

<https://doi.org/10.1038/s44303-024-00021-7>

# Deep learning-based virtual H&E staining from label-free autofluorescence lifetime images

Check for updates

Qiang Wang<sup>1,2</sup>✉, Ahsan R. Akram<sup>1,2</sup>, David A. Dorward<sup>1,3</sup>, Sophie Talas<sup>1,3</sup>, Basil Monks<sup>3</sup>, Chee Thum<sup>3</sup>, James R. Hopgood<sup>4</sup>, Malihe Javidi<sup>5,6</sup> & Marta Vallejo<sup>5</sup>

Label-free autofluorescence lifetime is a unique feature of the inherent fluorescence signals emitted by natural fluorophores in biological samples. Fluorescence lifetime imaging microscopy (FLIM) can capture these signals enabling comprehensive analyses of biological samples. Despite the fundamental importance and wide application of FLIM in biomedical and clinical sciences, existing methods for analysing FLIM images often struggle to provide rapid and precise interpretations without reliable references, such as histology images, which are usually unavailable alongside FLIM images. To address this issue, we propose a deep learning (DL)-based approach for generating virtual Hematoxylin and Eosin (H&E) staining. By combining an advanced DL model with a contemporary image quality metric, we can generate clinical-grade virtual H&E-stained images from label-free FLIM images acquired on unstained tissue samples. Our experiments also show that the inclusion of lifetime information, an extra dimension beyond intensity, results in more accurate reconstructions of virtual staining when compared to using intensity-only images. This advancement allows for the instant and accurate interpretation of FLIM images at the cellular level without the complexities associated with co-registering FLIM and histology images. Consequently, we are able to identify distinct lifetime signatures of seven different cell types commonly found in the tumour microenvironment, opening up new opportunities towards biomarker-free tissue histology using FLIM across multiple cancer types.

Fluorescence lifetime is characterised by a decay from the excited state to the ground state, which is independent of its intensity but extremely sensitive to the surrounding bio-environment. Fluorescence lifetime imaging microscopy (FLIM) has been shown to provide valuable information about the underlying metabolic state, pathological conditions, and constitution of biological samples by analysing endogenous fluorescence<sup>1,2</sup>. By measuring the fluorescence lifetime of fluorescent molecules within a sample, FLIM can provide insights into various microenvironmental factors, including pH, ion concentration, and molecular interactions. Due to lifetime contrast, this has led to various applications in biology and medicine, such as cancer detection and diagnosis<sup>1,3</sup>, cellular phenotyping<sup>4,5</sup>, pH calculation of the secretory pathway<sup>6</sup>, and the monitoring of drug delivery<sup>7</sup>.

Conventionally, quantitative analysis of lifetime contrast is prevailed by statistical methods, such as histogram of FLIM images or phasor characterisation<sup>2,4,8</sup>, with the assistance of reference data, e.g., histological

images. Due to the nature of statistical analysis, most current approaches can only reveal averaged lifetime of dominant components in FLIM images, rather than cellular-level information. Although a few papers<sup>9,10</sup> have addressed this by co-registering FLIM and histology images together, there are still outstanding challenges associating with this co-registration, such as underlying structure change caused by tissue preparation<sup>11</sup>. In addition, such reference data is not always available alongside FLIM images, and therefore, unable to provide instant interpretation immediately after FLIM imaging procedures.

The explosive development of deep learning (DL) technologies is transforming conventional biomedical imaging analysis<sup>12,13</sup>, leading to automated processing that could surpass human capabilities. One of the beneficial fields is virtual histological staining from microscopic images acquired by various imaging modalities<sup>14</sup> using DL models, among which U-Net<sup>15</sup>-based Generative Adversarial Networks (GANs)<sup>16</sup> are prominent. Rivenson et al.<sup>17</sup> demonstrated the effectiveness of translating label-free

<sup>1</sup>Centre for Inflammation Research, Institute of Regeneration and Repair, The University of Edinburgh, Edinburgh, UK. <sup>2</sup>Translational Healthcare Technologies Group, Centre for Inflammation Research, Institute of Regeneration and Repair, The University of Edinburgh, Edinburgh, UK. <sup>3</sup>Department of Pathology, Royal Infirmary of Edinburgh, Edinburgh, UK. <sup>4</sup>School of Engineering, The University of Edinburgh, Edinburgh, UK. <sup>5</sup>School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK. <sup>6</sup>Department of Computer Engineering, Quchan University of Technology, Quchan, Iran. ✉e-mail: q.wang@ed.ac.uk

autofluorescence intensity images to multiple histology staining for different organs using a supervised GAN with an extra constraint of total variation. Li et al.<sup>18</sup> synthesised bright-field images on unstained carotid artery samples to Haematoxylin and Eosin (H&E), Picrosirius Red, and Orcein stain using the pix2pix GAN<sup>19</sup>. Borhani et al.<sup>20</sup> applied a custom DL model to generate virtual H&E stain from a combination of 16-channel two-photon excited fluorescence images with 1-channel lifetime images collected by multimodal microscopy on unstained tissue. Cao et al. applied an unsupervised GAN model, namely CycleGAN<sup>21</sup>, to transform ultraviolet photoacoustic microscopy to virtual H&E staining<sup>22</sup>. Other examples include transferring H&E to virtual IHC (immunohistochemistry) stain<sup>23</sup> and IHC to virtual multiplex immunofluorescence<sup>24</sup>.

In this study, we will explore the use of DL techniques to generate virtual H&E staining from label-free FLIM images. Specifically, we will employ the supervised DL model, pix2pix GAN<sup>19</sup>, along with the advanced image quality metric, Deep Image Structure and Texture Similarity (DISTs)<sup>25</sup>, to synthesise H&E-stained histology images. By doing so, we aim to demonstrate the potential of this technique in enabling rapid and precise interpretation at the cellular level of FLIM images, which could allow for the identification of label-free autofluorescence lifetime signatures for various types of tumour cells. Additionally, we will compare the effectiveness of different input data formats, such as intensity and weighted lifetime images, to showcase the benefits of using FLIM for the synthesis of virtual H&E staining. Meanwhile, we will also evaluate our method for multiple cancer types to demonstrate the advantages of FLIM-based virtual H&E staining.

## Results

### Virtual H&E staining from FLIM

Figure 1 summarises examples of the virtual H&E staining from label-free FLIM images using an independent dataset collected on two TMA (tissue

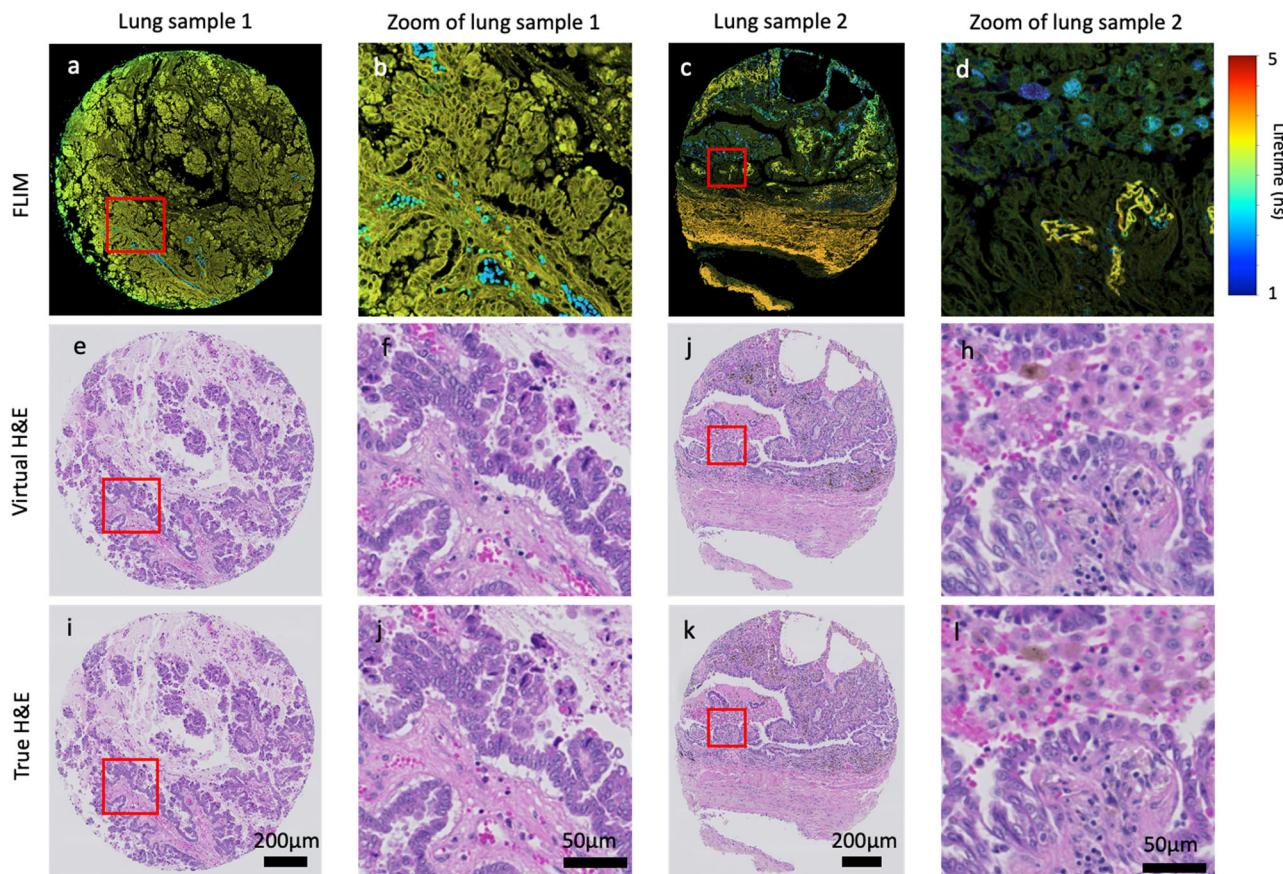
microarray) cores of lung cancer which were not utilised in training, where one contains relatively few distinct cell populations (Fig. 1a) while the other is a more complex mixture of cell types (Fig. 1c). Both are visualised as intensity-weighted false-colour lifetime images (see the section “Comparison of different image formats” for the details).

