1   **Title:** Quantification of vascular networks in photoacoustic mesoscopy

2

3   **Authors:**

4   Emma L. Brown[1,2,#], Thierry L. Lefebvre[1,2,#], Paul W. Sweeney[1,2,#], Bernadette J. Stolz[3], Janek

5   Gröhl[1,2], Lina Hacker[1,2], Ziqiang Huang[2] [+], Dominique-Laurent Couturier[2], Heather A.

6   Harrington[3], Helen M. Byrne[3] and Sarah E. Bohndiek[1,2] *.

7

8   [1] Department of Physics, University of Cambridge, JJ Thomson Avenue, Cambridge, CB3

9   0HE, U.K.

10

11   [2] Cancer Research UK Cambridge Institute, University of Cambridge, Robinson Way,

12   Cambridge, CB2 0RE, U.K.

13

14   [3] Mathematical Institute, University of Oxford, Woodstock Road, Oxford, OX2 6GG, U.K.

15

16   + Now at: European Molecular Biology Laboratory Imaging Centre, Heidelberg, Germany.

17

18   # These authors contributed equally

19   * Corresponding author: email seb53@cam.ac.uk, telephone +44 1223 337267.

20

21   **Keywords:** photoacoustic imaging, vasculature, segmentation, topology

22

23 **ABSTRACT**

24 Mesoscopic photoacoustic imaging (PAI) enables non-invasive visualisation of tumour

25 vasculature and has the potential to assess prognosis and therapeutic response. Currently,

26 evaluating vasculature using mesoscopic PAI involves visual or semi-quantitative 2D

27 measurements, which fail to capture 3D vessel network complexity, and lack robust ground

28 truths for assessment of segmentation accuracy. Here, we developed an *in silico*, phantom,

29 *in vivo*, and *ex vivo*-validated end-to-end framework to quantify 3D vascular networks captured

30 using mesoscopic PAI. We applied our framework to evaluate the capacity of rule-based and

31 machine learning-based segmentation methods, with or without vesselness image filtering, to

32 preserve blood volume and network structure by employing topological data analysis. We first

33 assessed segmentation performance against ground truth data of *in silico* synthetic

34 vasculatures and a photoacoustic string phantom. Our results indicate that learning-based

35 segmentation best preserves vessel diameter and blood volume at depth, while rule-based

36 segmentation with vesselness image filtering accurately preserved network structure in

37 superficial vessels. Next, we applied our framework to breast cancer patient-derived

38 xenografts (PDXs), with corresponding *ex vivo* immunohistochemistry. We demonstrated that

39 the above segmentation methods can reliably delineate the vasculature of 2 breast PDX

40 models from mesoscopic PA images. Our results underscore the importance of evaluating the

41 choice of segmentation method when applying mesoscopic PAI as a tool to evaluate vascular

42 networks *in vivo*.

43

INTRODUCTION

44

45   Tumour blood vessel networks are often chaotic and immature (Brown et al., 2019;

46 Corliss et al., 2019; Hanahan & Weinberg, 2011; Krishna Priya et al., 2016; Nagy & Dvorak,

47 2012), with inadequate oxygen perfusion and therapeutic delivery (Michiels et al., 2016;

48 Trédan et al., 2007). The association of tumour vascular phenotypes with poor prognosis

49 across many solid cancers (Brown et al., 2019) has generated substantial interest in non-

50 invasive imaging of the structure and function of tumour vasculature, particularly longitudinally

51 during tumour development. Imaging methods that have been tested to visualise the

52 vasculature include whole-body macroscopic methods, such as computed tomography and

53 magnetic resonance imaging, as well as localised methods, such as ultrasound and

54 photoacoustic imaging (PAI) (Brown et al., 2019). Microscopy methods can achieve much

55 higher spatial resolution but are typically depth limited, at up to ~1mm depth, and frequently

56 applied *ex vivo* (Brown et al., 2019; Jährling et al., 2009; Kelch et al., 2015; Keller & Dodt,

57 2012; Ntziachristos, 2010).

58   Of the available tumour vascular imaging methods, PAI is highly scalable and, as such,

59 applicable for studies from microscopic to macroscopic regimes. By measuring ultrasound

60 waves emitted from endogenous molecules, including haemoglobin, following the absorption

61 of light, PAI can reconstruct images of vasculature at depths beyond the optical diffraction limit

62 of ~1 mm (Beard, 2011; Ntziachristos, 2010; Ntziachristos et al., 2005; Wang & Yao, 2016).

63 State-of-the-art mesoscopic systems now bridge the gap between macroscopy and

64 microscopy, achieving ~20 µm resolution at up to 3 mm in depth (Omar et al., 2014, 2019).

65 Preclinically, mesoscopic PAI has been used to monitor the development of vasculature in

66 several tumour xenograft models (Haedicke et al., 2020; Omar et al., 2015; Orlova et al., 2019)

67 and can differentiate aggressive from slow-growing vascular phenotypes (Orlova et al., 2019).

68 Studies to-date, however, have been largely restricted to qualitative analyses due to the

69 challenges of accurate 3D vessel segmentation, quantification and robust statistical analyses

70 (Haedicke et al., 2020; Imai et al., 2017; Omar et al., 2015, 2019; Orlova et al., 2019; Rebling

71   et al., 2021). Instead, PAI quantification is typically manual and ad-hoc, with 2D measurements

72   often extracted from 3D PAI data (Haedicke et al., 2020; Imai et al., 2017; Lao et al., 2008;

73   Orlova et al., 2019; Soetikno et al., 2012), reducing repeatability and comparability across

74   datasets.

75   To assess the performance and accuracy of such vessel analyses, ground truth

76   datasets are needed with *a priori* known features (Krig & Krig, 2014). Creating full-network

77   ground truth reference annotations could be achieved through comprehensive manual

78   labelling of PAI data, but this is difficult due to: the lack of available experts to perform

79   annotation with a new imaging modality; the time taken to label images; and the inherent noise

80   and artefacts present in PAI data. Despite the numerous software packages available to

81   analyse vascular networks (Corliss et al., 2019), their performance in mesoscopic PAI has yet

82   to be evaluated, hence there is an unmet need to improve the quantification of vessel networks

83   in PAI, particularly given the increasing application of PAI in the study of tumour biology

84   (Haedicke et al., 2020; Omar et al., 2019; Orlova et al., 2019).

85   To quantify PAI vascular images and generate further insights into the role of vessel

86   networks in tumour development and therapy response, accurate segmentation of the vessels

87   must be performed (Corliss et al., 2019) (see step 1 in **Figure 1**). A plethora of segmentation

88   methods exist and can be broadly split into two categories: rule-based and machine learning-

89   based methods. Rule-based segmentation methods encompass techniques that automatically

90   delineate the vessels from the background based on a custom set of rules (F. Zhao et al.,

91   2019). These methods provide less flexibility and tend to consider only a few features of the

92   image, such as voxel intensity (Haedicke et al., 2020; Orlova et al., 2019; Raumonen &

93   Tarvainen, 2018; Soetikno et al., 2012) but they are easy-to-use, with no training dataset

94   requirements. On the other hand, machine learning-based methods, such as random forest

95   classifiers, delineate vessels based on self-learned features (Moccia et al., 2018; F. Zhao et

96   al., 2019). Nonetheless, learning-based methods are data-driven, requiring large and high-

97   quality annotated datasets for training and can have limited applicability to new datasets. To

98 tackle some of these issues, several software packages have been developed in recent years,

99 and have become increasingly popular in life science research (Berg et al., 2019; Corliss et

100 al., 2019; Sommer et al., 2011). Prior to segmentation, denoising and feature enhancement

101 methods, such as Hessian-matrix based filtering, can also be applied to overcome the

102 negative impact of noise and/or to enhance certain vessel structures within an image (Oruganti

103 et al., 2013; Ul Haq et al., 2016; H. Zhao et al., 2019).

104 Here, we establish ground truth PAI data based on simulations conducted using

105 synthetic vascular architectures generated *in silico* and, also using a photoacoustic string

106 phantom, composed of a series of synthetic blood vessels (strings) of known structure, which

107 can be imaged in real-time. Against these ground truths, we compare and validate the

108 performance of two common vessel segmentation methods, with or without the application of

109 3D Hessian matrix-based vesselness image filtering feature enhancement of blood vessels

110 (steps 2 & 3 in **Figure 1**). Following skeletonisation of the segmentation masks, we perform

111 statistical and topological analyses to establish how segmentation influences the architectural

112 characteristics of a vascular network acquired using PAI (steps 4 & 5 in **Figure 1**). Finally, we

113 apply our segmentation and analysis pipeline to two *in vivo* breast cancer models and

114 undertake a biological validation of the segmentation and subsequent statistical and

115 topological descriptors using *ex vivo* immunohistochemistry (IHC). Compared to a rule-based

116 auto-thresholding method, our findings indicate that a learning-based segmentation, via a

117 random forest classifier, is better able to account for the artefacts observed in our 3D

118 mesoscopic PAI datasets, providing a more accurate segmentation of vascular networks.

119 Statistical and topological descriptors of vascular structure are influenced by the chosen

120 segmentation method, highlighting a need to validate and standardise segmentation methods

121 in PAI for increased reproducibility and repeatability of mesoscopic PAI in biomedical

122 applications.

123

124
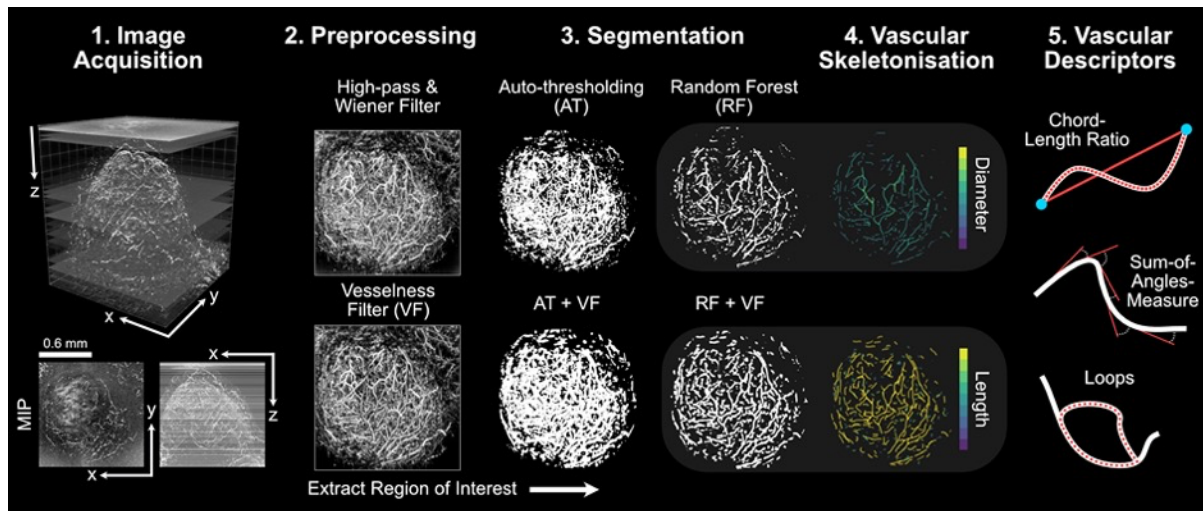
Figure 1. The mesoscopic photoacoustic image analysis pipeline. 1) Images are acquired and reconstructed at a resolution of 20 x 20 x 4 µm^3 (PDX tumour example shown with axial and lateral maximum intensity projections – MIPs). 2) Image volumes are pre-processed to remove noise and homogenise the background signal (high-pass and Wiener filtering followed by slice-wise background correction). Vesselness image filtering (VF) is an optional and additional feature enhancement method. 3) Regions of interest (ROIs) are extracted and segmentation is performed on standard and VF images using auto-thresholding (AT or AT + VF, respectively) or random forest-based segmentation with ilastik (RF or RF + VF, respectively). 4) Each segmented image volume is skeletonised (skeletons with diameter and length distributions shown for RF and RF + VF, respectively). 5) Statistical and topological analyses are performed on each skeleton to quantity vascular structures for a set of vascular descriptors. All images in steps 2-4 are shown as x-y MIPs.

