF.

The FWHM of IRFs has a Gaussian distribution with a mean value of 167 ps and a standard deviation of 60 ps. The training batch size is 200, and 20% of the total datasets are used for validation. An early stop callback with 20 patients is included to prevent overfitting. All the datasets are normalized to 1 before being fed into the neural network. Figure 2c shows the training and validation errors of the MLP-Mixer. The mean square errors (MSEs) of both training and validation losses decrease rapidly and reach the plateau

after 120 epochs. They converge at a small value of 1.8×10^{-4} , indicating that the network is well trained as the estimated lifetimes are close to their ground truths. Because of the simple network structure, the number of hyperparameters is only 26,081. The training time is half an hour on a typical desktop using a central processing unit (CPU).

3. Simulation Results

The MLP-Mixer was first evaluated on synthetic monoexponential signals with lifetimes of 1 and 4 ns. The photon counts range from 100 to 10,000, with 20 to 40 dB SNRs. We define the relative bias as $\text{Bias} = |\tau_{pre} - \tau_{truth}| / \tau_{truth}$, where τ_{pre} and τ_{truth} are the estimated τ_A by the methods and label of τ_A from Equation (4), respectively. We define the standard deviation, $\text{Std} = \sqrt{\frac{\sum_{i=1}^n (\tau_A^i - \tau_{truth})^2}{(n-1)}}$, where τ_A^i is the predicted parameter, and n is the number of simulated decay curves. Figure 3a,b shows the violin plots of the predicted results of the simulated data. The red dotted line stands for the ground truth of lifetimes. The result of the MLP-Mixer is closer to the ground truth, reaching 0.95, 0.96, and 0.98 ns at SNRs 20~32 dB, 32~38 dB, and 38~40 dB, respectively, for signals with a lifetime of 1 ns. Figure 3c,d is the corresponding bias and standard deviation plots. Both figures show that the MLP-Mixer and 1D-CNN outperform the traditional methods NLSM and VPM in terms of accuracy under different SNRs for both short and long lifetimes. The biases of the estimation with the MLP-Mixer are lower than 2% and 5% for short and long lifetimes, meaning that the MLP-Mixer is a more accurate estimator. In addition, as shown in Figure 3c, the standard deviations of all methods for short lifetimes are less than 0.06, which is much smaller than those shown in Figure 3d, which are above 0.1, indicating that methods for short lifetimes are more precise than long lifetimes under the same photon counts and SNRs. This is because, in TCSPC, the noise is governed by Poisson distribution, so the SNR is proportional to the square root of total photon counts. For a given photon count, for example, 2000 p.c., a larger lifetime means that photons are distributed in more time bins. Therefore, each time bin has fewer photons, and the SNR is smaller, so larger lifetimes yield a more significant deviation. In conclusion, the MLP-Mixer is more accurate than 1D-CNN, NLSM, and VPM with acceptable precision under different SNRs.

We also evaluate our method on histograms following a biexponential decay model. The short and long lifetime components, τ_1 and τ_2 , range from 0.5 to 1.5 ns and from 2.5 to 3.5 ns, respectively. α equals 0.2, 0.5, and 0.8. We compared the inference performances of the MLP-Mixer, 1D-CNN, NLSM, and VPM for amplitude-weighted lifetimes. In this simulation, 2000 simulated testing datasets were generated using Equation (4). Figure 4 shows the results for SNR 26 dB, 34 dB, and 38 dB and α ranging from 0.2 to 0.8. From Figure 4a,d,g, we can see that the MLP-Mixer has a low bias of 0.037, 0.057, and 0.006 at $\alpha = 0.2, 0.5$, and 0.8 , respectively, and NLSM and VPM have bias over 0.1 under the same conditions, indicating that the MLP-Mixer outperforms NLSM and VPM. It is obvious that the bias of the MLP-Mixer at $\alpha = 0.8$ is much smaller than that of 1D-CNN, NLSM, and VPM at $\alpha = 0.2, 0.5$, because larger α means the short lifetime components in the decay are dominant, which have a good agreement with the discussion in Figure 3. Moreover, NLSM and VPM are more sensitive to SNRs than the MLP-Mixer and 1D-CNN. However, when $\alpha = 0.5$, the MLP-Mixer and 1D-CNN have a very similar result.