The results demonstrate that the method described in the section “Methods” is able to reconstruct various cellular components in the lung tissue samples from FLIM images to match the corresponding H&E-stained brightfield images. Specifically, the method proves its effectiveness in faithfully reproducing the morphological and textural attributes of a wide range of cell types present within the lung TME (tumour micro-environment), including tumour cells, stromal components (such as fibroblast's) and inflammatory cells, as well as red blood cells. Each of these was confirmed with the True HE (Fig. 1i–l).

While the DL model was trained using input patches sized at  $256 \times 256$ , it is important to note that these models can accommodate input of various sizes during the inference stage. In addition, H&E-stained images were downsampled to match the pixel size of FLIM images, where the pixel size was increased from  $0.22 \mu\text{m}$  to  $0.455 \mu\text{m}$  using bicubic interpolation.

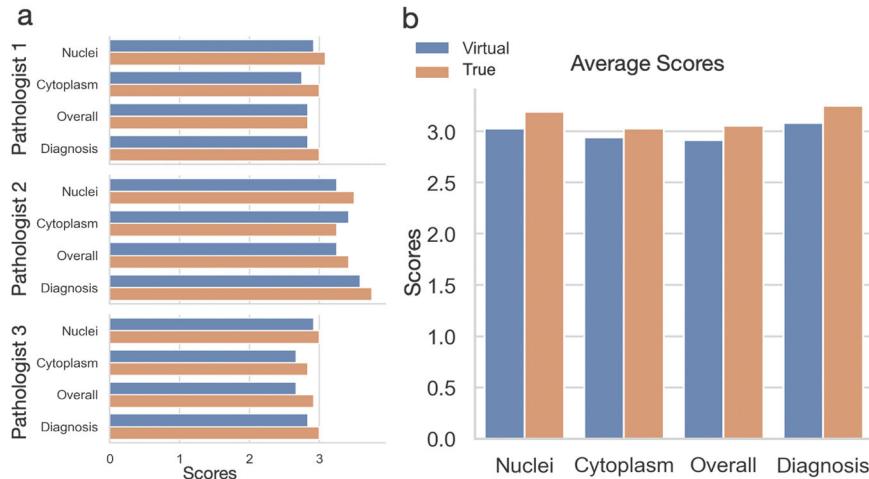
### Blind evaluation of virtual staining quality

In pathology laboratories, the visual inspection of histological images by experienced pathologists is an integral step in current clinical practice for reaching diagnostic decisions. Similar to refs. 17,26, a blind evaluation was conducted by three experienced pathologists to assess the subjective quality of virtual H&E staining. 12 pairs of true and virtual H&E-stained images were prepared using lung cancer tissues, with all images anonymised. To comprehensively evaluate the quality of virtually stained images, the pathologists examined each image in four key aspects, including nuclei detail via Haematoxylin stain, cytoplasm detail via Eosin stain, overall



**Fig. 1 | Virtual H&E staining from label-free FLIM images of non-small cell lung cancers.** (a) and (c) are the false-colour FLIM images inputted into the DL model, where (b) and (d) are the corresponding patches within the red rectangle, respectively. Corresponding virtual H&E images are shown (e–h), along with true H&E images (i–l).

**Fig. 2 | Blind evaluation of virtual and true H& E stained lung cancer cores.** **a** Subjective quantification of true and virtual H& E staining of 12 TMA cores by three pathologists. **b** Overall quantification by the inter-observer scores of true and virtual histological images. Each score combines the evaluation of nuclear and cytoplasmic detail, overall staining quality, and diagnostic confidence based on staining quality, quantified using a range of 1 (unacceptable) to 4 (perfect).



staining quality, and quality for diagnostic decision-making. For each dimension, pathologists assigned scores based on predetermined criteria, ranging from Unacceptable (1) to Acceptable (2), Excellent (3), and Perfect (4). This meticulous evaluation process offers valuable insights into the efficacy and dependability of virtual H&E staining compared to conventional staining methodologies.

Figure 2 depicts the mean scores provided by each pathologist on four image quality metrics. In general, the average scores shown in Fig. 2b demonstrate the outstanding quality of virtual histological images on all metrics, where only marginal differences were found between real and virtual images. Although each pathologist rated the images differently (Fig. 2a), their outcomes are consistent on the differences between true and virtual images. For example, pathologist 1 ranked both stains as above excellent, while pathologist 3 considered they are between acceptable (2) and excellent (3). Supplementary Table 1 presents the detailed ranking on each lung tissue sample. An interesting result is on sample 6 (lung sample 2 in Fig. 1), where all pathologists voted the virtual image over the true one on all four metrics. In summary, the subjective evaluation demonstrated a high degree of agreement between the techniques, where the pathologists were able to examine histological features on virtual staining and reach the same level of diagnostic decision-making.

#### Vitual H&E staining on various tissue samples

In addition to lung tissue samples, we expanded our experiments to include primary colorectal and endometrial cancers, and the results are presented in Figs. 3, 4 for colorectal and endometrial adenocarcinomas, respectively. Similar to the results shown in the section “Virtual H&E staining from FLIM”, the virtual images (Fig. 3b, e for colorectal cancer and Fig. 4b, e for endometrial cancer) exhibit significant consistency with the real ones (Fig. 3c, f for colorectal cancer and Fig. 4c, f for endometrial cancer), where most cellular details are accurately reconstructed, which are suitable for both visual and quantitative analysis. A notable phenomenon is observed in the FLIM image of endometrial cancer, where the fluorescence lifetime appears dramatically shorter compared to other samples, resulting in a visually dimmer appearance. However, it's important to note that this shorter lifetime does not affect the virtual result obtained through our method.

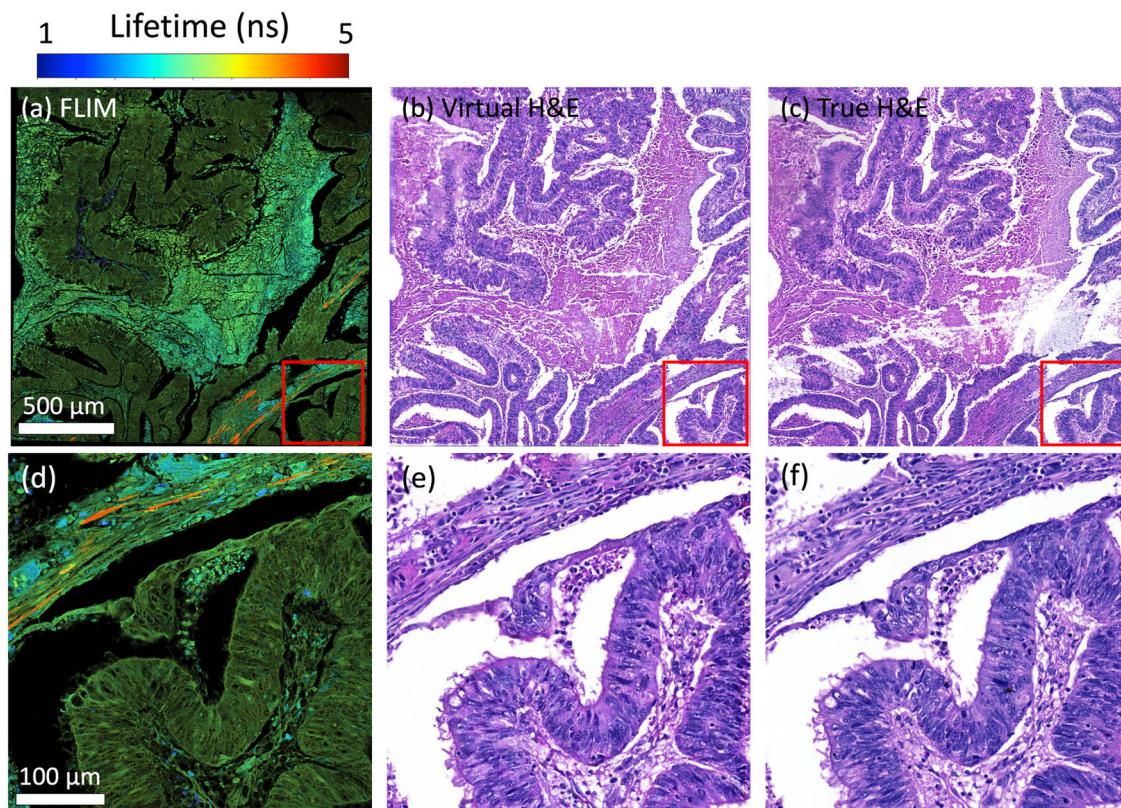
We also expanded our experiments to include Formalin-Fixed Paraffin-Embedded (FFPE) lung biopsies. Figure 5 summarise the results of an FFPE lung biopsy. The virtually stained image (Fig. 5d) of the lung biopsy exhibits consistency with the real one (Fig. 5f), revealing cellular details within the area. Despite the performance of our method on the biopsy is not as remarkable as demonstrated on TMA slides in the section “Virtual H&E staining from FLIM”, the virtual image does not compromise clinical decision-making.