137

138   **RESULTS**

139   ***In silico* simulations of synthetic vasculature enable segmentation precision to be**

140   **evaluated against a known ground truth**

141   Our ground truth consisted of a reference dataset of synthetic vascular network binary

142   masks (n=30) generated from a Lindenmayer System, referred to as L-nets (**Figure 2;**

143   **Supplementary Movie 1** for 3D visualisation). We simulated PAI mesoscopy data from these

144   L-nets (**Figure 2A**) and subsequently used vesselness filtering (VF) as an optional and

145   additional feature enhancement method (**Figure 2B**). The four segmentation pipelines

146   selected for testing (**Figure 1**) were applied to the simulated PAI data (**Figure 2C**), that is, all

147   images were segmented with:

148   1. Auto-thresholding using a moment preserving method (AT);

149   2. Auto-thresholding using a moment preserving method with vesselness filtering pre-

150   segmentation (AT+VF);

151   3. Random forest classifier (RF);

152   4. Random forest classifier with vesselness filtering pre-segmentation (RF+VF).

153   Visually, RF methods appear to segment a larger portion of synthetic blood vessels (**Figure**

154   **2C**) and they are particularly good at segmenting vessels at depths furthest from the simulated

155   light source (**Figure 2D**). A key image quality metric in the context of segmentation is the

156   signal-to-noise (SNR), which is degraded at greater depth (**Figure 3A**). To evaluate the

157   relative performance of the methods, we compared the segmented and skeletonised blood

158   volumes (BV) from the simulated PAI data to the known ground truth from the L-net. Here, we

159   found that the learning-based RF segmentation outperformed the others in making the

160   segmentation masks, with significantly higher R^2 (segmented BV: AT: 0.68, AT+VF: 0.58, RF:

161   0.84, RF+VF: 0.89, **Figure 3B** skeleton BV: AT: 0.59, AT+VF: 0.73, RF: 0.90, RF+VF: 0.93,

162   **Figure 3C**) and lower mean-squared error (MSE) (**Figure 3D**), with respect to the ground truth

163   L-net volumes, compared to both AT methods (p<0.0001 for all comparisons). Bland-Altman

164   plots, which we used to illustrate the level of agreement between segmented and ground truth

165   vascular volumes, showed a mean difference compared to the reference volume of 0.61 mm^3

166   (limits of agreement, LOA -0.48 to 1.7 mm^3, **Figure 3E**) and F1 score of 0.73 ± 0.11 (0.49-

167   0.88) for RF segmentation, albeit with a wide variation indicated by the LOA. RT+VF

168   segmentation resulted in a similar mean difference 0.74 mm^3 (LOA -0.50 to 2.0 mm^3, **Figure**

169   **3F**) and F1 score of 0.66 ± 0.11 (0.44-0.84). In comparison, the rule-based AT segmentation

170   showed poor performance in segmenting vessels at depth (**Figure 2C, Supplementary Movie**

171   **1**), yielding a mean difference of 1.1 mm^3 (LOA -0.60 to 2.8 mm^3) and as with RT+VF, AT+VF

172   did not improve the result, yielding the same mean difference of 1.1 mm^3 (LOA -0.52 to 2.8

173   mm^3) (**Figure 3G,H**). F1 scores were poor for both AT methods, with 0.39 ± 0.10 (0.21-0.59)

174   for AT and 0.37 ± 0.09 (0.16-0.52) for AT+VF.

175   In all cases, the mean difference shown in Bland-Altman plots increased with ground

176   truth vascular volume, especially in the rule-based AT segmentation, which would be expected

177   due to the restricted illumination geometry of photoacoustic mesoscopy. Since more vessel

178   structures lie at a greater distance from the simulated light source in larger L-nets, they suffer

179   from the depth-dependent decrease in SNR (**Figure 3A**). RF segmentation was better able to

180   cope with the SNR degradation, particularly at distances beyond ~1.5 mm, compared to the

181   AT segmentation, which consistently underestimated the vascular volume.

182   Next, we skeletonised each segmentation mask to enable us to perform statistical and

183   topological data analysis (TDA) to test how each segmentation method quantitatively

184   influences a core set of vessel network descriptors (Stolz et al., 2020). These descriptors

185   allowed us to evaluate the performance of the different segmentation methods in respect of

186   the biological characterisation of the tumour networks. We used the following statistical

187   descriptors: vessel diameters and lengths, vessel tortuosity (sum-of-angles measure, SOAM)

188   and vessel curvature (chord-to-length ratio, CLR). Our topological network descriptors are

189   connected components (Betti number β_0) and looping structures (1D holes, Betti number β_1)

190   (see **Supplementary Table 1** for descriptor descriptions).

191   Here, the accuracy and strength of relationship between the segmented and ground

192   truth vascular descriptors, calculated by MSE (see **Figure 3D**) and R^2 values (**Supplementary**

193   **Figure 1A-I**) respectively, gave the same conclusions. Across all skeletons, we measured an

194   increased number of connected components (β_0) and changes to the number of looping

195   structures (β_1) from the simulated compared to the ground truth L-nets, resulting in low R^2 and

196   high MSE for all methods (**Figure 3D**). The observed changes in these topological descriptors

197   arise due to depth-dependent SNR and PAI echo artefacts. For all other descriptors, AT+VF

198   outperformed the other segmentation methods in its ability to accurately preserve the

199   architecture of the L-nets, with higher R^2 and lowest MSE values for vessel lengths, CLR,

200   SOAM, number of edges and number of nodes (**Figure 3D**).

201   Vessel diameters are accurately preserved by both RF segmentation methods,

202   supporting our observation that these methods perform accurate vascular volume

203   segmentation. We note that the number of edges and nodes are also well preserved by RF

204   and RF+VF. This further supports the high accuracy of both RF methods to segment vascular
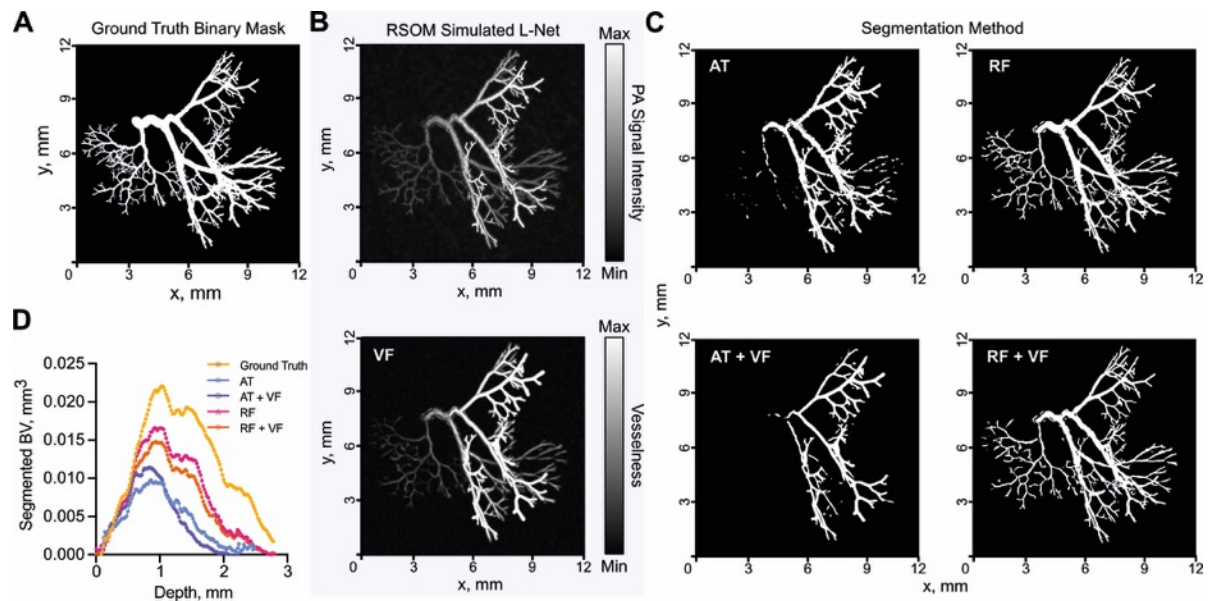
205   structures.

206

207

208

209

210

211

212

213

Figure 2. Exemplar vascular architectures generated *in silico* and processed through our photoacoustic image analysis pipeline. (A-C) XY maximum intensity projections of L-net vasculature. (A) Ground truth L-Net binary mask used to simulate raster-scanning optoacoustic mesoscopy (RSOM) image shown in (B, top) and subsequent optional vesselness filtering (VF) (B, bottom). (C) Segmented binary masks generated using either auto-thresholding (AT), auto-thresholding after vesselness filtering (AT + VF), random forest classification (RF); or random forest classification after vesselness filtering (RF+VF). (D) Segmented blood volume (BV) average across L-net image volumes, plotted against image volume depth (mm). For (D) n=30 L-nets. See Supplementary Movie 1 for 3D visualisation.