Table 1 summarizes the main characteristic parameters of the MLP-Mixer and 1D-CNN. The number of hyperparameters of the MLP-Mixer is threefold more significant than that of 1D-CNN, whereas the training time of the MLP-Mixer (1871.0 s) is slightly less than that of 1D-CNN (2270.62 s). The total numbers of floating-point operations (FLOPs) for a single sample are very similar, 5.9 billion and 5.4 billion for the MLP-Mixer and 1D-CNN, respectively. The inference time of 1D-CNN is 34.06 s, whereas the MLP-Mixer only consumed 12.54 s, showing that the efficiency of the MLP-Mixer is better than 1D-CNN.

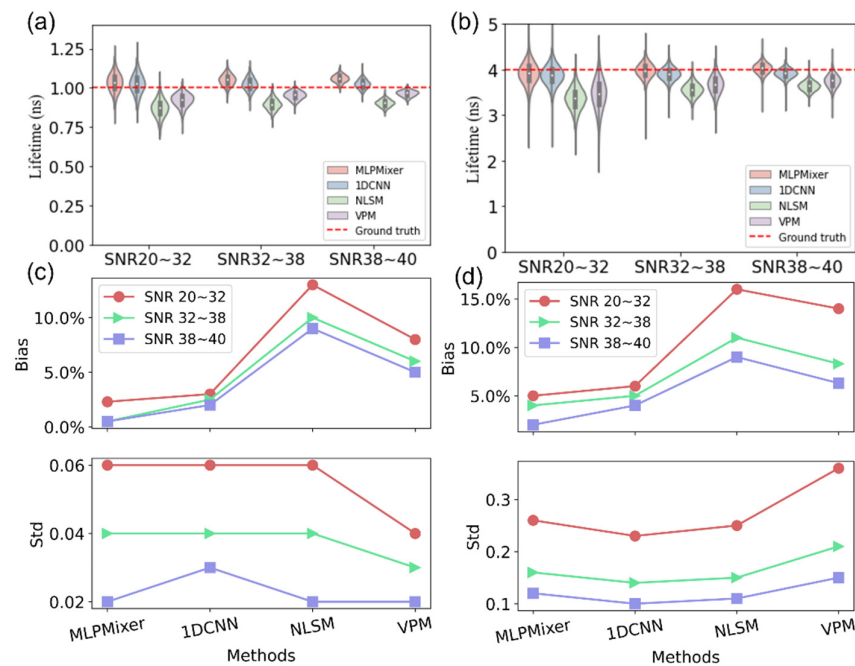


Figure 3. Boxplots of estimations under different signal-to-noise ratios (SNRs) with the four methods, in which the median and quartiles with whiskers reach up to 1.5 times the interquartile range. The violin plot outlines illustrate kernel probability density; for example, the width of the shaded area represents the proportion of the data. (a) Short lifetime (set at 1 ns) and (b) long lifetime (set at 4 ns) estimations from synthetic data; (c,d) corresponding bias and standard deviation plots for (a) short and (b) long lifetimes.