#### Lifetime signatures of various cell types in lung tissue

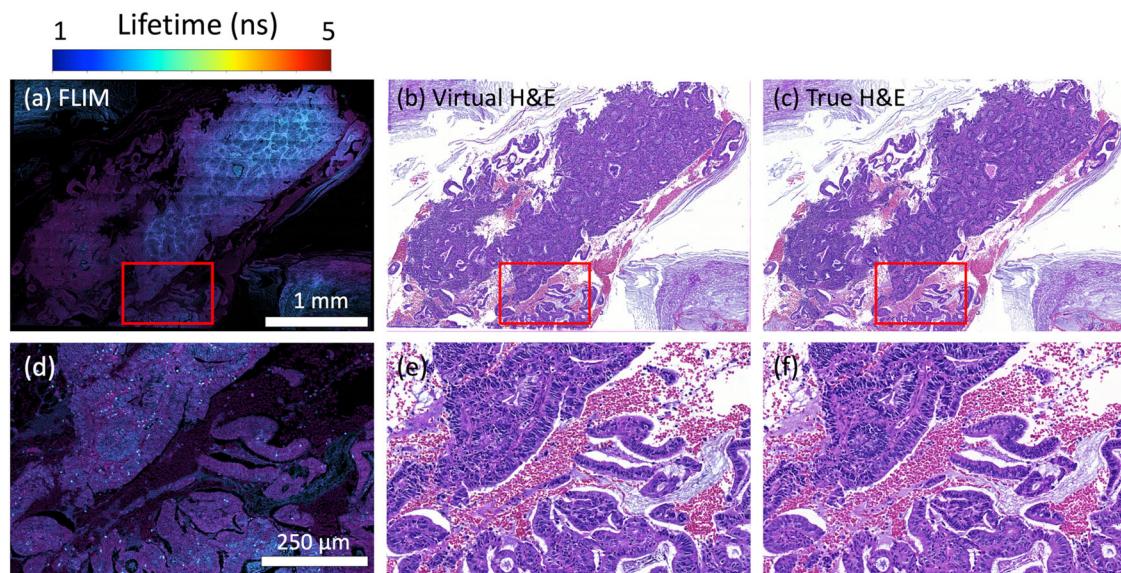
By leveraging virtual H&E staining, we can directly map cellular components annotated by pathologists to FLIM images, enabling the identification of lifetime signatures of diverse cell types, without the need for conventional staining and co-registration processing procedures. Figure 6 depicts an example of label-free autofluorescence lifetime signatures of 7 different cell types, where the lifetime is determined by the peak values in the histogram of pixel lifetime values within the marked cells.

Figure 6a is a FLIM tile patched from Fig. 1c, Fig. 6c is the corresponding virtual H&E image, and Fig. 6b is the true H&E image as the reference. An experienced pathologist annotated 7 different cells commonly found in the TME on virtual histology image and confirmed the true one, including tumour, fibroblasts, lymphocytes, plasma cells, neutrophils, macrophages, and RBC. Afterwards, lifetime histograms are derived for all the cells annotated, and averaged lifetime were extracted. Figure 6d clearly demonstrates the lifetime differences of the cells. Interestingly, the averaged lifetimes for various cell types follow a progressive increase: RBCs have an average lifetime of 0.251 ns, fibroblasts 0.376 ns, tumours 0.424 ns, lymphocytes 0.471 ns, neutrophils 0.659 ns, plasma cells 0.69 ns, and macrophages have the longest average lifetime at 2.42 ns. RBC show the shortest lifetime and appear too dark in the FLIM image (depicted as dark blue dots in Fig. 6a) to be discerned. On the other hand, macrophages have the longest lifetime and are easily identifiable in the FLIM image (represented as large blue dots in Fig. 6a). Plasma cells and neutrophils possess similar lifetime values, resulting in their visual similarity (highlighted by green dots indicated by yellow and blue arrows in Fig. 6a). In comparison, lymphocytes are less visually distinguishable than these two cell types. Tumour cells and fibroblasts exhibit shorter lifetime values, making them appear dimmer in FLIM images compared to other cells. While most tumour cells can be identified in FLIM with the aid of the H&E image as a reference, pinpointing fibroblasts remains a challenging task on the FLIM image.

Figure 6d illustrates the distribution of lifetime of the cells. This variability across different cell types or even within the same cell type is anticipated, which arises from the diverse composition of molecules within individual cells, each with its own characteristic lifetime. In the context of Fig. 6, a specific lifetime suggests the presence of a specific molecule shared among all 7 cell types, with the counts on the y-axis reflecting the relative abundance of this molecule within the cells. In practice, distinguishing between different cell types often involves leveraging both intensity and lifetime data. While intensity provides valuable morphological information, lifetime measurements offer insights into molecular composition. By analysing cellular-level histograms of lifetime, users can gain a deeper understanding of the functional characteristics of various cell types, complementing the morphological details provided by intensity data.



**Fig. 3 | Virtual H&E staining from FLIM on colorectal adenocarcinoma.** (a) and (d) are intensity-weighted FLIM images. (b) and (e) are the virtually stained images. (c) and (f) are the true H&E stain as the reference.