223
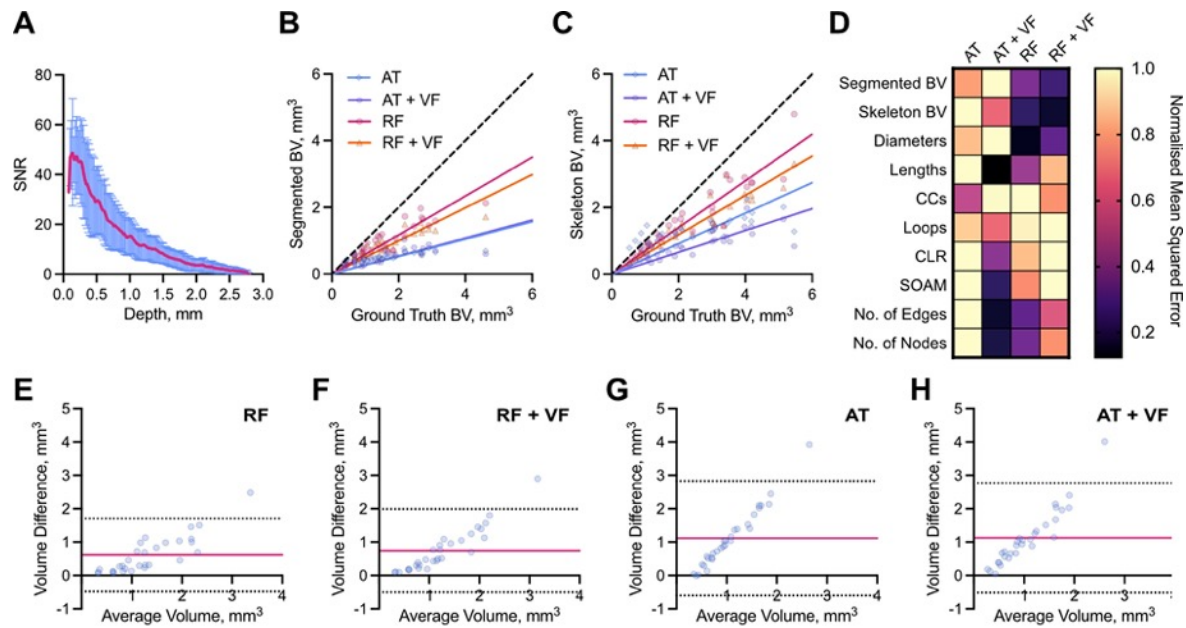
Figure 3. Learning-based random forest classifier outperforms rule-based auto-thresholding in segmenting simulated PAI vascular networks. (A) Depth-wise comparison of signal-to-noise ratio (SNR) measured in PAI-simulated L-nets across depth. (B,C) A comparison between ground truth blood volume (BV) and (B) segmented or (C) skeletonised blood volumes (BV). The dashed line indicates a 1:1 relationship. (D) Heat map displaying normalised (with respect to the maximum of each individual descriptor) mean-squared error comparing our vascular descriptors, calculated from segmented and skeletonised L-nets compared to ground truth L-nets, to each segmentation method. Abbreviations defined: connected components, $\beta 0$ (CC), chord-to-length ratio (CLR), sum-of-angle measure (SOAM). (E-H) Bland-Altman plots comparing blood volume measurements from ground truth L-nets with that of each segmentation method: (E) RF, (F) RF+VF, (G) AT, (H) AT+VF. Pink lines indicate mean difference to ground truth, whilst dotted black lines indicate limits of agreement (LOA). For all subfigures n=30 L-nets.

239 **Random forest classifier accurately segments a string phantom**

240   We next designed a phantom test object (**Supplementary Figure 2**) to further

241 compare the performance of our segmentation pipelines in a ground truth scenario. Agar

242 phantom images (n=7) were acquired using a photoacoustic mesoscopy system and

243 contained three strings of the same known diameter (126 μm), length (~8.4 mm) and

244 consequently volume (104.74 $μm^3$), positioned at 3 different depths, 0.5 mm, 1 mm, and 2

245 mm, respectively (**Figure 4A,B; Supplementary Movie 2**). Consistent with our *in silico*

246 experiments, the accuracy of skeletonised string volumes decreased as a function of depth

247 across all methods (**Figure 4C**), due to the decreased SNR with depth (**Figure 4D**).

248 Interestingly, the significance of this decrease was very high for all comparisons (top vs.

249 middle, top vs. bottom and middle vs. bottom) in both AT methods (all $p<0.001$), but we

250 observed an improvement in string volume predictions across depth for both RF methods,

251 such that middle vs. bottom string volumes were not significantly different in RF+VF ($p=0.42$).

252   The illumination geometry of the photoacoustic mesoscopy system means that vessels

253 or strings are underrepresented when detected as the illumination source is located at the top

254 surface of the tissue or phantom (**Figure 4E**). As a result, all string volumes computed from

255 the segmented images are inaccurate relative to ground truth suggesting that blood volume

256 cannot be accurately predicted from segmented PA images (**Figure 4F**). Skeletonisation

257 provides a more accurate prediction of vessel and string volume as it approximates the

258 undetected section by representing these objects as axisymmetric tubes (**Figure 4C,F**).

259

260
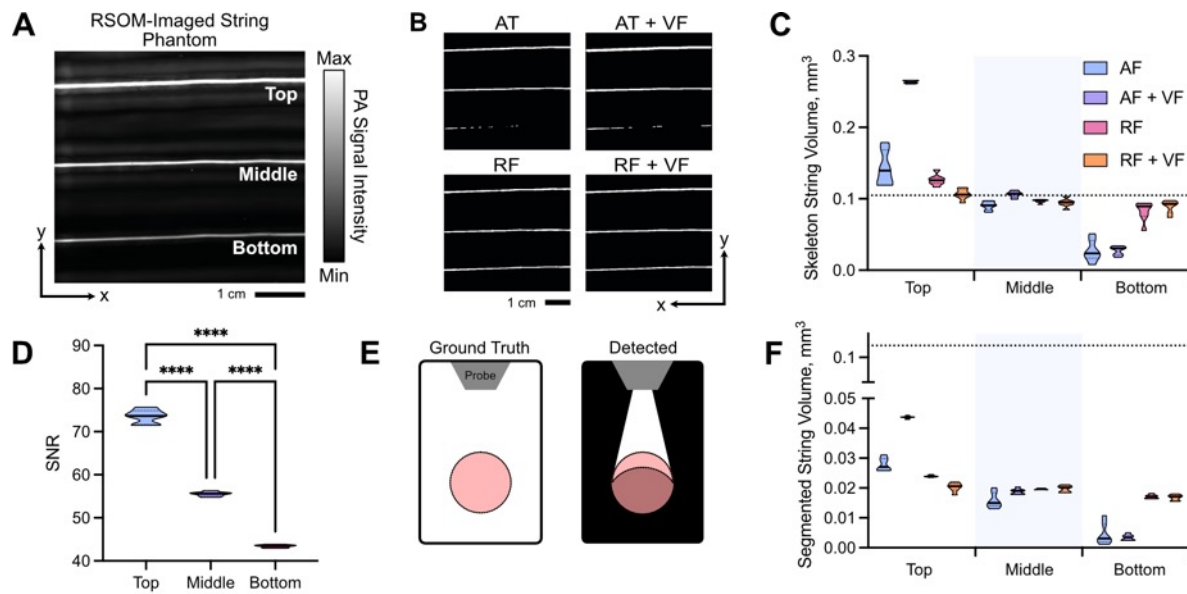
Figure 4. Random forest classifier outperforms auto-thresholding in segmenting a string phantom. XY maximum intensity projections of string phantom imaged with RSOM show that random forest-based segmentation outmatches auto-thresholding when correcting for depth-dependent SNR. (A) Photoacoustic mesoscopy (RSOM) image shows measured string PA signal intensity with top (0.5 mm), middle (1 mm) and bottom (2 mm) strings labelled. (B) Binary masks are shown following segmentation using: (AT) auto-thresholding; (RF) Random forest classifier; (AT+VF) vesselness filtered strings with auto-thresholding; and (RF+VF) vesselness filtered strings with random-forest classifier. (C) Skeletonised string volume calculated from segmented images of 3 strings placed at increasing depths in an agar phantom. Results from all 4 segmentation pipelines are shown. All volume comparisons (top vs. middle, top vs. bottom, middle vs. bottom) where significant (p<0.05) except middle vs. bottom for RF+VF (p=0.42). (D) SNR decreases with increasing depth. (E) Illumination geometry: known cross-section of string outlined (left); during measurement, signal is detected from the partially illuminated section (outlined) resulting in an underestimation in string volume (right). (F) String volume calculated pixel-wise from the segmented binary mask. (C,D,F) Data represented by truncated violin plots with interquartile range (bold) and median (dotted), ****=p<0.0001 (n=7 scans). (C,F) Dotted line indicates ground truth volume 0.105 mm^3. See Supplementary Movie 2 for 3D visualisation.

279

Vesselness filtering of in vivo tumour images impacts computed blood volume

281   Having established the performance of our AT- and RF-based segmentation methods

282   *in silico* and in a string phantom, next we sought to determine the influence of the chosen

283   method in quantifying tumour vascular networks from size-matched breast cancer patient-

284   derived xenograft (PDX) tumours of two subtypes (ER- *n*=6; ER+ *n*=8, total *n*=14).

285   Visual inspection of the tumour networks subjected to our processing pipelines

286   suggests that VF increases vessel diameters *in vivo* (**Figure 5A-C; see Supplementary**

287   **Movie 3** for 3D visualisation). This could be due to acoustic reverberations observed

288   surrounding vessels *in vivo*, which VF scores with high vesselness, spreading the apparent

289   extent of a given vessel and ultimately increased volume. Our quantitative analysis confirmed

290   this hypothesis, where significantly higher skeletonised blood volumes were calculated in the

291   AT+VF and RF+VF masks compared to AT and RF alone (**Figure 5D**).

292

293
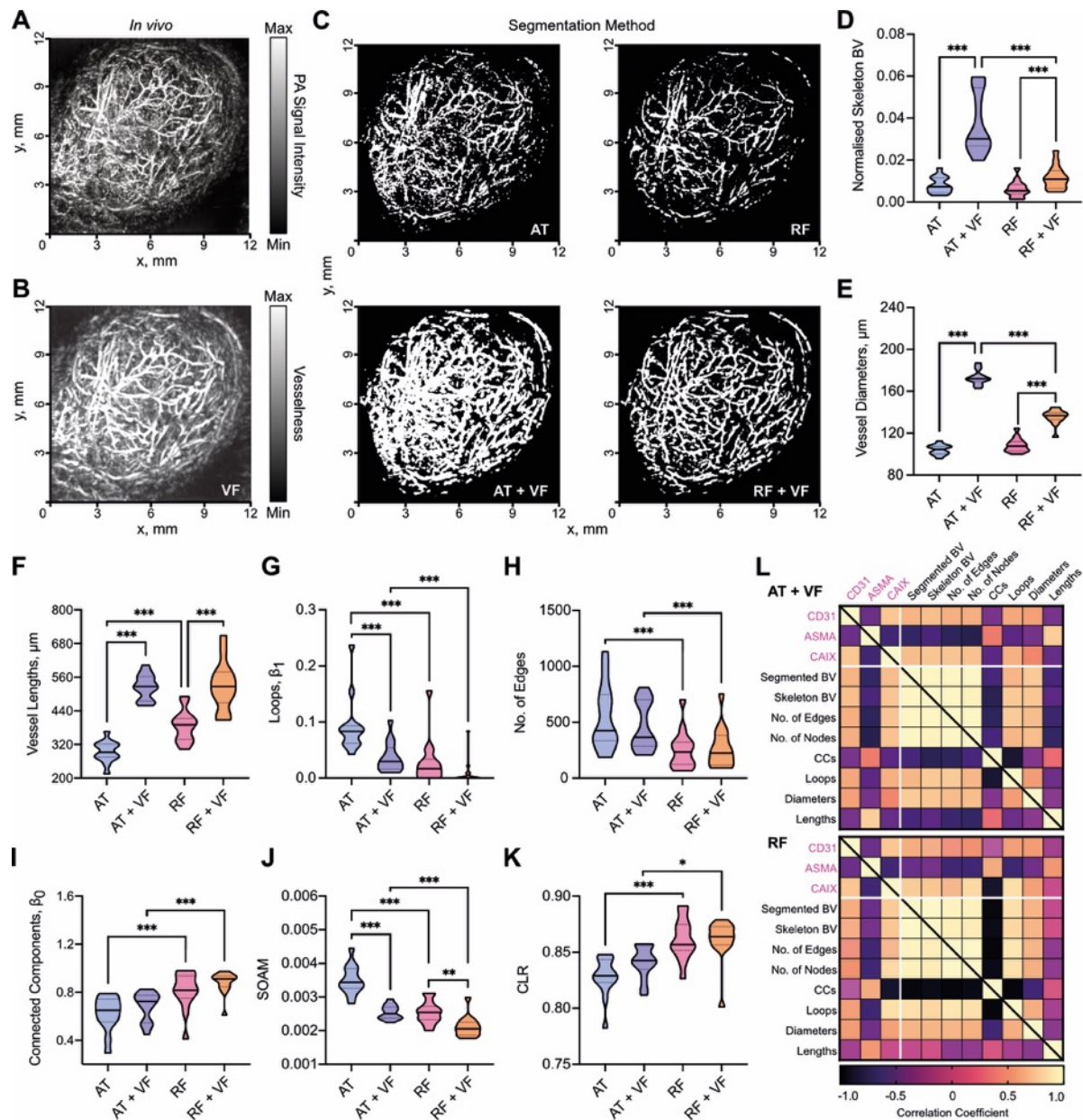
Figure 5. Vesselness filtering increases blood volume calculations from *in vivo* tumour images. XY Maximum intensity projections of breast PDX tumours imaged with RSOM: (A) original image before segmentation; (B) original image with vesselness filtering (VF) applied; (C) a panel showing segmentation with each method (AT: auto-thresholding, AT+VF: auto-thresholding with VF, RF: random forest classifier, and RF + VF: random forest with VF). (D) Skeletonised tumour blood volume (BV) from all 4 segmentation methods normalised to ROI volume. Statistical and topological data analyses were performed on skeletonised tumour vessel vascular networks for the following descriptors: (E) Total number of edges; (F)