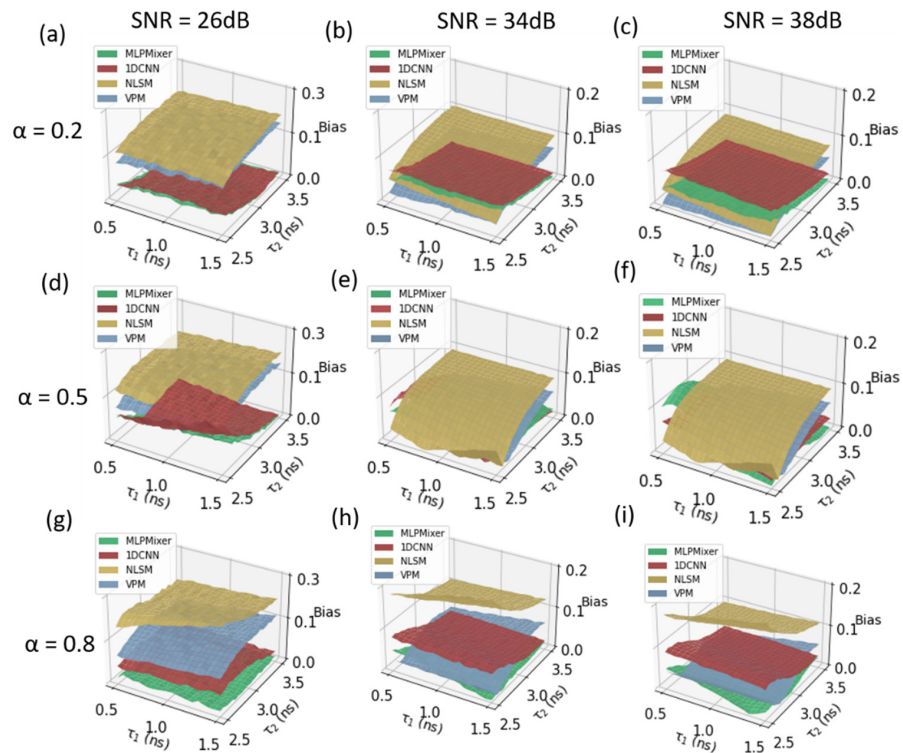


Figure 4. τ_A estimation error planes of the MLP-Mixer, 1D-CNN, NLSM, and VPM at different fraction ratios α under different signal-to-noise ratios. For signals following a biexponential decay model where $\tau_1 \in (0.5, 1.5)$ ns, $\tau_2 \in (2.5, 3)$ ns. The error planes (a,d,g) SNR 26 dB, (b,e,h) SNR 34 dB, (c,f,i) SNR 38 dB, and $\alpha = 0.2, 0.5, 0.8$.

Table 1. Main characteristics of the MLP-Mixer and 1D-CNN lifetime processor.

Algorithm	Total Parameters	FLOPs	Total Training Time (s)	Inference Time
MLP-Mixer	61,306	5.9 B	1871.0	12.54 s
1D-CNN	17,409	5.4 B	2270.65	34.06 s

4. Experimental Results and Discussion

Dye solutions and plant cell samples were used for experimental validation. FLIM datasets were obtained with a commercial two-photon FLIM system.

4.1. Rhodamine 6G and Rhodamine B Solutions

Rhodamine B and 6G saturated aqueous solutions were placed on a glass slide with coverslips. The FLIM system consists of a confocal microscope (LSM 510, Carl Zeiss, Oberkochen, Germany) and a TCSPC card (SPC-830, Becker & Hickl GmbH, Berlin, Germany). The femtosecond Ti: sapphire laser (Chameleon, Coherent, Santa Clara, CA, USA) has an excitation wavelength of 800 nm and a pulse width of less than 200 fs with an 80 MHz repetition rate. The bin width of TCSPC is 0.039 ns, and each measured histogram contains 256-time bins. The emission light passes through a 500–550 nm bandpass filter and then is collected by a 60× water-immersion objective lens (N.A = 1.0). We experimented on rhodamine 6G and rhodamine B at different SNRs to mimic low-light conditions. The results are shown in Table 2.

Table 2. Rhodamine B and 6G lifetimes with the MLP-Mixer, 1D-CNN, NLSM, and VPM.

Rhodamine B		SNR 26~32	SNR 32~38	SNR 38~44	τ_{REF} (ns)
MLP-Mixer	τ_{mean} (ns)	1.70	1.73	1.71	1.72 [37]
	Std	0.40	0.46	0.50	
	Bias	1.16%	0.58%	0.58%	
1D-CNN	τ_{mean} (ns)	1.68	1.65	1.69	
	Std	0.59	0.54	0.52	
	Bias	2.33%	4.07%	1.74%	
NLSM	τ_{mean} (ns)	1.55	1.61	1.63	
	Std	0.39	0.24	0.27	
	Bias	9.88%	6.40%	5.32%	
VPM	τ_{mean} (ns)	1.62	1.64	1.67	
	Std	0.52	0.47	0.45	
	Bias	5.81%	4.56%	2.96%	
Rhodamine 6G		SNR 26~32	SNR 32~38	SNR 38~44	τ_{REF} (ns)
MLP-Mixer	τ_{mean} (ns)	4.00	4.04	4.07	4.06 [38]
	Std	0.59	0.46	0.62	
	Bias	1.48%	0.5%	0.25%	
1D-CNN	τ_{mean} (ns)	4.22	3.93	3.94	
	Std	0.78	0.7	0.72	
	Bias	3.94%	3.20%	2.95%	
NLSM	τ_{mean} (ns)	3.58	3.68	3.75	
	Std	0.53	0.27	0.34	
	Bias	11.82%	9.36%	7.64%	
VPM	τ_{mean} (ns)	3.79	3.89	3.92	
	Std	0.87	0.61	0.48	
	Bias	6.65%	4.19%	3.45%	