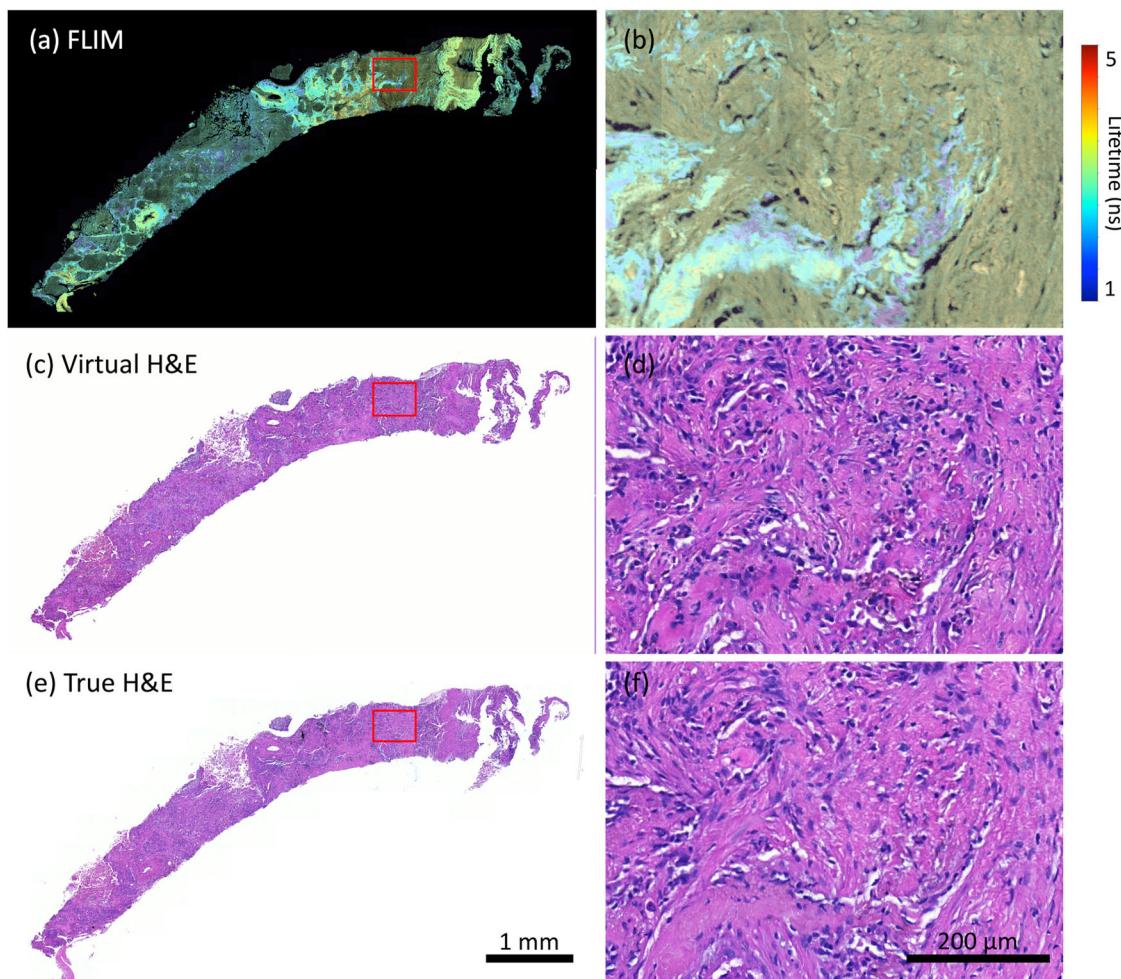


**Fig. 4 | Virtual H&E staining from FLIM on endometrial adenocarcinoma.** (a) and (d) are intensity-weighted FLIM images. (b) and (e) are the virtually stained images. (c) and (f) are the true H&E stain as the reference.

### Comparison of different image formats

Various DL-based approaches have been proposed to leverage label-free autofluorescence images for the generation of virtual histological images, such as<sup>17</sup>. Since FLIM images also contain autofluorescence intensity, we will compare synthetic histology images generated from three different formats. These formats include autofluorescence intensity images in greyscale, false-colour lifetime images with normalised intensity as the alpha channel ( $\alpha$ -FLIM)<sup>27,28</sup>, and intensity-weighted false-colour lifetime images (IW-FLIM).

For intensity images, we followed the post-processing described in<sup>17</sup>. In the case of false-colour lifetime images, the process involves initially converting greyscale lifetime images into 3-channel RGB images through colour mapping, with a fixed range of [1.0 ns, 5 ns]. Subsequently, the normalised intensity is appended as the alpha channel to create  $\alpha$ -FLIM, resulting in four-channel images. For the generation of IW-FLIM images, the normalised intensity is employed as a soft weight, undergoing pixel-wise multiplication with each channel in the false-colour lifetime images. Note that



**Fig. 5 | Virtual H&E staining on an Formalin-Fixed Paraffin-Embedded lung biopsy sample.** (a) and (b) are FLIM images. (c) and (d) are the virtually stained images. (e) and (f) are the true H&E stain as the reference.

despite different processing methods,  $\alpha$ -FLIM and IW-FLIM images are visually identical. These three formats were then inputted into the DL model with the same hyper-parameters for training on the same training dataset.

Figure 7 depicts the synthesis of virtual H&E staining from the aforementioned three image formats. Visual inspection by three pathologists is carried out initially to assess the overall quality of the synthesis. This evaluation reveals that all three formats can generate satisfactory virtual H&E-stained images, without compromising the detection and diagnosis of lung cancer. However, some mis-reconstructed details are observed in intensity-based synthesis, which can impact the biological analysis of the tissue.

In our study, a noticeable limitation of intensity-based reconstruction is its inability to locate some macrophages. In Fig. 7, the red arrows point to clusters of macrophages that are misinterpreted as tumour cells in virtual H&E staining from intensity (Fig. 7i), whereas they are clearly visible in the virtual images generated from  $\alpha$ -FLIM (Fig. 7j) and IW-FLIM (Fig. 7k). However, macrophages can be accurately identified in intensity images when they exhibit strong signals, as indicated by the bright dots highlighted by the blue and cyan arrows in the intensity images (Fig. 7s). Additionally, a few immune cells are misclassified as other cell types in the intensity-based virtual image, as shown by the pink arrow in Fig. 7s. In contrast, both  $\alpha$ -FLIM (Fig. 7t) and IW-FLIM (Fig. 7u) have excellent capabilities for virtual H&E staining.

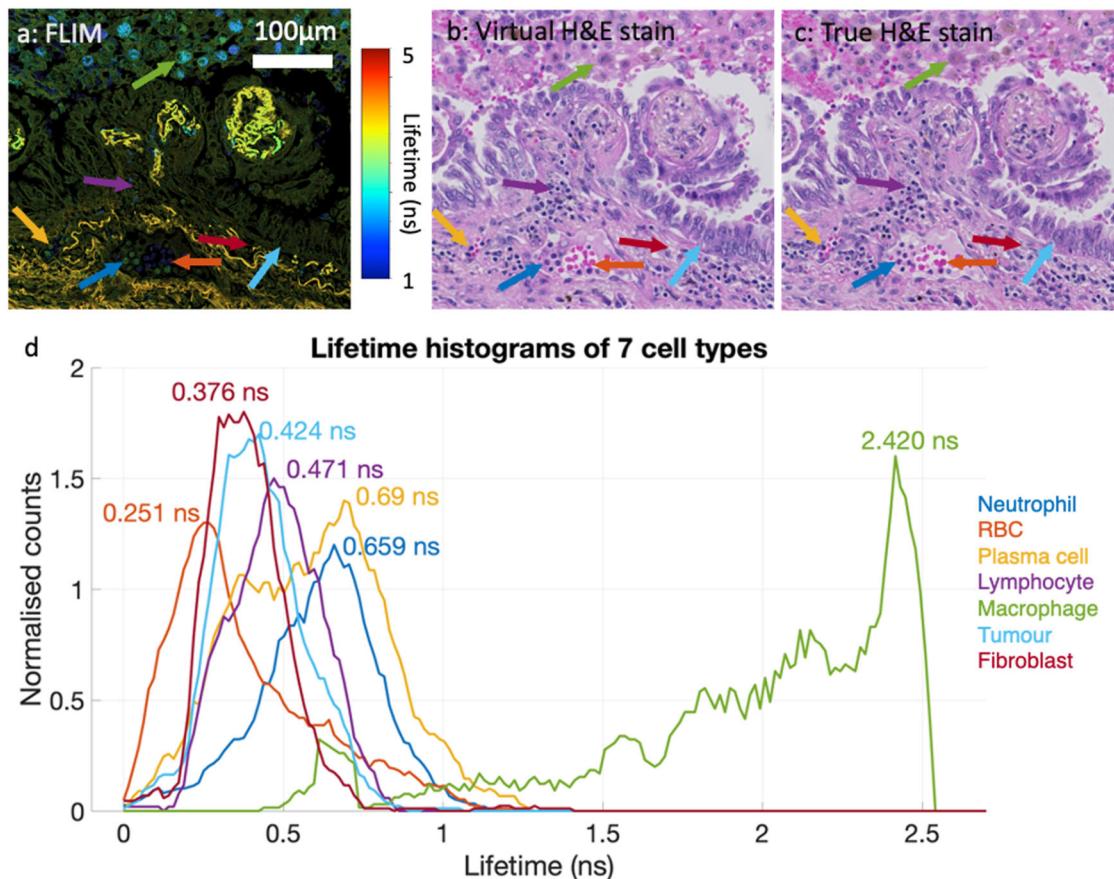
To quantify the comparison, we applied four predominant metrics<sup>29</sup> for measuring the similarity between true and virtual H&E-stained images, including the normalised root mean-squared error (NRMSE), the

normalised mutual information (NMI), the peak signal-to-noise ratio (PSNR), and the mean structural similarity index metric (NSSIM). The results are listed in Table 1 and the best one is highlighted in bold for each metric. Overall, the quantitative results are consistent with the visual results shown in Fig. 7, where all values are comparable. Intensity images are less effective than false-colour FLIM images for synthesising virtual images, with IW-FLIM being the most suitable format for achieving optimal virtual H&E staining.  $\alpha$ -FLIM also produces competitive results similar to IW-FLIM, with marginal differences compared to IW-FLIM.

## Discussion

This study described a method to generate clinical-grade virtual H&E staining from label-free FLIM images acquired on unstained tissue samples, with the assistance of the pix2pix network<sup>19</sup> and the DISTS loss<sup>25</sup>. Despite a variety of virtual histological staining methodologies having been proposed, our study differs from existing methods in several aspects. Our FLIM images were obtained using a confocal microscope at the optimal excitation and emission wavelengths determined by a  $\lambda$ -to- $\lambda$ /spectral scan on the tissue provided by the Leica FLIM system, containing the intensity and the corresponding lifetime images. Although Borhani et al. utilised lifetime images in their study, their input images were collected using a multi-modality microscope, resulting in a combination of 16-channel spectral intensity and single-channel lifetime<sup>20</sup>.