302 Connected components normalised by network volume, $\beta 0$; (G) loops normalised by network

303 volume, $\beta 1$; (H) sum-of-angle measure (SOAM); (I) vessel lengths; (J) vessel diameters; (K)

304 chord-to-length ratio (CLR). In (D-K), data are represented by truncated violin plots with

305 interquartile range (dotted) and median (bold). Pairwise comparisons of AT vs. AT+VF, AT vs.

306 RF, RF vs. RF+VF and AT+VF vs. RF+VF calculated using a linear mixed effects model (*=

307 $p < 0.05$, **$= p < 0.01$, ***$= p < 0.001$,). L) Matrix of correlation coefficients for comparisons

308 between IHC, BV and vascular descriptors for (top) AT+VF and (bottom) RF segmented

309 networks. Pearson or spearman coefficients are used as appropriate, depending on data

310 distribution. For (D) n=14, (E-K) n=13 due to imaging artefact in one image which will impact

311 our vascular descriptors. For (L) comparisons involving BV n=14, all other vascular descriptors

312 n=13. See Supplementary Movie 3 for 3D visualisation.

313

314 **Network structure analyses and comparisons to *ex vivo* immunohistochemistry of**

315 **tumour vasculature are impacted by the choice of segmentation method**

316   Next, we computed vascular descriptors for our dataset of segmented *in vivo* images.

317 As expected from our initial *in silico* and phantom evaluations, VF led to increased vessel

318 diameters and lengths (**Figure 5E,F**), as well as blood volume. Our *in silico* analysis indicated

319 that AT performs poorly in differentiating vessels from noise and introduces many vessel

320 discontinuities (**Supplementary Table 1**). This was exacerbated *in vivo* where more complex

321 vascular networks and real noise lead to an increase in segmented blood volume ($p<0.01$),

322 looping structures (**Figure 5G**), a greater number of edges (**Figure 5H**), and reduced number

323 of connected components (**Figure 5I**).

324   Our prior *in silico* and phantom experiments indicate that RF-based methods have a

325 greater capacity to segment vessels at depth. Similarly, we observe more connected

326 components for RF-based methods *in vivo* (**Figure 5I**) along with lower SOAM (**Figure 5J**)

327 and higher CLR (**Figure 5K**), suggesting that RF-segmented vessels have reduced tortuosity

328 and curvature compared to AT+VF segmented vessels. These *in vivo* findings support our

329 observations from *in silico* and phantom studies where RF-based methods provide the most

330 reliable prediction of vascular volume, whereas AT+VF best preserves architecture towards

331 the tissue surface.

332   Next, we sought to assess how our vascular metrics correlated with the following *ex*

333 *vivo* IHC descriptors: CD31 staining area (to mark vessels), ASMA vessel coverage (as a

334 marker of pericyte/smooth muscle coverage and vessel maturity) and CAIX (as a marker of

335 hypoxia) to provide *ex vivo* biological validation of our *in vivo* descriptors. Our *in silico,*

336 phantom and *in vivo* analyses indicate that AT+VF and RF are the top performing

337 segmentation methods and so we focussed on these (results for AT and RF+VF can be found

338 in **Supplementary Figure 3**). We note that none of the vascular metrics derived from AT

339 segmented networks correlated with IHC descriptors.

340   Both AT+VF and RF skeletonised blood volume correlate with CD31 staining area

341   (r=0.54, p=0.05; and r=0.61, p=0.02 respectively; **Figure 5L**). This is as expected as elevated

342   CD31 indicates a higher number of blood vessels and, consequently, higher vascular volume.

343   The following correlations are observed for ASMA vessel coverage: vessel diameters (r=-0.41,

344   p=0.17; and r=-0.43, p=0.14, respectively); looping structures (r=-0.68, p=0.01; and r=-0.58,

345   p=0.04, respectively); number of edges (r=-0.69, p=0.01; and r=-0.65, p=0.02, respectively);

346   number of nodes (r=-0.70, p=0.01; and r=-0.65, p=0.02, respectively); vessel lengths (r=0.76,

347   p=0.03; and r=0.5, p=0.08, respectively); connected components (r=0.38, p=0.22; and r=0.59,

348   p=0.03, respectively). Considering the strengths of AT+VF and RF, these results are

349   biologically intuitive as tumour vessel maturation may lead to higher pericyte coverage, lower

350   vessel density and the pruning of redundant vessels. Elevated pericyte coverage is known to

351   decrease vessel diameters (Barlow et al., 2013), whereas high vessel density resulting from

352   high angiogenesis rates can result in immature vessel networks (Brown et al., 2019). Pruning

353   may lead to a reduction in looping structures and, consequently, an increase in vessel lengths

354   or vascular subnetworks.

355   Finally, levels of hypoxia in the tumours, measured by CAIX IHC, positively correlated

356   in both AT+VF and RF methods with skeletonised blood volume (r=0.72, p=0.007; and r=0.72,

357   p=0.004, respectively), number of edges (r=0.59, p=0.04; and r=0.84, p<0.001, respectively),

358   nodes (r=0.72, p=0.007; and r=0.84, p<0.001, respectively) and looping structures (r=0.61,

359   p=0.03; and r=0.85, p<0.001, respectively). In the case of blood volume, edges and nodes,

360   these results are expected as it has been shown that breast cancer tumours with dense but

361   immature and dysfunctional vasculatures exhibit elevated hypoxia(Brown et al., 2019; Quiros-

362   Gonzalez et al., 2018), likely due to poor perfusion. CAIX negatively correlated with connected

363   components for RF networks (r=-0.87, p<0.001) (**Figure 5L**), reflecting results for ASMA

364   vessel coverage. Our cross-validation between *ex vivo* IHC and vascular descriptors indicate

365   that RF and AT+VF segmentation methods can reliably capture biological characteristics in

366   tumours.

367 **Ex vivo immunohistochemistry and network structural analyses highlight distinct**

368 **vascular networks between ER- and ER+ breast patient-derived xenograft tumours**

369   Finally, we quantified and compared IHC and our vascular descriptors between the

370 two breast cancer subtypes represented (RF in **Figure 6**; AT+VF in **Supplementary Figure**

371 **4**; similar trends and significances are observed unless stated otherwise). From analysis of

372 IHC images (**Figure 6A**), ER- tumours had higher CD31 staining area (**Figure 6B**), poorer

373 ASMA+ pericyte vessel coverage (**Figure 6C**) and higher CAIX levels (**Figure 6D**) compared

374 to ER+ tumours. Our IHC data supports our RF-derived vascular descriptors, where we found

375 that ER- tumours had denser networks, with higher blood volume, diameter and looping

376 structures (**Figure 6E,F,G**). ER+ tumours have a sparse network but showed more

377 subnetworks (**Figure 6H**) with significantly longer vessels in AT+VF segmented networks

378 ($p < 0.05$, **Supplementary Figure 4C**), which could indicate a more mature vessel network

379 based on our prior correlative analyses. No significant differences between the two models

380 were observed for blood vessel tortuosity and curvature (**Figure 6J,K**).

381

382
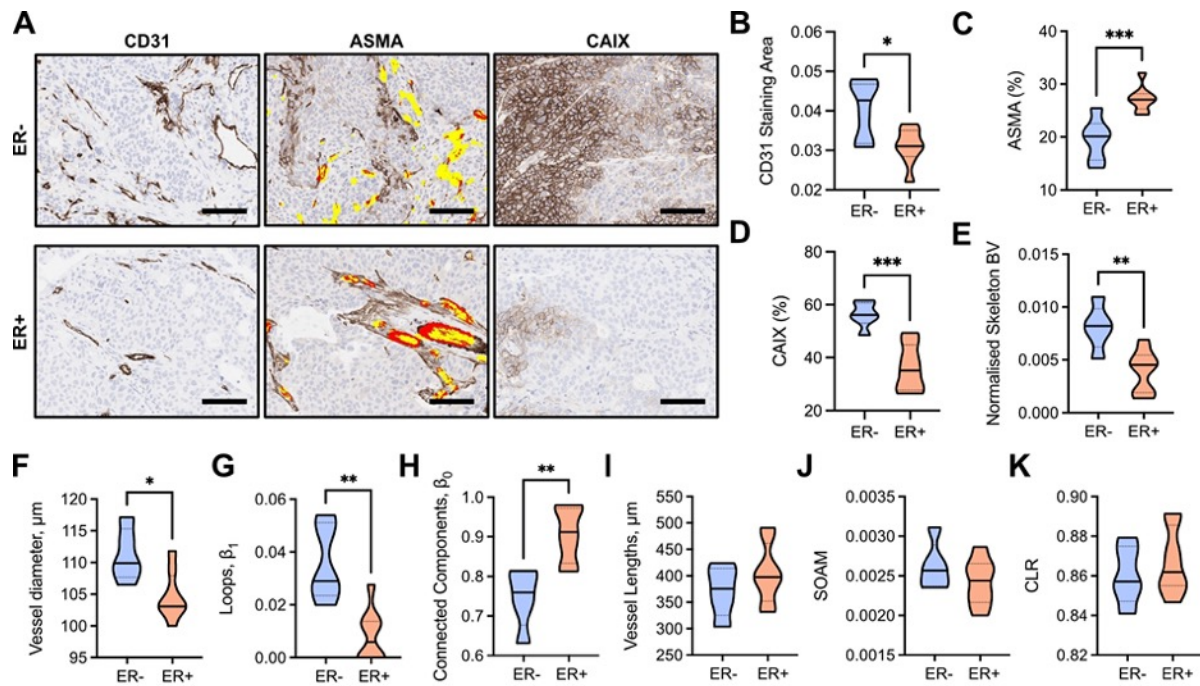
Figure 6. ER- PDX tumours have dense and immature vascular networks which result in hypoxic tumour tissue. (A) Exemplar IHC images of CD31, ASMA and CAIX stained ER- and ER+ tumours. Scale bar=100µm. Brown staining indicates positive expression of marker. ASMA sections display CD31 overlay, where red indicates areas where CD31 and ASMA are colocalised (ASMA vessel coverage) and yellow indicates areas where CD31 is alone. (B) CD31 staining area quantified from CD31 IHC sections and normalised to tumour area. (C) ASMA vessel coverage of CD31+ vessels (number of red pixels/number of red+yellow pixels, expressed as a percentage) on ASMA IHC sections. (D) CAIX total positive pixels as a percentage of the total tumour area pixels on CAIX IHC sections. (E-K) Statistical and topological data analyses comparing ER- and ER+ tumours. Data are represented by truncated violin plots with interquartile range (dotted black) and median (solid black). Comparisons between ER- and ER+ tumours made with unpaired t-test. *= p<0.05, **=p<0.01, ***=p<0.001. For (B-E) ER- n=6, ER+ n=8. For (F-K) ER- n=5, ER+ n=8, one ER- image excluded with artefact that would impact the measured vascular descriptors.