Table 2 summarizes the results calculated by the MLP-Mixer, NLSM, VPM, and 1D-CNN. τ_{REF} is the reference lifetime of the tested samples (rhodamine 6G and rho-

damine B) [37,38]. In our study, a threshold (10% of total counts) was applied to mask pixels with low photon counts. The MLP-Mixer and 1D-CNN outperform NLSM and VPM in terms of accuracy. Results show that the MLP-Mixer is better than others, even under a low SNR. That is because a lower SNR means a lower photon count, which is not enough for fitting methods.

4.2. *Convallaria majalis* Cells

We further evaluated the MLP-Mixer on *Convallaria majalis* cell samples. The sample was measured with ‘high photon counts’ (HPC) at a 30 s acquisition time, ‘middle photon counts’ (MPC) at a 15 s acquisition time, and ‘low photon counts’ (LPC) at a 3 s acquisition time. Figure 5 shows the results for three cases. LPC data analysis is challenging for NLSM due to the low SNR. The performances of these results were evaluated via the structural similarity index (SSIM) [39], summarized in Table 3. The MLP-Mixer significantly outperforms NLSM, especially in low SNR situations, where the SSIM values are 0.95 and 0.81 for the MLP-Mixer and NLSM, respectively. From the lifetime images, we can see that Figure 5e,i is very similar to Figure 5h,l, indicating that the MLP-Mixer and 1D-CNN performed well even at low counts. That is because both DL algorithms extract the features of decay histograms in high dimensional space, making the algorithms insensitive to different SNR levels. In Figure 5x, the lifetime distributions show that all methods obtain similar results at a high SNR.

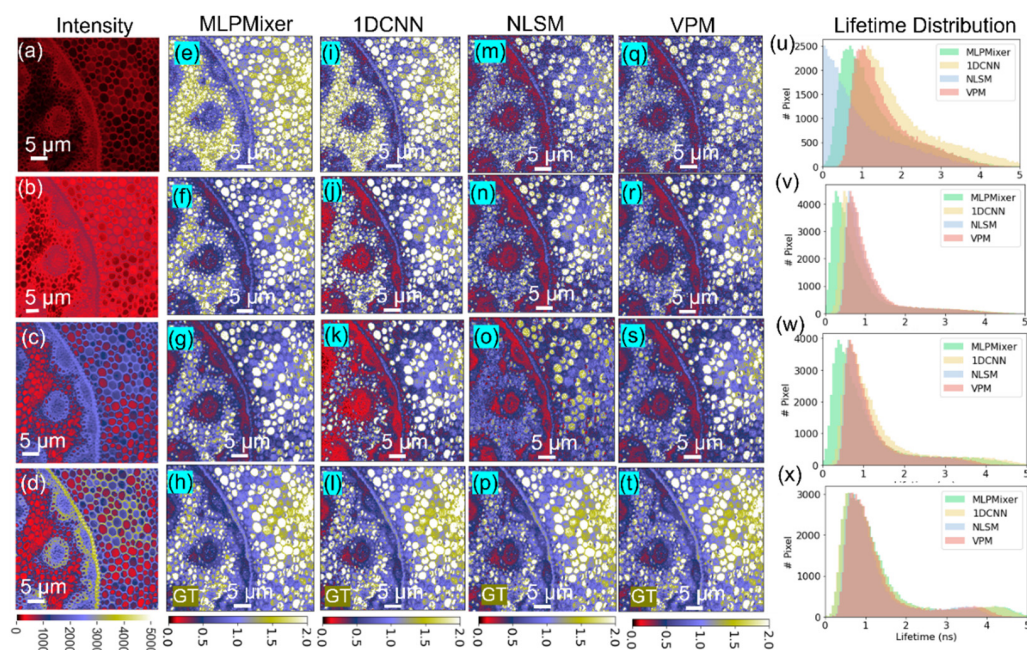


Figure 5. Two-photon FLIM experiments of *Convallaria majalis* cells. (a–d) Intensity images under different SNRs (3 s, 15 s, 30 s, and 90 s acquisition time). Lifetime images recovered from (e–h) the MLP-Mixer, (i–l) 1D-CNN, (m–p) NLSM, and (q–t) VPM. (u–x) Lifetime histograms for each case.