Using FLIM for virtual H&E staining yields superior results, compared to existing methods, such as autofluorescence intensity images, which is primarily attributed to the additional information introduced by



**Fig. 6 | Lifetime signatures of 7 cellular components in lung cancerous tissue, including tumours, lymphocytes, plasma cells, macrophages, neutrophils, fibroblasts, and red blood cells (RBC). DL-generated virtual H&E (b) has the capability to reconstruct those cell types from the FLIM image (a), enabling precise**

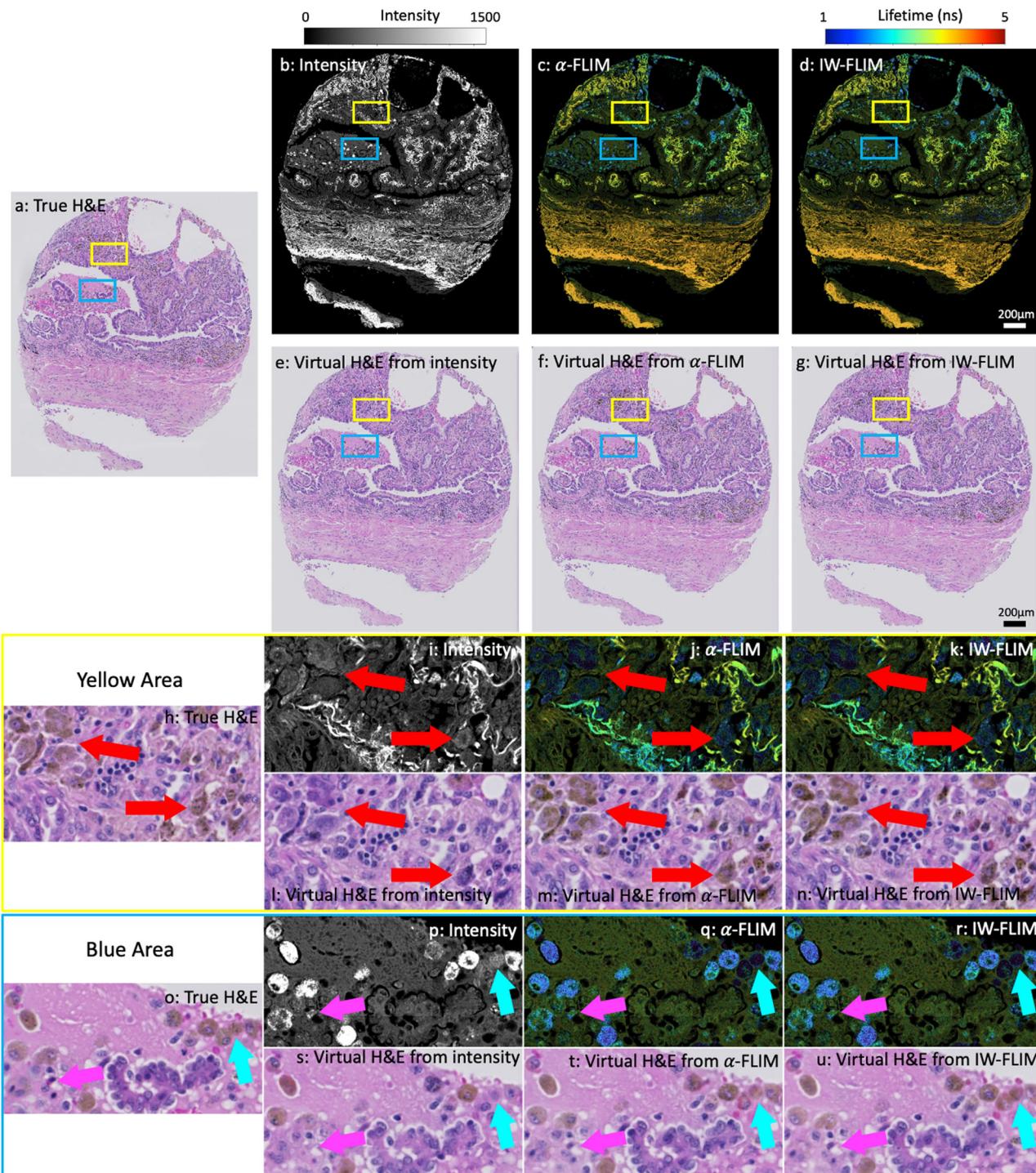
and timely identification of lifetime signatures of cellular components in the tissue. True H&E image is also presented for comparison. (d) demonstrates that 7 cellular components annotated have distinct lifetime values.

fluorescence lifetime. An initial scan of the unstained tissue using a commercial FLIM imaging system can identify the optimal excitation/emission wavelengths best suited for the samples, ensuring the acquisition of high-quality FLIM images. Since FLIM contains both intensity and lifetime information, our previous work and this study demonstrate that combining both information can achieve optimal outcomes<sup>30</sup>. In addition, the section “Comparison of different image formats” indicated that intensity-weighted false-colour lifetime images are most suited for the virtual staining. Consequently, our method applies intensity-weighted false-colour images as the input into a supervised GAN model for paired image-to-image translation.

Despite the prevalence of FLIM in biomedical and clinical applications, most FLIM image analysis methods primarily rely on statistical techniques, such as histogramming, to summarise the data. These approaches are effective at identifying dominant components within FLIM images but may fall short when it comes to providing cellular-level information. This limitation becomes particularly apparent when attempting to discern the intricate details of cellular components within complex environments like the tumour microenvironment (TME). This may be overcome by co-registering FLIM with histology images<sup>10</sup>, which is still challenging, and sometimes may be infeasible. Virtual H&E staining can address the limitations by offering a rapid and precise reference for FLIM images at the cellular level. This reference enables the identification of distinct lifetime signatures associated with various cell types within tissue samples, as illustrated in Fig. 6. By providing this valuable reference instantly, virtual H&E staining greatly expands the potential applications of FLIM beyond its current scope. This advancement has great potential for label-free characterisation of both morphological and phenotypical features of cells, opening up new possibilities for research and clinical purposes.

Compared to other imaging modalities used in the literature for label-free virtual H&E staining, our FLIM microscope is more complicated and usually requires longer time to image tissue samples. For example, the images presented in Fig. 1, 3, 4, and Supplementary Fig. 1a-d were acquired within 40 min per TMA core with a pixel size of 455 nm. However, following the optimisation of scanning parameters, the scanning time reduced to less than 15 min while maintaining the same quality. This allows us to scan a TMA core within 30 min at a much higher resolution (a pixel size of 200 nm), as shown in Supplementary Fig. 1e-j. On the other hand, H&E stain is a routine test in biology and pathology laboratories, which takes about 30 min for the process. It seems counterproductive to use FLIM for virtual H&E staining, given the time on FLIM imaging. Considering the quality of FLIM-based virtual H&E staining, we believe it is worthwhile. Additionally, FLIM has demonstrated the capability for providing metabolic information, without requiring exogenous biomarkers<sup>14–7,31</sup>. All these together have great potential to address challenges in managing small biopsies and limited tissue samples in advanced-stage lung cancer patients undergoing only small biopsies rather than resection procedures. By leveraging FLIM, we can optimise tissue utilisation by reducing the necessity for additional ancillary tests while achieving the same level of diagnostic and phenotypic characterisation.

In the process of sample preparation and data acquisition, a common practice for virtual histological staining from autofluorescence images involves first scanning unstained samples using autofluorescence imaging modalities and subsequently staining the same samples<sup>17,18,32</sup>. This approach offers a significant advantage in achieving optimal co-registration at the pixel level with minimal distortion on both images, which is crucial for the effective utilisation of supervised DL techniques.



**Fig. 7 | Comparison of virtual H & E staining using intensity, false-colour FLIM with the intensity image as the alpha channel ( $\alpha$ -FLIM), and intensity-weighted false-colour FLIM (IW-FLIM). (a), (h), and (o) are the true H & E staining. (e), (i), and (s) are the virtual H & E staining from intensity (b), (i), and (p), respectively. (f), (m), and (t) are the virtual H & E staining from  $\alpha$ -FLIM (c), (j), and (q), respectively. (g), (n), and (u) are the virtual H & E staining from IW-FLIM (d), (k), and (r), respectively. The yellow and blue rectangles highlight variations in the synthetic**

results. The red arrows point out macrophages that are entirely absent in the intensity-based synthesis (i) but appear in the  $\alpha$ -FLIM (j) and IW-FLIM-based (k) outcomes. Synthesis based on intensity can reconstruct macrophages when the signal is strong (bright dots pointed out by the cyan arrows in (p)), but it struggles to identify those with dimmer signals (the grey dot surrounded by the bright dots, as indicated by the cyan arrow in p). The pink arrows also indicate a few missing immune cells in intensity-based synthesis.