399 **DISCUSSION**

400

401   Mesoscopic PAI enables longitudinal visualisation of blood vessel networks at high

402 resolution, non-invasively and at depths beyond the optical diffraction limit of 1 mm

403 (Ntziachristos, 2010; Ntziachristos et al., 2005; Omar et al., 2019; Wang & Yao, 2016). To

404 quantify the vasculature, PA images need to be accurately segmented. Manual annotation of

405 vasculature in 3D PAI is difficult due to depth-dependent signal-to-noise and imaging artefacts.

406 Whilst a plethora of vascular segmentation techniques are available (Corliss et al., 2019;

407 Moccia et al., 2018), their application in PAI has been limited due to a lack of an available

408 ground truth for comparison and validation.

409   In this study, we first sought to address the need for ground truth data in PAI

410 segmentation. We generated two ground truth datasets to assess the performance of rule-

411 based and machine learning-based segmentation approaches with or without feature

412 enhancement via vesselness filtering. The first is an *in silico* dataset where PAI was simulated

413 on 3D synthetic vascular architectures; the second is an experimental dataset acquired from

414 a vessel-like string phantom. These allowed us to evaluate the ability of different segmentation

415 methods to preserve blood volume and vascular network structure.

416   Our first key finding is that machine learning-based segmentation using RF

417 classification provided the most accurate segmentation of vessel volumes across our *in silico*,

418 phantom and *in vivo* datasets, particularly at depths beyond ~1.5mm, where SNR diminishes

419 due to optical attenuation. Compared to the AT approaches, RF-based segmentation partially

420 overcomes the depth dependence of PAI SNR since it identifies and learns edge and texture

421 features of vessels at different scales and contrasts. Such intrinsic depth-dependent

422 limitations are often ignored in the literature, where analyses are typically performed on 2D

423 maximum intensity projections for simplicity (Haedicke et al., 2020; Imai et al., 2017; Lao et

424 al., 2008; Omar et al., 2015; Orlova et al., 2019; Soetikno et al., 2012), suggesting that a fully

425   3D machine learning-based segmentation is needed to accurately recapitulate the complexity

426   of *in vivo* vasculatures measured using PAI.

427   As blood vessel networks can be represented as complex, interconnected graphs, we

428   performed statistical and topological data analyses (Chung et al., 2019; Stolz et al., 2020) to

429   further assess the strengths and weaknesses of our chosen segmentation methods.

430   Our second key finding is that AT methods struggle to segment vessels with low SNR,

431   but adding VF outperforms all other methods in preserving vessel lengths, loops, curvature

432   and tortuosity. Additionally, where intensity varies across a vessel structure, this results in

433   many disconnected vessels when segmenting with AT alone, as only the highest intensity

434   voxels will pass the threshold. Only when vesselness filtering is applied does AT do well at

435   preserving topology. VF alters the intensity values from a measure of PA signal to a prediction

436   of 'vesselness', generating a more homogeneous intensity across the vessel structures and

437   ultimately a more continuous vessel structure to segment. This likely explains why AT+VF best

438   preserves vessel length and, subsequently, network structure, while AT alone performs poorly.

439   For AT, VF improved BV predictions *in silico* via better preservation of lengths but not

440   diameters, as our phantom experiments indicate that AT+VF overestimates diameter.

441   Owing to the homogenous intensity of vessels introduced by VF, one could therefore

442   assume that RF+VF would be the most accurate method at preserving network structure (by

443   combining the machine-learning accuracy of segmentation with the shape enhancement of

444   VF). However, this is not the case: RF alone can account for discontinuities in vessel intensity,

445   unlike AT, meaning it does not rely on VF to enhance structural preservation, which is our third

446   key finding. In fact, the slight inaccuracy in diameter preservation introduced by VF *in silico*

447   appears to decrease topology preservation in RF+VF compared to RF alone. As expected, all

448   methods led to an increase in the number of subnetworks (connected components) *in silico*,

449   as these segmentation methods cannot reconnect vessel subnetworks that were disconnected

450   due to poor SNR or imaging artefacts. Given the better segmentation at depth by RF-methods,

451   we hypothesise that these increasingly small subnetworks might have biased the

452   segmentations to underperform in our vascular descriptors. This could be explored in future

453   work, for example, by developing string phantoms with more complex topologies.

454   Taken together, our results suggest that RF performs feature detection across scales

455   in the manually labelled voxels to learn discriminating characteristics for vessel classification

456   and segmentation. Adding VF before RF segmentation may confound this segmentation

457   framework, because VF systematically smooths images and removes non-cylindrical raw

458   image information, which may have been vital in the RF learning of vascular structures on the

459   training dataset.

460   Applying statistical and topological analyses to our *in vivo* tumour PDX dataset we

461   observed trends consistent with our *in silico and* phantom experiments. Cross-validating our

462   vascular descriptors with *ex vivo* IHC confirmed that we can extract biologically relevant

463   information from mesoscopic PA images. For example, predictions of BV correlated with

464   endothelial cell and hypoxia markers via CD31 and CAIX staining, respectively; and

465   descriptors relating to the maturation of vascular structures correlated with ASMA vessel

466   coverage. Applying our segmentation pipeline to compare ER- and ER+ breast cancer PDX

467   models showed that descriptors of network structure can capture the higher density and

468   immaturity of ER- vessel networks which result in decreased oxygen delivery and high hypoxia

469   levels in comparison to ER+ tumours.

470   While our pipeline yields encouraging correlations to the underlying tumour

471   vasculature, avenues of further development exist to: improve the realism of our ground truth

472   data, including advances in simulation complexity, and tissue-specific synthetic and phantom

473   vasculatures. While our *in silico* PAI dataset incorporated the effects of depth-dependent SNR

474   and gaussian noise found in *in vivo* PAI mesoscopic data, further development of the optical

475   simulations could, for example, recapitulate the raster-scanning motion of illumination optical

476   fibres, instead of approximating a simultaneous illumination plane of single-point sources. The

477   limited aperture of the raster-scanning ultrasound transducer could not be simulated in k-Wave

478   as it is not yet implemented for 3D structures. In terms of vascular complexity, our string

479   phantom represents a highly idealised vessel networks but future work could introduce more

480   complex and interconnected vessel-like networks in order to replicate more realistic vascular

481   topologies (Dantuma et al., 2019). Our *ex vivo* IHC descriptors were used to confirm our *in*

482   *vivo* tumour analyses but did not exhibit correlations across all vascular descriptors. This may

483   be expected as the 2D IHC analysis does not fully encompass the 3D topological

484   characteristics of the vascular network. 3D IHC, microCT or light sheet fluorescence

485   microscopy may provide improved *ex vivo* validation using exogenous labelling to identify 3D

486   vascular structures, such as tortuosity, at endpoint (Epah et al., 2018; Hlushchuk et al., 2019).

487   It should also be noted that we cannot discount the effect of unconscious biases on

488   segmentation performance when manually labelling images with and without VF to train the

489   classifier. The segmentation accuracy of classifiers trained by multiple users could be explored

490   in future work to formally investigate these effects.

491   Furthermore, the past decade has seen the rise of a multitude of blood vessel

492   segmentation methods using convolutional neural networks and deep learning (Jia & Zhuang,

493   2021). Applying deep learning to mesoscopic PAI could provide a means to overcome several

494   equipment-related limitations such as: vessel discontinuities induced by breathing motion *in*

495   *vivo*; vessel orientation relative to the ultrasound transducer; shadow and reflection artefacts;

496   or underestimation of vessel diameter in the z-direction due to surface illumination. Whilst we

497   found that skeletonisation addressed diameter underestimation and observed the influence of

498   discontinuities on the extracted statistical and topological descriptors, they were not deeply

499   characterised or corrected. Nonetheless, whilst deep learning may provide superior

500   performance when fine-tuned to specific tasks, the resulting methods may lack generalisability

501   across tissues with differing SNR and blood structures, requiring large datasets for training. In

502   this study we chose to use software that is open-source and widely accessible to biologists in

503   the life sciences. We believe that such a platform shows more potential to be employed widely

504   with limited computational expertise.

505   In summary, we developed an *in silico,* phantom*, in vivo*, and *ex vivo*-validated end-to-

506   end framework for the segmentation and quantification of vascular networks captured using

507   mesoscopic PAI. We created *in silico* and string phantom ground truth PAI datasets to validate

508   segmentation of 3D mesoscopic PA images. We then applied a range of segmentation

509   methods to these and images of breast PDX tumours obtained *in vivo*, including cross-

510   validation of *in vivo* images with *ex vivo* IHC. We have shown that learning-based

511   segmentation, via a random forest classifier, best accounted for the artefacts present in

512   mesoscopic PAI, providing a robust segmentation of blood volume at depth in 3D and a good

513   approximation of vessel network structure. Despite the promise of the learning-based

514   approach to account for depth-dependent variation in SNR, auto-thresholding with vesselness

515   filtering more accurately represents statistical and topological characteristics in the superficial

516   blood vessels as it better preserves vessel lengths. Therefore, when quantifying PA images,

517   users need to consider the relative importance of each descriptor as the choice of

518   segmentation method can directly impact the resulting analyses. We have highlighted the

519   potential of statistical and topological analyses to provide a detailed parameterisation of

520   tumour vascular networks, from classic statistical descriptors such as vessel diameters and

521   lengths to more complex descriptors of network topology characterising vessel connectivity

522   and loops. Our results further underscore the potential of photoacoustic mesoscopy as a tool

523   to provide biological insight into studying vascular network *in vivo* by providing life scientists

524   with a readily deployable and cross-validated pipeline for data analysis.