Table 3. SSIM values under different photon counts in Figure 5.

Data	MLP-Mixer	1D-CNN	NLSM	VPM
HPC	0.98	0.95	0.94	0.95
MPC	0.96	0.94	0.91	0.94
LPC	0.95	0.93	0.81	0.93

5. Conclusions

In summary, we present the design and training of the MLP-Mixer for FLIM analysis. The proposed algorithm has a simple architecture, which is easy to implement and hardware

friendly. The simulated FLIM data analysis shows that the MLP-Mixer performs better, especially for small lifetime components. The trained model was then employed to analyze two-photon FLIM images of rhodamine 6G, rhodamine B, and *Convallaria majalis* cells. The results also suggest the MLP-Mixer's excellent performance and robustness. Due to its simple architecture, it has the potential to be implemented in FPGA devices to accelerate the analysis for real-time FLIM applications.

Author Contributions: Conceptualization and methodology, Q.W. and D.X.; software, Q.W., D.X. and Y.L.; validation, Q.W.; formal analysis, Q.W.; investigation, Q.W.; writing—original draft preparation, Q.W.; writing—review and editing, Q.W., D.X., Y.L., Y.C. and D.D.U.L.; visualization, Q.W., Z.J. and Z.Z.; supervision, D.D.U.L.; project administration, D.D.U.L.; funding acquisition, D.D.U.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) under grant number BB/V019643/1.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data underlying the results presented in this paper are not publicly available but may be obtained from the authors upon reasonable request.

Acknowledgments: We would like to thank students Daniel Doveiko and Natakorn Sapermsap, Department of Physics, University of Strathclyde, for their support in dye solution preparation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Berezin, M.Y.; Achilefu, S. Fluorescence Lifetime Measurements and Biological Imaging. *Chem. Rev.* **2010**, *110*, 2641–2684. [[CrossRef](#)] [[PubMed](#)]
2. Suhling, F.; Sahlén, G.; Gorb, S.; Kalkman, V.J.; Dijkstra, K.-D.B.; van Tol, J. Order Odonata. In *Thorpe and Covich's Freshwater Invertebrates*; Elsevier: Amsterdam, The Netherlands, 2015; pp. 893–932.
3. Ning, Y.; Cheng, S.; Wang, J.-X.; Liu, Y.-W.; Feng, W.; Li, F.; Zhang, J.-L. Fluorescence Lifetime Imaging of Upper Gastrointestinal PH in Vivo with a Lanthanide Based near-Infrared τ Probe. *Chem. Sci.* **2019**, *10*, 4227–4235. [[CrossRef](#)]
4. Agronskaia, A.V.; Tertoolen, L.; Gerritsen, H.C. Fast Fluorescence Lifetime Imaging of Calcium in Living Cells. *J. Biomed. Opt.* **2004**, *9*, 1230. [[CrossRef](#)] [[PubMed](#)]
5. Zheng, K.; Jensen, T.P.; Rusakov, D.A. Monitoring Intracellular Nanomolar Calcium Using Fluorescence Lifetime Imaging. *Nat. Protoc.* **2018**, *13*, 581–597. [[CrossRef](#)] [[PubMed](#)]
6. Hosny, N.A.; Lee, D.A.; Knight, M.M. Single Photon Counting Fluorescence Lifetime Detection of Pericellular Oxygen Concentrations. *J. Biomed. Opt.* **2012**, *17*, 016007. [[CrossRef](#)] [[PubMed](#)]
7. Peng, X.; Yang, Z.; Wang, J.; Fan, J.; He, Y.; Song, F.; Wang, B.; Sun, S.; Qu, J.; Qi, J.; et al. Fluorescence Ratiometry and Fluorescence Lifetime Imaging: Using a Single Molecular Sensor for Dual Mode Imaging of Cellular Viscosity. *J. Am. Chem. Soc.* **2011**, *133*, 6626–6635. [[CrossRef](#)]
8. Loura, L. FRET in Membrane Biophysics: An Overview. *Front. Physiol.* **2011**, *2*, 82. [[CrossRef](#)] [[PubMed](#)]
9. Bajar, B.; Wang, E.; Zhang, S.; Lin, M.; Chu, J. A Guide to Fluorescent Protein FRET Pairs. *Sensors* **2016**, *16*, 1488. [[CrossRef](#)] [[PubMed](#)]
10. Lakowicz, J.R.; Szmajda, H.; Nowaczyk, K.; Johnson, M.L. Fluorescence Lifetime Imaging of Calcium Using Quin-2. *Cell Calcium* **1992**, *13*, 131–147. [[CrossRef](#)]
11. Yahav, G.; Gershanov, S.; Salmon-Divon, M.; Ben-Zvi, H.; Mircus, G.; Goldenberg-Cohen, N.; Fixler, D. Pathogen Detection Using Frequency Domain Fluorescent Lifetime Measurements. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 2731–2741. [[CrossRef](#)] [[PubMed](#)]
12. Becker, W. *Advanced Time-Correlated Single Photon Counting Techniques*; Springer Series in Chemical Physics; Springer: Berlin/Heidelberg, Germany; New York, NY, USA, 2005; ISBN 978-3-540-26047-9.
13. Peter, M.; Ameer-Beg, S.M. Imaging Molecular Interactions by Multiphoton FLIM. *Biol. Cell* **2004**, *96*, 231–236. [[CrossRef](#)] [[PubMed](#)]
14. Li, Y.; Natakorn, S.; Chen, Y.; Safar, M.; Cunningham, M.; Tian, J.; Li, D.D.-U. Investigations on Average Fluorescence Lifetimes for Visualizing Multi-Exponential Decays. *Front. Phys.* **2020**, *8*, 576862. [[CrossRef](#)]
15. Periasamy, N. Analysis of Fluorescence Decay by the Nonlinear Least Squares Method. *Biophys. J.* **1988**, *54*, 961–967. [[CrossRef](#)]
16. Bajzer, Ž.; Therneau, T.M.; Sharp, J.C.; Prendergast, F.G. Maximum Likelihood Method for the Analysis of Time-Resolved Fluorescence Decay Curves. *Eur. Biophys. J.* **1991**, *20*, 247–262. [[CrossRef](#)]