To achieve this, diverse techniques have been applied, such as multi-stage elastic-based registration<sup>17</sup> or RANSAC-based warping<sup>20</sup>. Nonetheless, our study suggests that with the same sample preparation and data collection procedure, employing an affine transformation for the co-registration is adequate to achieve satisfactory virtual staining, thus,

substantially reduce technical uncertainties for enhanced reproducibility. Supplementary Fig. 2 illustrates that the synthetic results with additional elastic registration do not contribute to better virtual staining, which may attribute to the sample preparation described in the section “Sample preparation and data collection”.

**Table 1 | Quantitative comparison of virtual H&E staining on two TMA sections of lung cancerous tissues from intensity,  $\alpha$ -FLIM, and IW-FLIM using the normalised root-mean-squared error (NRMSE), the normalised mutual information (NMI), the peak signal-to-noise ratio (PSNR), and the mean structural similarity index metric (MSSIM).**

		NRMSE	NMI	PSNR	MSSIM
Lung sample 1 (Fig. 1a)	Intensity	0.127	1.093	20.085	0.529
	a-FLIM	0.127	1.091	20.121	0.53
	IW-FLIM	<b>0.126</b>	<b>1.096</b>	<b>20.202</b>	<b>0.533</b>
Lung sample 2 (Fig. 1c)	Intensity	0.095	1.154	22.767	0.689
	a-FLIM	0.093	1.161	22.962	0.695
	IW-FLIM	<b>0.092</b>	<b>1.164</b>	<b>23.026</b>	<b>0.697</b>

Note that smaller NRMSE and NMI indicate more similar images, whereas larger PSNR and MSSIM represent better similarity between images.

It is important to note that although we leveraged pix2pix and DISTS as our DL method, we believe that other advanced models could also be exploited for the purpose. Indeed, many models have proved to be effective for virtual H&E staining, such as UNet-based GAN<sup>17,32,33</sup>, FCNN-p2p and VGG-a2p<sup>20</sup>, or conditional GAN<sup>18</sup>. Generally, all DL models described in<sup>34</sup> for image-to-image translation may be applicable to DL-based virtual histological staining. However, direct applications may not be feasible for this purpose. We experimented this with three advanced DL models without any optimisation, including the original pix2pix GAN, ResVit<sup>35</sup>, and Denoising diffusion GAN (DDGAN, <sup>36</sup>), and the results are depicted in Supplementary Fig. 3. Despite the success of these models, they are unable to generate acceptable virtual H&E staining. Therefore, incorporating advanced loss functions to further refine the learning process would also be beneficial for achieving optimal results. For example, the perceptual loss<sup>37</sup> can be employed for high-level feature reconstruction while preserving realistic textural information, and the texture loss<sup>38</sup> can help retain fine textural details from the original images. The DISTS loss<sup>25</sup> harnessed in this study combines spatial information with textural features, making it a more favourable choice for enhancing the reconstruction.

Various virtual histological staining techniques have been proposed over the past few years<sup>14</sup>, and our approach perfectly complements these methods in a novel aspect. Existing methods for virtual H&E staining usually facilitate autofluorescence images acquired at specific excitation/emission wavelengths using specific filters, such as DAPI and Cy5 channels<sup>17</sup>. In contrast, our FLIM image acquisition utilised excitation and emission wavelengths determined through a  $\lambda$ -to- $\lambda$  scan of the lung tissue samples, resulting in optimal excitation (485 nm) and emission ([500 nm, 720 nm]) wavelengths for our study. This represents a fundamental difference in our approach compared to other proposals, which is independent of specific pre-set filters, allowing us to collect high-quality images at optimal excitation/emission wavelengths. This enhancement was confirmed through blinded evaluations by experienced pathologists and quantitative assessments using commonly used similarity metrics (see the section “Comparison of different image formats”).

## Methods

### Sample preparation and data collection

Surgically resected, early stage non-small cell lung carcinomas were identified from the clinical archival and a TMA was constructed as previously described<sup>39</sup>. FLIM images were collected using a confocal FLIM system (Leica STELLARIS 8 FALCON FLIM Microscope) with an 20x/0.75NA objective. The excitation and emission wavelengths were set at 485 nm and [500 nm, 720 nm], which was determined by a  $\lambda$ -to- $\lambda$  scan of the tissue. The image size was fixed at 512 × 512 pixels with various pixel sizes. Following FLIM acquisition the same sample underwent rehydration in a series of

ethanol dilutions, staining with Haematoxylin and Eosin and then dehydration and subsequent imaging on a bright-field slide scanner (ZEISS Axio Scan.Z1), with an 20x/0.75NA objective. The digitalised H&E-stained images have a pixel size of 0.22  $\mu\text{m}$ . In total, 84 lung cancer tissue samples (including 69 TMA cores scanned with the pixel size 0.455  $\mu\text{m}$ , 5 biopsies scanned with the pixel size at 0.4  $\mu\text{m}$ , and another 10 TMA cores from a separate cohort scanned with the pixel size at 0.25  $\mu\text{m}$ ), 4 slides of colorectal cancer scanned with the pixel size 0.4  $\mu\text{m}$ , and four sections of endometrial cancer scanned with the pixel size 0.4  $\mu\text{m}$  were facilitated for this study.

### Data post-processing

FLIM images were reconstructed by an exponential fitting on photon histogram counts to estimate lifetimes using the Leica LAX-X software and exported as intensity and lifetime images separately. Note that to ensure the consistency when visualising FLIM images, raw intensity data is limited to [0, 2000] and raw lifetime is to [0.0 ns, 5.0 ns]. Afterwards, all tiles per core within the tumour microarray (TMA) were stitched using MIST<sup>40</sup>. To achieve optimal registration, only intensity images were exploited. Similar to refs. 17, intensity images were colour inverted for the white background. H&E images were converted to greyscale and contrast-enhanced by saturating 1% of the values at both low and high bounds of the intensity range. The processed intensity and H&E images were used as the input for the co-registration. Since the tissue processing did not change the underlying structural and cellular information, Affine transformation in MATLAB was employed for the co-registration of FLIM and H&E images (Fig. 8b).

The co-registered images are then resampled to the patches of 256 × 256 pixels, where those patches with over 75% background were disregarded. For the input into the pix2pix network, three different formats were tested. The first is the intensity images, the second is the false-colour lifetime images with the corresponding normalised intensity images as the alpha channel, and the third is the false-colour lifetime images pixel-multiplied with the corresponding intensity images. Note that intensity images are globally normalised per TMA core before applying to the lifetime images.

### GAN Architecture

In this study, the pix2pix network<sup>19</sup> was utilised and a conceptual diagram is depicted in Fig. 8c. Essentially, the pix2pix GAN is a conditional model, where the generator is U-Net-like architecture and the discriminator is a multi-layer classifier using the input FLIM images as the condition. The objective of the discriminator is given by:

$$\mathcal{L}(G, D) = \mathbb{E}_{f,s}[\log D(s|f)] + \mathbb{E}_{f,h}[\log(1 - D(h|f))] \quad (1)$$

where  $f$ ,  $s$ , and  $h$  are FLIM, synthetic H&E, and true H&E images, respectively, and  $\mathbb{E}$  is the expectation. The objective of the generator is:

$$\mathcal{L}(G) = \mathbb{E}_{f,s,h}[||h - G(f, s)||_1] \quad (2)$$

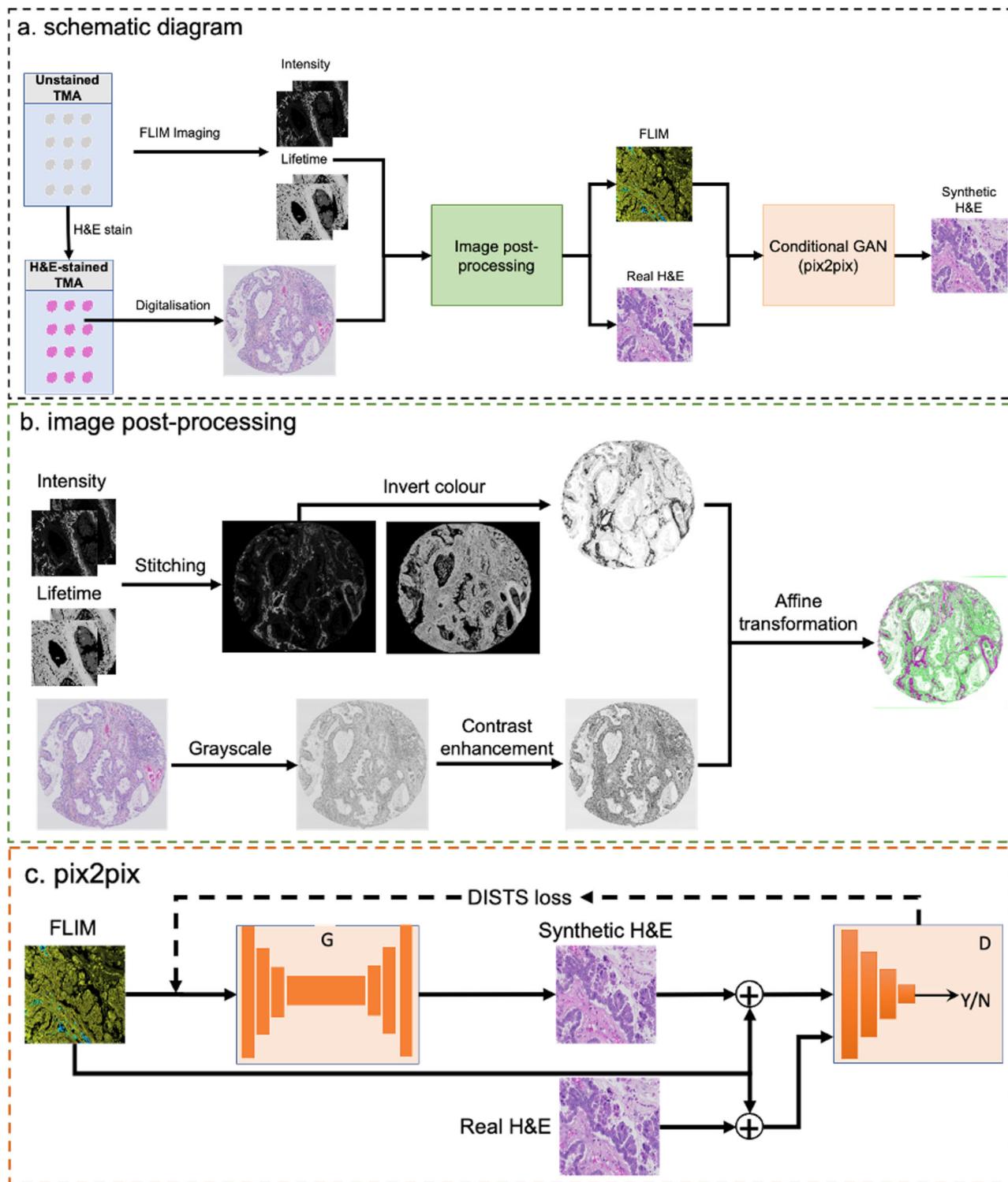
where  $L_1$  distance is applied. Accordingly, the overall objective for the GAN is:

$$\mathcal{L} = \alpha \mathcal{L}(G, D) + \beta \mathcal{L}(G) \quad (3)$$

where  $\alpha$  and  $\beta$  are the regularisation parameter to adjust the weight of the discriminator and generator, respectively.

### DISTS Loss

The regularisation terms in the original GAN networks are usually insufficient to synthesise acceptable virtual histological staining and various extra constraints have been applied to overcome the challenges, such as total variations (TV)<sup>17,32,41</sup> and Structural Similarity Index Measure (SSIM)<sup>20,42</sup>. To achieve the optimal synthesis, the Deep Image Structure and Texture Similarity (DISTS) metric is employed in this study, which yields superiority to existing single metrics by combining structural and texture information



**Fig. 8 | Schematic diagram of the proposed method.** (a) is the overall procedure of the method. (b) is the image post-processing to generate paired FLIM and H&E images. (c) is the conditional GAN (pix2pix) and DISTS loss for the synthesis.

together and has proven to be tolerant to mild local and global structural distortion<sup>25</sup>.

Given  $r$  and  $s$  are true and synthetic images, respectively, and  $\mathcal{F}$  is a function to retrieve feature maps of input images from a trained VGG model, the extracted features  $\tilde{x}_j^{(i)}$  can be defined as  $\mathcal{F}(x) = \{\tilde{x}_j^{(i)}; i \in [0, 5]; j \in [1, n_i]; x \in [r, s]\}$ , where  $\tilde{x}^{(0)} = x$ ,  $i$  is  $i$ th convolution layer, and  $n_i$  is feature maps in the  $i$ th convolution layer. DISTS

is defined as:

$$\mathcal{L}_{DISTS}(r, s, \eta, \theta) = 1 - \sum_{i=0}^m \sum_{j=1}^{n_i} (\eta_{ij} \mathcal{T}(\mathcal{F}(r), \mathcal{F}(s)) + \theta_{ij} \mathcal{S}(\mathcal{F}(r), \mathcal{F}(s))) \quad (4)$$

where  $\eta$  and  $\theta$  are learnable weights, satisfying  $\sum_{i=0}^m \sum_{j=1}^{n_i} (\eta_{ij} + \theta_{ij}) = 0$ ,  $\mathcal{T}$  and  $\mathcal{S}$  are quality measurement for the texture and structure,

respectively, governed by:

$$\mathcal{T}(\mathcal{F}(r), \mathcal{F}(s)) = \frac{2\mu_{\tilde{r}_j}^{(i)}\mu_{\tilde{s}_j}^{(i)} + c_1}{\left(\mu_{\tilde{r}_j}^{(i)}\right)^2 + \left(\mu_{\tilde{s}_j}^{(i)}\right)^2 + c_1} \quad (5)$$

and

$$\mathcal{S}(\mathcal{F}(r), \mathcal{F}(s)) = \frac{2\sigma_{\tilde{r}_j \tilde{s}_j}^{(i)} + c_2}{\left(\sigma_{\tilde{r}_j}^{(i)}\right)^2 + \left(\sigma_{\tilde{s}_j}^{(i)}\right)^2 + c_2} \quad (6)$$

where  $\mu_{\tilde{r}_j}^{(i)}$ ,  $\mu_{\tilde{s}_j}^{(i)}$ ,  $\sigma_{\tilde{r}_j}^{(i)}$ , and  $\sigma_{\tilde{s}_j}^{(i)}$  are the global means and variances of the extracted feature maps for  $r$  and  $s$ , respectively,  $\sigma_{\tilde{r}_j \tilde{s}_j}^{(i)}$  is the global covariance between the feature maps, and  $c_1$  and  $c_2$  are tiny values for numeric stability.

Combining Eq. (3) and Eq. (4), the final objective of the network is:

$$\mathcal{L} = a\mathcal{L}(G, D) + \beta\mathcal{L}(G) + \lambda\mathcal{L}_{DISTS}(r, s, \eta, \theta) \quad (7)$$

where  $\lambda$  is another regularisation parameter for DISTS loss.

### Blinded evaluation of virtual and true H&E staining

A set of 12 lung TMA cores were facilitated for the blinded evaluation, where all images were anonymised to allow three experienced pathologists to perform unbiased inspection. To holistically appraise the quality of the images, pathologists meticulously examined each image's four fundamental aspects: nuclei clarity via Haematoxylin staining, cytoplasm precision via Eosin staining, overall staining integrity, and suitability for diagnostic decision-making. For each parameter, pathologists judiciously assigned scores aligned with established criteria, ranging from Unacceptable (1) to Acceptable (2), Excellent (3), and Exemplary (4).

### Implementation details

The model was implemented using PyTorch and adapted for three input formats, including greyscale intensity images, false-colour lifetime images with normalised intensity as the alpha channel, and false-colour lifetime images pixelwise weighted by normalised intensity. Adaptive Moment Estimation optimiser<sup>43</sup> was deployed with  $\beta_1$  as 0.5 and  $\beta_2$  as 0.999. Learning rate was initialised at 0.0001 and decayed by 10 at very 60 epochs. Total epochs were set to 300.  $\alpha$ ,  $\beta$ , and  $\lambda$  in Eq. (7) were set to 0.1, 1, and 5, respectively.

72 lung cancer samples (including 70 TMA cores and 2 biopsies) were utilised for the training and 15 independent lung cancer sections (including 12 TMA cores and 3 biopsies) were split out as the independent testing set. For colorectal and endometrial cancers, both were trained via transfer learning on 2 sections and evaluated on two separate sections. The patches in the training stage were fixed at  $256 \times 256$ . Simple data augmentations, including horizontal flipping and 15-degree rotation, were applied during the training. Training was performed on NVIDIA V100 GPUs provided by the EPSRC Tier-2 National HPC Services Cirrus hosted by EPCC (<https://www.epcc.ed.ac.uk/hpc-services/cirrus>) at The University of Edinburgh.

### Data availability

All data generated or analysed during this study are included in this published article.

### Code availability

The pix2pix used in this study is publicly available at <https://github.com/mrzhu-cool/pix2pix-pytorch>. The DISTS loss is available at <https://github.com/dingkeyan93/DISTS>. FLIM images were stitched using Fiji MIST stitching plugin (<https://github.com/usnistgov/MIST>). MATLAB was used for affine transformation to co-register FLIM and true histology images.