525

526 **MATERIALS AND METHODS**

527 **Generating ground truth vascular architectures in silico**

528   To generate an *in silico* ground truth vascular network, we utilised Lindenmayer

529 systems (L-Systems, see **Supplementary Figure 5**) (Lindenmayer, 1968). L-Systems are

530 language-theoretic models that were originally developed to model cellular interactions but

531 have been extended to model numerous developmental processes in biology (Rozenberg &

532 Arto Salomaa, 1992). Here, we apply L-Systems to generate realistic, 3D vascular

533 architectures (Galarreta-Valverde, 2012; Galarreta-Valverde et al., 2013) (referred to as L-

534 nets) and corresponding binary image volumes. A stochastic grammar was used (Galarreta-

535 Valverde, 2012) to create a string that was evaluated using a lexical and syntactic analyser to

536 build a graphical representation of each L-net. To transfer the L-net to a discretised binary

537 image volume, we used a modified Bresenham's algorithm (Bresenham, 1965) for 3D to create

538 a vessel skeleton. Voxels within a vessel volume were then identified using the associated

539 vessel diameter for each centreline (**Supplementary Figure 5**).

540 **Photoacoustic image simulation of synthetic ground truths**

541   To test the accuracy of the segmentation pipelines, the L-nets were then used to

542 simulate *in vivo* photoacoustic vascular networks embedded in muscle tissue using the

543 Simulation and Image Processing for Photoacoustic Imaging (SIMPA) python package

544 (SIMPA v0.1.1, https://github.com/CAMI-DKFZ/simpa) (Janek Gröhl, Kris K. Dreher, Melanie

545 Schellenberg, Alexander Seitel, 2021) and the k-Wave MATLAB toolbox (k-Wave v1.3,

546 MATLAB v2020b, MathWorks, Natick, MA, USA) (Treeby & Cox, 2010). Planar illumination of

547 the L-nets on the XY plane was achieved using Monte-Carlo eXtreme (MCX v2020, 1.8)

548 simulation on the L-net computational grid of size 10.24 x 10.24 x 2.80 mm^3 with 20 µm

549 isotropic resolution. The optical forward modelling was conducted at 532 nm using the optical

550 absorption spectrum of 50% oxygenated haemoglobin for vessels (an approximation of tumour

551 vessel oxygenation based on previously collected photoacoustic data (Quiros-Gonzalez et al.,

552   2018) and of water for muscle. Next, 3D acoustic forward modelling was performed on the

553   illuminated L-nets assuming a speed of sound of 1500 ms^{-1} in k-Wave. The photoacoustic

554   response of the illuminated L-nets was measured with a planar array of sensors positioned on

555   the surface of the XY plane with transducer elements of bandwidth central frequency of 50

556   MHz (100% bandwidth) and using a 1,504 time steps, where a time step is 5×10^{-8} Hz^{-1}). Finally,

557   the 3D initial PA wave-field was reconstructed using fast Fourier transform-based

558   reconstruction (Treeby & Cox, 2010), after adding uniform gaussian noise on the collected

559   wave-field.

560   **String phantom**

561   We used a string phantom as a ground truth structure (see **Supplementary**

562   **Materials**). The agar phantom was prepared as described previously (Joseph et al., 2017)

563   including intralipid (I141-100ML, Merck, Gillingham, UK) to mimic tissue-like scattering

564   conditions. Red-coloured synthetic fibres (Smilco, USA) were embedded at three different

565   depths defined by the frame of the phantom to provide imaging targets with a known diameter

566   of 126 μm. The top string was positioned at 0.5 mm from the agar surface, the middle one at

567   1 mm, and the bottom one at 2 mm, as shown in **Supplementary Figure 2**.

568   **Animals**

569   All animal procedures were conducted in accordance with project and personal

570   licences, issued under the United Kingdom Animals (Scientific Procedures) Act, 1986 and

571   approved locally under compliance forms CFSB1567 and CFSB1745. For *in vivo* vascular

572   tumour models, cryopreserved breast PDX tumour fragments in freezing media composed of

573   heat-inactivated foetal bovine serum (10500064, Gibco^{TM}, Fisher Scientific, Göteborg

574   Sweden) and 10% dimethyl sulfoxide (D2650, Merck) were defrosted at 37°C, washed with

575   Dulbecco's Modified Eagle Medium (41965039, Gibco) and mixed with matrigel (354262,

576   Corning®, NY, USA) before surgical implantation. One estrogen receptor negative (ER-, n=6)

577   PDX model and one estrogen receptor positive (ER+, n=8) PDX model were implanted

578 subcutaneously into the flank of 6-9 week-old NOD scid gamma (NSG) mice (#005557, Jax

579 Stock, Charles River, UK) as per standard protocols (Bruna et al., 2016). Once tumours had

580 reached ~1cm mean diameter, tumours were imaged and mice sacrificed afterwards, with

581 tumours collected in formalin for IHC.

582 **Photoacoustic imaging**

583 Mesoscopic PAI was performed using the raster-scan optoacoustic mesoscopy

584 (RSOM) Explorer P50 (iThera Medical GmbH, Munich, Germany). The system uses a 532 nm

585 laser for excitation. Two optical fibre bundles are arranged either side of a transducer, which

586 provide an elliptical illumination beam of approximately 4 mm x 2 mm in size. The transducer

587 and lasers collectively raster-scan across the field-of-view. A high-frequency single-element

588 transducer with a centre frequency of 50 MHz (>90% bandwidth) detects ultrasound. The

589 system achieves a lateral resolution of 40 μm, an axial resolution of 10 μm and a penetration

590 depth of up to ~3 mm (Omar et al., 2013).

591 For image acquisition of both phantom and mice, degassed commercial ultrasound gel

592 (AquaSonics Parker Lab, Fairfield, NJ, USA) was applied to the surface of the imaging target

593 for coupling to the scan interface. Images were acquired over a field of view of 12 × 12 mm^2

594 (step size: 20 μm) at either 100% (phantom) or 85% (mice) laser energy and a laser pulse

595 repetition rate of 2 kHz (phantom) or 1 kHz (mice). Image acquisition took approximately 7

596 min. Animals were anaesthetised using 3-5% isoflurane in 50% oxygen and 50% medical air.

597 Mice were shaved and depilatory cream applied to remove fur that could generate image

598 artefacts; single mice were placed into the PAI system, on a heat-pad maintained at 37°C.

599 Respiratory rate was maintained between 70-80 bpm using isoflurane (~1-2% concentration)

600 throughout image acquisition.

601

602

603

604   **Segmentation and extraction of structural and topological vascular descriptors**

605   All acquired data were subjected to pre-processing prior to segmentation,

606   skeletonisation and structural analyses of the vascular network, with an optional step of

607   vesselness filtering also tested (**Figure 1**). Prior to segmentation, data were filtered in the

608   Fourier domain in XY plane to remove reflection lines, before being reconstructed using a

609   backprojection algorithm in viewRSOM software (v2.3.5.2 iThera Medical GmbH) with motion

610   correction for *in vivo* images with a voxel size of 20 x 20 x 4 μm^3 (X,Y,Z). To reduce background

611   noise and artefacts from the data acquisition process, reconstructed images were subjected

612   to a high-pass filter, to remove echo noise, followed by a Wiener filter in MATLAB (v2020b,

613   MathWorks, Natick, MA, USA) to remove stochastic noise. Then, a built-in slice-wise

614   background correction (Sternberg, 1983) was performed in Fiji(Schindelin et al., 2012) to

615   achieve a homogenous background intensity (see exemplars of each pre-processing step in

616   **Supplementary Figure 6**).

617   **Image segmentation using auto-thresholding or a random forest classifier**

618   Using two common tools adopted in the life sciences, we tested both a rule-based

619   moment preserving thresholding method (included in Fiji v2.1.0) and a learning-based

620   segmentation method based on random forest classifiers (with ilastik v1.3.3 (Berg et al.,

621   2019)). These popular packages were chosen to enable widespread application of our

622   findings. Moment preserving thresholding, referred to as *auto-thresholding* (AT) for the

623   remainder of this work, computes the intensity moments of an image and segments the image

624   while preserving these moments (Tsai, 1985). Training of the random forest (RF) backend was

625   performed on 3D voxel features in labelled regions, including intensity features, as with the

626   AT method, combined with edge filters, to account for the intensity gradient between vessels

627   and background, and texture descriptors, to discern artefacts in the background from the

628   brighter and more uniform vessel features, each evaluated at different scales (up to a sigma

629   of 5.0).

630   A key consideration in the machine learning-based segmentation is the preparation of

631   training and testing data (**Supplementary Table 2**). For the *in silico* ground truth L-net data,

632   all voxel labels are known. All vessel labels were used for training, however, only partial

633   background labels were supplied to minimise computational expense by labelling the 10 voxel

634   radius surrounding all vessels as well as 3 planes parallel to the Z-axis (edges and middle) as

635   background (**Supplementary Figure 7A,B)**. For the phantom data, manual segmentation of

636   the strings from background was performed to provide ground truth. Strings were segmented

637   in all slices on which they appeared and background was segmented tightly around the string

638   (**Supplementary Figure 7C**). For the *in vivo* tumour data, manual segmentation of vessels

639   was made by a junior user (TLL) supervised by an experienced user (ELB), including images

640   of varying signal-to-noise ratio (SNR) to increase the robustness of the algorithm for

641   application in a range of unseen data. Up to 10 XY slices per image stack in the training

642   dataset were segmented with pencil size 1 at different depths to account for depth-dependent

643   SNR differences (**Supplementary Figure 7D**).

644   Between pre-processing and segmentation, feature enhancement was tested as a

645   variable in our segmentation pipeline. In Fiji, we adapted a modified version of Sato filtering

646   (α=0.25) (Sato et al., 1998) to calculate vesselness from Hessian matrix eigenvalues (Frangi

647   et al., 1998) across multiple scales. Five scales in a linear Gaussian normalized scale space

648   were used, from which the maximal response was measured to produce the final vesselness

649   filtered images (20, 40, 60, 80, and 100 μm) (Sato et al., 1998).

650   Finally, all segmented images (either from Fiji or ilastik) were passed through a built-

651   in 3D median filter in Fiji, to remove impulse noises (**Supplementary Figure 8**). To summarise

652   the pipeline (**Figure 1**), the methods under test for all datasets were:

653   1.   Auto-thresholding using a moment preserving method (AT);

654   2.   Auto-thresholding using a moment preserving method with vesselness filtering pre-

655   segmentation (AT+VF);

656   3.   Random forest classifier (RF);

657   4.   Random forest classifier with vesselness filtering pre-segmentation (RF+VF).

658   Computation times are summarised in Supplementary Table 3.

Extracting tumour ROIs using a 3D CNN

660   To analyse the tumour data in isolation from the surrounding tissue required

661   delineation of tumour regions of interest (ROIs). To achieve this, we trained a 3D convolutional

662   neural network (CNN) to fully automate extraction of tumour ROIs from PAI volumes. The 3D

663   CNN is based on the U-Net architecture (Ronneberger et al., 2015) extended for volumetric

664   delineation (Çiçek et al., 2016). Details on the CNN architecture and training are provided in

665   the **Supplementary Materials** and **Supplementary Figures 9-10**.

Network Structure and Topological Data Analysis

667   Topological data analysis (TDA) of the vascular networks was performed using

668   previously reported software that performs TDA and structural analyses on vasculature

669   (Chung et al., 2019; Stolz et al., 2020). Prior to these analyses, segmented image volumes

670   were skeletonised using the open-source package Russ-learn (Bates, 2017, 2018). Our

671   vascular descriptors comprised a set of statistical descriptors: vessel diameters and lengths,

672   vessel tortuosity (sum-of-angles measure, SOAM) and curvature (chord-to-length ratio, CLR),

673   In addition, the following descriptors were used to define network topology: the number of

674   connected components (Betti number β_0) and looping structures (1D holes, Betti number β_1).