17. Rowley, M.I.; Barber, P.R.; Coolen, A.C.C.; Vojnovic, B. *Bayesian Analysis of Fluorescence Lifetime Imaging Data*; Periasamy, A., König, K., So, P.T.C., Eds.; SPIE: San Francisco, CA, USA, 2011; p. 790325.
18. Jo, J.A.; Fang, Q.; Marcu, L. Ultrafast Method for the Analysis of Fluorescence Lifetime Imaging Microscopy Data Based on the Laguerre Expansion Technique. *IEEE J. Select. Top. Quantum Electron.* **2005**, *11*, 835–845. [\[CrossRef\]](#)
19. Stringari, C.; Cinquin, A.; Cinquin, O.; Digman, M.A.; Donovan, P.J.; Gratton, E. Phasor Approach to Fluorescence Lifetime Microscopy Distinguishes Different Metabolic States of Germ Cells in a Live Tissue. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13582–13587. [\[CrossRef\]](#)
20. Li, D.D.-U.; Arlt, J.; Tyndall, D.; Walker, R.; Richardson, J.; Stoppa, D.; Charbon, E.; Henderson, R.K. Video-Rate Fluorescence Lifetime Imaging Camera with CMOS Single-Photon Avalanche Diode Arrays and High-Speed Imaging Algorithm. *J. Biomed. Opt.* **2011**, *16*, 096012. [\[CrossRef\]](#)
21. Redford, G.I.; Clegg, R.M. Polar Plot Representation for Frequency-Domain Analysis of Fluorescence Lifetimes. *J. Fluoresc.* **2005**, *15*, 805–815. [\[CrossRef\]](#)
22. Clayton, A.H.A.; Hanley, Q.S.; Verveer, P.J. Graphical Representation and Multicomponent Analysis of Single-Frequency Fluorescence Lifetime Imaging Microscopy Data. *J. Microsc.* **2004**, *213*, 1–5. [\[CrossRef\]](#)
23. Zhang, C.; Li, P.; Sun, G.; Guan, Y.; Xiao, B.; Cong, J. Optimizing FPGA-Based Accelerator Design for Deep Convolutional Neural Networks. In Proceedings of the 2015 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays, Monterey, CA, USA, 22–24 February 2015; ACM: Monterey, CA, USA; pp. 161–170.
24. Belthangady, C.; Royer, L.A. Applications, Promises, and Pitfalls of Deep Learning for Fluorescence Image Reconstruction. *Nat. Methods* **2019**, *16*, 1215–1225. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Wu, G.; Nowotny, T.; Zhang, Y.; Yu, H.-Q.; Li, D.D.-U. Artificial Neural Network Approaches for Fluorescence Lifetime Imaging Techniques. *Opt. Lett.* **2016**, *41*, 2561. [\[CrossRef\]](#)
26. Zickus, V.; Wu, M.-L.; Morimoto, K.; Kapitany, V.; Fatima, A.; Turpin, A.; Insall, R.; Whitelaw, J.; Machesky, L.; Bruschini, C.; et al. Fluorescence Lifetime Imaging with a Megapixel SPAD Camera and Neural Network Lifetime Estimation. *Sci. Rep.* **2020**, *10*, 20986. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Smith, J.T.; Yao, R.; Sinsuebphon, N.; Rudkouskaya, A.; Un, N.; Mazurkiewicz, J.; Barroso, M.; Yan, P.; Intes, X. Fast Fit-Free Analysis of Fluorescence Lifetime Imaging via Deep Learning. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 24019–24030. [\[CrossRef\]](#)