Received: 10 January 2024; Accepted: 11 June 2024;

Published online: 28 June 2024

### References

1. Marcu, L. Fluorescence lifetime techniques in medical applications. *Ann. Biomed. Eng.* **40**, 304–31 (2012).
2. Datta, R., Heaster, T. M., Sharick, J. T., Gillette, A. A. & Skala, M. C. Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications. *J. Biomed. Opt.* **25**, 1–43 (2020).
3. Wang, Y. L. et al. Rapid, label-free, and highly sensitive detection of cervical cancer with fluorescence lifetime imaging microscopy. *IEEE J. Sel. Top. Quantum Electron.* **22** (2016).
4. Alfonso-Garcia, A. et al. Label-free identification of macrophage phenotype by fluorescence lifetime imaging microscopy. *J. Biomed. Opt.* **21**, 46005 (2016).
5. Walsh, A. J. et al. Classification of t-cell activation via autofluorescence lifetime imaging. *Nat. Biomed. Eng.* **5**, 77–88 (2021).
6. Linders, P. T., Ioannidis, M., Ter Beest, M. & van den Bogaart, G. Fluorescence lifetime imaging of ph along the secretory pathway. *ACS Chem. Biol.* **17**, 240–251 (2022).
7. Yan, Y. et al. Fluorescence intensity and lifetime imaging of lipofuscin-like autofluorescence for label-free predicting clinical drug response in cancer. *Redox Biol.* **59**, 102578 (2023).
8. Rahim, M. K. et al. Phasor analysis of fluorescence lifetime enables quantitative multiplexed molecular imaging of three probes. *Anal. Chem.* **94**, 14185–14194 (2022).
9. Unger, J. et al. Real-time diagnosis and visualization of tumor margins in excised breast specimens using fluorescence lifetime imaging and machine learning. *Biomed. Opt. Express* **11**, 1216–1230 (2020).
10. Wang, Q. et al. Deep learning-assisted co-registration of full-spectral autofluorescence lifetime microscopic images with h&e-stained histology images. *Commun. Biol.* **5**, 1119 (2022).
11. Madabhushi, A. & Lee, G. Image analysis and machine learning in digital pathology: challenges and opportunities. *Med. Image Anal.* **33**, 170–175 (2016).
12. Adhikari, M., Houhou, R., Hniopek, J. & Bocklitz, T. Review of fluorescence lifetime imaging microscopy (flim) data analysis using machine learning. *J. Exp. Theor. Anal.* **1**, 44–63 (2023).
13. Gouzou, D. et al. Applications of machine learning in time-domain fluorescence lifetime imaging: a review. *Methods Appl. Fluorescence* **12**, 022001 (2023).
14. Bai, B. et al. Deep learning-enabled virtual histological staining of biological samples. *Light Sci. Appl.* **12**, 57 (2023).
15. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention*, Pt III, Vol. 9351, 234–241 (Springer International Publishing, 2015).
16. Goodfellow, I. et al. in *Advances in Neural Information Processing Systems* Vol. 27 (eds. Ghahramani, Z. et al.) (Curran Associates, Inc., 2014).
17. Rivenson, Y. et al. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nat. Biomed. Eng.* **3**, 466–477 (2019).
18. Li, D. et al. Deep learning for virtual histological staining of bright-field microscopic images of unlabeled carotid artery tissue. *Mol. Imaging Biol.* **22**, 1301–1309 (2020).
19. Isola, P., Zhu, J. Y., Zhou, T. H. & Efros, A. A. Image-to-image translation with conditional adversarial networks. In: *30th IEEE Conference on Computer Vision and Pattern Recognition (Cvpr 2017)* 5967–5976 (IEEE Computer Society, 2017).
20. Borhani, N., Bower, A. J., Boppart, S. A. & Psaltis, D. Digital staining through the application of deep neural networks to multi-modal multi-photon microscopy. *Biomed. Opt. Express* **10**, 1339–1350 (2019).
21. Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proc. IEEE International Conference On Computer Vision* (eds. Ikeuchi, K., Pelillo, M. & Medioni, G.) 2223–2232 (IEEE, 2017).

22. Cao, R. et al. Label-free intraoperative histology of bone tissue via deep-learning-assisted ultraviolet photoacoustic microscopy. *Nat. Biomed. Eng.* **7**, 124–134 (2023).
23. Hong, Y. et al. Deep learning-based virtual cytokeratin staining of gastric carcinomas to measure tumor–stroma ratio. *Sci. Rep.* **11**, 19255 (2021).
24. Ghahremani, P. et al. Deep learning-inferred multiplex immunofluorescence for immunohistochemical image quantification. *Nat. Mach. Intell.* **4**, 401–412 (2022).
25. Ding, K. Y., Ma, K. D., Wang, S. Q. & Simoncelli, E. P. Image quality assessment: Unifying structure and texture similarity. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 2567–2581 (2022).
26. Zhang, G. et al. Image-to-images translation for multiple virtual histological staining of unlabeled human carotid atherosclerotic tissue. *Mol. Imaging Biol.* **24**, 31–41 (2022).
27. McGinty, J. et al. Wide-field fluorescence lifetime imaging of cancer. *Biomed. Opt. Express* **1**, 627–640 (2010).
28. Williams, G. O. et al. Full spectrum fluorescence lifetime imaging with 0.5 nm spectral and 50 ps temporal resolution. *Nat. Commun.* **12**, 6616 (2021).
29. Zhou, S. K., Rueckert, D. & Fichtinger, G. *Handbook of Medical Image Computing and Computer Assisted Intervention* (Academic Press, 2019).
30. Wang, Q. et al. A layer-level multi-scale architecture for lung cancer classification with fluorescence lifetime imaging endomicroscopy. *Neural Comput. Appl.* **34**, 18881–18894 (2022).
31. Ozcan, A. et al. Method and system for digital staining of microscopy images using deep learning (2023). US Patent App. 17/783,260.
32. Zhang, Y. J. et al. Digital synthesis of histological stains using microstructured and multiplexed virtual staining of label-free tissue. *Light Sci. Appl.* **9** (2020).
33. Li, J. X. et al. Biopsy-free in vivo virtual histology of skin using deep learning. *Light Sci. Appl.* **10** (2021).
34. Pang, Y., Lin, J., Qin, T. & Chen, Z. Image-to-image translation: methods and applications. *IEEE Trans. Multimedia* **24**, 3859–3881 (2022).
35. Dalmaz, O., Yurt, M. & Çukur, T. Resvit: residual vision transformers for multimodal medical image synthesis. *IEEE Trans. Med. Imaging* **41**, 2598–2614 (2022).
36. Xiao, Z., Kreis, K. & Vahdat, A. Tackling the generative learning trilemma with denoising diffusion gans. In: *International Conference on Learning Representations* (eds. Katja Hofmann, A. S. R.) (ICLR, 2021).
37. Johnson, J., Alahi, A. & Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In: *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings Part II* 14 694–711 (Springer International Publishing, 2016).
38. Gatys, L. A., Ecker, A. S. & Bethge, M. Image style transfer using convolutional neural networks. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 2414–2423 (IEEE Computer Society, 2016).
39. Koppensteiner, L. et al. Location of cd39+ t cell subpopulations within tumors predict differential outcomes in non-small cell lung cancer. *J. ImmunoTherapy Cancer* **11** (2023).
40. Chalfoun, J. et al. Mist: accurate and scalable microscopy image stitching tool with stage modeling and error minimization. *Sci. Rep.* **7** (2017).
41. Rivenson, Y. et al. Phastain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light Sci. Appl.* **8** (2019).
42. Liu, S. T. et al. Unpaired stain transfer using pathology-consistent constrained generative adversarial networks. *IEEE Trans. Med. Imaging* **40**, 1977–1989 (2021).
43. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. In: *International Conference on Learning Representations* (2015).

## Acknowledgements

This study was funded by Cancer Research UK's Early Detection and Diagnosis Primer Award (EDDPMA-May22/100054), NVIDIA Academic Hardware Grant Program, The University of Edinburgh's Wellcome Institutional Translational Partnership Accelerator Fund and Cancer Research Horizons Seed Fund (PIII140), Cancer Research UK Clinician Scientist Fellowship (A24867), the UK Engineering and Physical Sciences Research Council (EP/R005257/1 and EP/S025987/). The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

## Author contributions

Q.W. proposed the ideas, conducted the experiments, and prepared the manuscript. A.A., D.D. and S.T. prepared and processed the tissue, and provided valuable comments from the clinical perspective. D.D., B.M. and C.T. blindly evaluated the results. Q.W. and M.J. collected and post-processed FLIM images. M.V. and J.H. offered valuable comments on lifetime reconstruction and deep learning to consolidate the ideas. Q.W., A.A., M.V. and J.H. supervised the project. All authors contributed to the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44303-024-00021-7>.

**Correspondence** and requests for materials should be addressed to Qiang Wang.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024