675   Full descriptions of the vascular descriptors are provided in **Supplementary Table 1** while

676   outputs are shown in **Supplementary Tables 4-7.**

Immunohistochemistry

678   For *ex vivo* validation, formalin-fixed paraffin-embedded (FFPE) tumour tissues were

679   sectioned. Following deparaffinising and rehydration, IHC was performed for the following

680   antibodies: CD31 (anti-mouse 77699, Cell signalling, London, UK), α-smooth muscle actin

681 (ASMA) (anti-mouse ab5694, abcam, Cambridge, UK), carbonic anhydrase-IX (CAIX) (anti-

682 human AB1001, Bioscience Slovakia, Bratislava, Slovakia) at 1:100, 1:500 and 1:1000,

683 respectively, using a BOND automated stainer with a bond polymer refine detection kit (Leica

684 Biosystems) and 3,3'-diaminobenzadine as a substrate. Stained FFPE sections were scanned

685 at 20x magnification using an Aperio ScanScope (Leica Biosystems, Milton Keynes, UK) and

686 analysed using ImageScope software (Leica Biosystems) or HALO Software (v2.2.1870,

687 Indica Labs, Albuquerque, NM, USA). ROIs were drawn over the whole viable tumour area

688 and built-in algorithms customised to analyse the following: CD31 positive area (μm^2)

689 normalised to the ROI area (μm^2) (referred to as CD31 vessel area), area of CD31 positive

690 pixels (μm^2) colocalised on adjacent serial section with ASMA positive pixels/CD31 positive

691 area (μm^2) (reported as ASMA vessel coverage (%)) and CAIX positive pixel count per total

692 ROI pixel count (reported as CAIX (%)).

693 **Statistical analysis**

694   Statistical analyses were conducted using Prism (v9, GraphPad Software, San Diego,

695 CA, USA) and R (v4.0.1(R Core, 2021), R Foundation, Vienna, Austria). We used the mean

696 square error and R-squared statistics to quantify the accuracy and strength of the relationship

697 between the segmented networks to the ground truth L-nets. For each outcome of interest,

698 we predicted the ground truth (on a scale compatible with the normality assumption according

699 to model checks) by means of each method estimates through a linear model. As model

700 performance statistics are typically overestimated when assessing the model fit on the same

701 data used to estimate the model parameters, we used bootstrapping (R = 500) to correct for

702 the optimism bias and obtain unbiased estimates (Harrell, 2016). Bland-Altman plots were

703 produced for each paired comparison of segmented volume to the ground truth volume in L-

704 nets and associated bias and limits of agreement (LOA) are reported. For L-nets, F1 scores

705 were calculated (Dice, 1945). PAI quality pre-segmentation was quantified by measuring SNR,

706 defined as the mean of signal over the standard deviation of the background signal.

707 Comparisons of string volume, as well as SNR, were completed using one-way ANOVA with

708 Tukey multiplicity correction.

709   For each outcome of interest, *in vivo* data was analysed as follows: A linear mixed

710 effect model was fitted on a response scale (log, square root or cube root) compatible with the

711 normality assumption according to model checks with the segmentation methods as a 4-level

712 fixed predictor and animal as random effect, to take the within mouse dependence into

713 account. Noting that the residual variance was sometimes different for each segmentation

714 group, we also fitted a heteroscedastic linear mixed effect allowing the variance to be a

715 function of the segmentation group. The results of the heteroscedastic model were preferred

716 to results of the homoscedastic model when the likelihood ratio test comparing both models

717 led to a p-value <0.05. Two multiplicity corrections were performed to achieve a 5% family-

718 wise error rate for each dataset: For each outcome, a parametric multiplicity correction on the

719 segmentation method parameters was first used (Bretz et al., 2010). A conservative Bonferroni

720 p-value adjustment was then added to it to account for the number of outcomes in the entire

721 *in vivo* dataset. The following pairwise comparisons were considered: AT vs. AT+VF, AT vs.

722 RF, RF vs. RF+VF and AT+VF vs. RF+VF. Comparisons of our vascular descriptors between

723 ER- and ER+ tumours were completed with an unpaired student's t-test. All p-values <0.05

724 after multiplicity correction were considered statistically significant.

725 **Code Availability**

726 Code to generate synthetic vascular trees (LNets) is available on GitHub

727 (https://github.com/psweens/V-System). *In silico* photoacoustic simulations were performed

728 using the SIMPA toolkit (https://github.com/CAMI-DKFZ/simpa). Both the trained 3D CNN to

729 extract tumour ROIs from RSOM images (https://github.com/psweens/Predict-RSOM-ROI)

730 and vascular TDA package are available on GitHub (https://github.com/psweens/Vascular-

731 TDA).

732 **Data Availability**

733 Exemplar datasets for the *in silico*, phantom, and *in vivo* data can be found at

734 https://doi.org/10.17863/CAM.78208. The authors declare that all data supporting the findings

735 of this study is available upon request.

736 **ACKNOWLEDGEMENTS**

753

754 **COMPETING INTERESTS**

755 The authors have no conflict of interest related to the present manuscript to disclose.

756

757 **AUTHOR CONTRIBUTIONS**

758 Conceptualization: PWS, ELB, TLL, SEB

759 Methodology: PWS, ELB, LH, TLL, ZH, SEB

760 Software: PWS, BJS, JG, TLL, ZH

761 Validation: PWS, ELB, LH, TLL

762 Formal Analysis: PWS, ELB, TLL

763 Investigation: PWS, ELB, LH, TLL

764 Resources: HAH, HMB

765 Data Curation: PWS, JG, TLL

766 Writing – original draft: PWS, ELB, LH, DLC, TLL, SEB

767 Writing – review & editing: PWS, ELB, JG, LH, BJS, HAH, HMB, TLL, SEB

768 Visualisation: PWS, ELB, TLL

769 Supervision: SEB

770 Project Administration: PWS, SEB

771 Funding Acquisition: PWS, TLL, SEB

772

773

774

775

776 **REFERENCES**

777   Abadi, M., Agarwal, A., Paul Barham, E. B., Zhifeng Chen, Craig Citro, Greg S.

778 Corrado, A. D., Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, I. G., Andrew Harp, Geoffrey

779 Irving, Michael Isard, Rafal Jozefowicz, Y. J., Lukasz Kaiser, Manjunath Kudlur, Josh

780 Levenberg, Dan Mané, M. S., Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, J. S.,

781 Benoit Steiner, Ilya Sutskever, Kunal Talwar, P. T., Vincent Vanhoucke, Vijay Vasudevan, F.

782 V., Oriol Vinyals, Pete Warden, Martin Wattenberg, M. W., & Yuan Yu, and X. Z. (2015).

783 TensorFlow: Large-scale machine learning on heterogeneous systems. *12th USENIX*

784 *Symposium on Operating Systems Design and Implementation*, 265–278.

785 *Autonomio Talos*. (2019).

786 Barlow, K. D., Sanders, A. M., Soker, S., Ergun, S., & Metheny-Barlow, L. J. (2013).

787   Pericytes on the tumor vasculature: jekyll or hyde? *Cancer Microenvironment : Official*

788   *Journal of the International Cancer Microenvironment Society*, *6*(1), 1–17.

789   https://doi.org/10.1007/s12307-012-0102-2

790 Bates, R. (2017). *Learning to Extract Tumour Vasculature: Techniques in Machine Learning*

791   *for Medical Image Analysis*. University of Oxford.

792 Bates, R. (2018). *Russ-learn: set of tools for application and training of deep learning*

793   *methods for image segmentation and vessel analysis*. GitLab.

794 Beard, P. (2011). Biomedical Photoacoustic Imaging. *Interface Focus*, *1*(4), 602–631.

795   https://doi.org/10.1098/rsfs.2011.0028

796 Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., Schiegg, M.,

797   Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J. I., Xu, B., Beuttenmueller, F.,

798   Wolny, A., Zhang, C., Koethe, U., Hamprecht, F. A., & Kreshuk, A. (2019). ilastik:

799   interactive machine learning for (bio)image analysis. *Nature Methods*, *16*(12), 1226–

800   1232. https://doi.org/10.1038/s41592-019-0582-9

801 Bresenham, J. E. (1965). Algorithm for computer control of a digital plotter. *IBM Systems*

802   *Journal*, *4*(1). https://doi.org/10.1147/sj.41.0025

803 Bretz, F., Hothorn, T., & Westfall, P. (2010). *Multiple comparisons Using R*. Chapman and

804   Hall/CRC. https://doi.org/10.1201/9781420010909

805   Brown, E., Brunker, J., & Bohndiek, S. E. (2019). Photoacoustic imaging as a tool to probe

806   the tumour microenvironment. *Disease Models &Amp; Mechanisms*, *12*(7),

807   dmm039636. https://doi.org/10.1242/dmm.039636

808   Bruna, A., Rueda, O. M., Greenwood, W., Batra, A. S., Callari, M., Batra, R. N., Pogrebniak,

809   K., Sandoval, J., Cassidy, J. W., Tufegdzic-Vidakovic, A., Sammut, S.-J., Jones, L.,

810   Provenzano, E., Baird, R., Eirew, P., Hadfield, J., Eldridge, M., McLaren-Douglas, A.,

811   Barthorpe, A., … Caldas, C. (2016). A Biobank of Breast Cancer Explants with

812   Preserved Intra-tumor Heterogeneity to Screen Anticancer Compounds. *Cell*, *167*(1),

813   260-274.e22. https://doi.org/10.1016/j.cell.2016.08.041

814   Chollet, F., & Others. (2015). *Keras*. Keras.Io.

815   Chung, M. K., Lee, H., DiChristofano, A., Ombao, H., & Solo, V. (2019). Exact topological

816   inference of the resting-state brain networks in twins. *Network Neuroscience*, *3*(3), 674–

817   694. https://doi.org/10.1162/NETN_A_00091

818   Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-net:

819   Learning dense volumetric segmentation from sparse annotation. *Medical Image*

820   *Computing and Computer-Assisted Intervention – MICCAI 2016. MICCAI 2016. Lecture*

821   *Notes in Computer Science*, *9901*. https://doi.org/10.1007/978-3-319-46723-8_49

822   Corliss, B. A., Mathews, C., Doty, R., Rohde, G., & Peirce, S. M. (2019). Methods to label,

823   image, and analyze the complex structural architectures of microvascular networks.

824   *Microcirculation (New York, N.Y. : 1994)*, *26*(5), e12520–e12520.

825   https://doi.org/10.1111/micc.12520

826   Crum, W. R., Camara, O., & Hill, D. L. G. (2006). Generalized overlap measures for

827   evaluation and validation in medical image analysis. *IEEE Transactions on Medical*

828   *Imaging*, *25*(11), 1451–1461. https://doi.org/10.1109/TMI.2006.880587

829   Dantuma, M., van Dommelen, R., & Manohar, S. (2019). Semi-anthropomorphic

830   photoacoustic breast phantom. *Biomed. Opt. Express*, *10*(11), 5921–5939.

831   https://doi.org/10.1364/BOE.10.005921

832   Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species.