28. Zang, Z.; Xiao, D.; Wang, Q.; Li, Z.; Xie, W.; Chen, Y.; Li, D.D.U. Fast Analysis of Time-Domain Fluorescence Lifetime Imaging via Extreme Learning Machine. *Sensors* **2022**, *22*, 3758. [\[CrossRef\]](#)
29. Chen, Y.-I.; Chang, Y.-J.; Liao, S.-C.; Nguyen, T.D.; Yang, J.; Kuo, Y.-A.; Hong, S.; Liu, Y.-L.; Rylander, H.G.; Santacruz, S.R.; et al. Generative Adversarial Network Enables Rapid and Robust Fluorescence Lifetime Image Analysis in Live Cells. *Commun. Biol.* **2022**, *5*, 18. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Xiao, D.; Chen, Y.; Li, D.D.-U. One-Dimensional Deep Learning Architecture for Fast Fluorescence Lifetime Imaging. *IEEE J. Select. Top. Quantum Electron.* **2021**, *27*, 1–10. [\[CrossRef\]](#)
31. Tolstikhin, I.; Houlsby, N.; Kolesnikov, A.; Beyer, L.; Zhai, X.; Unterthiner, T.; Yung, J.; Steiner, A.; Keysers, D.; Uszkoreit, J.; et al. MLP-Mixer: An All-MLP Architecture for Vision 2021. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 24261–24272.
32. Zhang, Y.; Cuyt, A.; Lee, W.; Lo Bianco, G.; Wu, G.; Chen, Y.; Li, D.D.-U. Towards Unsupervised Fluorescence Lifetime Imaging Using Low Dimensional Variable Projection. *Opt. Express* **2016**, *24*, 26777. [\[CrossRef\]](#)
33. Hendrycks, D.; Gimpel, K. Gaussian Error Linear Units (GELUs) 2020. *arXiv* **2020**, arXiv:1606.08415.
34. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition 2015. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016.
35. Available online: <http://pytorch.org/> (accessed on 27 October 2017).
36. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization 2017. *arXiv* **2014**, arXiv:1412.6980.
37. Kristoffersen, A.S.; Erga, S.R.; Hamre, B.; Frette, Ø. Testing Fluorescence Lifetime Standards Using Two-Photon Excitation and Time-Domain Instrumentation: Rhodamine B, Coumarin 6 and Lucifer Yellow. *J. Fluoresc.* **2014**, *24*, 1015–1024. [\[CrossRef\]](#)
38. Elder, A.D.; Frank, J.H.; Swartling, J.; Dai, X.; Kaminski, C.F. Calibration of a Wide-Field Frequency-Domain Fluorescence Lifetime Microscopy System Using Light Emitting Diodes as Light Sources. *J. Microsc.* **2006**, *224*, 166–180. [\[CrossRef\]](#)
39. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [\[CrossRef\]](#)