833   *Ecology*, *26*(3), 297–302. https://doi.org/https://doi.org/10.2307/1932409

834   Epah, J., Pálfi, K., Dienst, F. L., Malacarne, P. F., Bremer, R., Salamon, M., Kumar, S., Jo,

835   H., Schürmann, C., & Brandes, R. P. (2018). 3D imaging and quantitative analysis of

836   vascular networks: A comparison of ultramicroscopy and micro-computed tomography.

837   *Theranostics*, *8*(8), 2117–2133. https://doi.org/10.7150/THNO.22610

838   Frangi, A. F., Niessen, W. J., Vincken, K. L., & Viergever, M. A. (1998). Multiscale vessel

839   enhancement filtering. In W. M. Wells, A. Colchester, & S. Delp (Eds.), *Medical Image*

840   *Computing and Computer-Assisted Intervention --- MICCAI'98* (pp. 130–137). Springer

841   Berlin Heidelberg.

842   Galarreta-Valverde, M. A. (2012). *Geração de redes vasculares sintéticas tridimensionais*

843   *utilizando sistemas de Lindenmayer estocásticos e parametrizados* [Instituto de

844   Matemática e Estatística]. https://doi.org/10.11606/D.45.2012.tde-30112012-172822

845   Galarreta-Valverde, M. A., Macedo, M. M. G., Mekkaoui, C., & Jackowski, M. P. (2013).

846   Three-dimensional synthetic blood vessel generation using stochastic L-systems.

847   *Medical Imaging 2013: Image Processing*, 86691I. https://doi.org/10.1117/12.2007532

848   Haedicke, K., Agemy, L., Omar, M., Berezhnoi, A., Roberts, S., Longo-Machado, C., Skubal,

849   M., Nagar, K., Hsu, H.-T., Kim, K., Reiner, T., Coleman, J., Ntziachristos, V., Scherz, A.,

850   & Grimm, J. (2020). High-resolution optoacoustic imaging of tissue responses to

851   vascular-targeted therapies. *Nature Biomedical Engineering*, *4*(3), 286–297.

852   https://doi.org/10.1038/s41551-020-0527-8

853   Hanahan, D., & Weinberg, R. a. (2011). Hallmarks of cancer: the next generation. *Cell*,

854   *144*(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

855   Harrell, F. E. (2016). *Regression modeling strategies*. Springer International Publishing.

856   Hlushchuk, R., Haberthür, D., & Djonov, V. (2019). Ex vivo microangioCT: Advances in

857   microvascular imaging. *Vascular Pharmacology*, *112*, 2–7.

858   https://doi.org/10.1016/J.VPH.2018.09.003

859   Imai, T., Muz, B., Yeh, C.-H., Yao, J., Zhang, R., Azab, A. K., & Wang, L. (2017). Direct

860   measurement of hypoxia in a xenograft multiple myeloma model by optical-resolution

861   photoacoustic microscopy. *Cancer Biology & Therapy*, *18*(2), 101–105.

862   https://doi.org/10.1080/15384047.2016.1276137

863   Jährling, N., Becker, K., & Dodt, H.-U. (2009). 3D-reconstruction of blood vessels by

864   ultramicroscopy. *Organogenesis*, *5*(4), 227–230. https://doi.org/10.4161/org.5.4.10403

865   Janek Gröhl, Kris K. Dreher, Melanie Schellenberg, Alexander Seitel, L. M.-H. (2021).

866   SIMPA: an open source toolkit for simulation and processing of photoacoustic images.

867   *Photons Plus Ultrasound: Imaging and Sensing 2021*.

868   Jia, D., & Zhuang, X. (2021). Learning-based algorithms for vessel tracking: A review.

869   *Computerized Medical Imaging and Graphics*, *89*, 101840.

870   https://doi.org/https://doi.org/10.1016/j.compmedimag.2020.101840

871   Joseph, J., Tomaszewski, M. R., Quiros-Gonzalez, I., Weber, J., Brunker, J., & Bohndiek, S.

872   E. (2017). Evaluation of Precision in Optoacoustic Tomography for Preclinical Imaging

873   in Living Subjects. *Journal of Nuclear Medicine* , *58*(5), 807–814.

874   Kelch, I. D., Bogle, G., Sands, G. B., Phillips, A. R. J., LeGrice, I. J., & Rod Dunbar, P.

875   (2015). Organ-wide 3D-imaging and topological analysis of the continuous

876   microvascular network in a murine lymph node. *Scientific Reports*, *5*(1), 16534.

877   https://doi.org/10.1038/srep16534

878   Keller, P. J., & Dodt, H. U. (2012). Light sheet microscopy of living or cleared specimens. In

879   *Current Opinion in Neurobiology* (Vol. 22, Issue 1, pp. 138–143). Curr Opin Neurobiol.

880   https://doi.org/10.1016/j.conb.2011.08.003

881   Krig, S., & Krig, S. (2014). Ground Truth Data, Content, Metrics, and Analysis. In *Computer*

882   *Vision Metrics* (pp. 283–311). Apress. https://doi.org/10.1007/978-1-4302-5930-5_7

883   Krishna Priya, S., Nagare, R. P., Sneha, V. S., Sidhanth, C., Bindhya, S., Manasa, P., &

884   Ganesan, T. S. (2016). Tumour angiogenesis—Origin of blood vessels. *International*

885   *Journal of Cancer*, *139*(4), 729–735. https://doi.org/10.1002/ijc.30067

886   Lao, Y., Xing, D., Yang, S., & Xiang, L. (2008). Noninvasive photoacoustic imaging of the

887   developing vasculature during early tumor growth. *Physics in Medicine & Biology*,

888   *53*(15), 4203. http://stacks.iop.org/0031-9155/53/i=15/a=013

889   Lindenmayer, A. (1968). Mathematical models for cellular interactions in development I.

890   Filaments with one-sided inputs. *Journal of Theoretical Biology*, *18*(3), 280–299.

891   https://doi.org/10.1016/0022-5193(68)90079-9

892   Michiels, C., Tellier, C., & Feron, O. (2016). Cycling hypoxia: A key feature of the tumor

893   microenvironment. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1866*(1),

894   76–86. https://doi.org/https://doi.org/10.1016/j.bbcan.2016.06.004

895   Moccia, S., De Momi, E., El Hadji, S., & Mattos, L. S. (2018). Blood vessel segmentation

896   algorithms — Review of methods, datasets and evaluation metrics. *Computer Methods*

897   *and Programs in Biomedicine*, *158*, 71–91.

898   https://doi.org/https://doi.org/10.1016/j.cmpb.2018.02.001

899   Nagy, J. A., & Dvorak, H. F. (2012). Heterogeneity of the tumor vasculature: the need for

900   new tumor blood vessel type-specific targets. *Clinical & Experimental Metastasis*, *29*(7),

901   657–662. https://doi.org/10.1007/s10585-012-9500-6

902   Ntziachristos, V. (2010). Going deeper than microscopy: the optical imaging frontier in

903   biology. *Nature Methods*, *7*(8), 603–614. https://doi.org/10.1038/nmeth.1483

904   Ntziachristos, V., Ripoll, J., Wang, L. V, & Weissleder, R. (2005). Looking and listening to

905   light: the evolution of whole-body photonic imaging. *Nature Biotechnology*, *23*, 313.

906   http://dx.doi.org/10.1038/nbt1074

907   Omar, M., Aguirre, J., & Ntziachristos, V. (2019). Optoacoustic mesoscopy for biomedicine.

908   *Nature Biomedical Engineering*, *3*(5), 354–370. https://doi.org/10.1038/s41551-019-

909   0377-4

910   Omar, M., Gateau, J., & Ntziachristos, V. (2013). Raster-scan optoacoustic mesoscopy in

911   the 25–125 MHz range. *Opt. Lett.*, *38*(14), 2472–2474.

912   https://doi.org/10.1364/OL.38.002472

913   Omar, M., Schwarz, M., Soliman, D., Symvoulidis, P., & Ntziachristos, V. (2015). Pushing

914   the Optical Imaging Limits of Cancer with Multi-Frequency-Band Raster-Scan

915   Optoacoustic Mesoscopy (RSOM). *Neoplasia*, *17*(2), 208–214.

916   https://doi.org/https://doi.org/10.1016/j.neo.2014.12.010

917   Omar, M., Soliman, D., Gateau, J., & Ntziachristos, V. (2014). Ultrawideband reflection-

918   mode optoacoustic mesoscopy. *Optics Letters*, *39*(13), 3911–3914.

919   https://doi.org/10.1364/OL.39.003911

920   Orlova, A., Sirotkina, M., Smolina, E., Elagin, V., Kovalchuk, A., Turchin, I., & Subochev, P.

921   (2019). Raster-scan optoacoustic angiography of blood vessel development in colon

922   cancer models. *Photoacoustics*, *13*, 25–32. https://doi.org/10.1016/j.pacs.2018.11.005

923   Oruganti, T., Laufer, J. G., & Treeby, B. E. (2013). Vessel filtering of photoacoustic images.

924   *Proc.SPIE*, *8581*. https://doi.org/10.1117/12.2005988

925   Quiros-Gonzalez, I., Tomaszewski, M. R., Aitken, S. J., Ansel-Bollepalli, L., McDuffus, L.-A.,

926   Gill, M., Hacker, L., Brunker, J., & Bohndiek, S. E. (2018). Optoacoustics delineates

927   murine breast cancer models displaying angiogenesis and vascular mimicry. *British

928   Journal of Cancer*, *118*(8), 1098–1106. https://doi.org/10.1038/s41416-018-0033-x

929   R Core, T. (2021). *R: A language and environment for statistical computing.* R Foundation

930   for Statistical Computing, Vienna, Austria. https://www.r-project.org/

931   Raumonen, P., & Tarvainen, T. (2018). Segmentation of vessel structures from

932   photoacoustic images with reliability assessment. *Biomedical Optics Express*, *9*(7),

933   2887–2904. https://doi.org/10.1364/BOE.9.002887

934   Rebling, J., Greenwald, M. B.-Y., Wietecha, M., Werner, S., & Razansky, D. (2021). Long-

935   Term Imaging of Wound Angiogenesis with Large Scale Optoacoustic Microscopy.

936   *Advanced Science*, *8*(13), 2004226. https://doi.org/10.1002/ADVS.202004226

937   Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for

938   Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted

939   Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science*, *9351*,

940   234–241. https://doi.org/10.1007/978-3-319-24574-4

941   Rozenberg, G., & Arto Salomaa (Eds.). (1992). *Lindenmayer Systems: Impacts on

942   theoretical computer science, computer graphics, and development biology*. Springer-

943   Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-58117-5

944   Sato, Y., Nakajima, S., Shiraga, N., Atsumi, H., Yoshida, S., Koller, T., Gerig, G., & Kikinis,

945   R. (1998). Three-dimensional multi-scale line filter for segmentation and visualization of

946   curvilinear structures in medical images. *Medical Image Analysis*, *2*(2), 143–168.

947   https://doi.org/https://doi.org/10.1016/S1361-8415(98)80009-1

948   Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,

949   Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J.,

950   Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source

951   platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682.

952   https://doi.org/10.1038/nmeth.2019

953   Soetikno, B., Hu, S., Gonzales, E., Zhong, Q., Maslov, K., Lee, J.-M., & Wang, L. V. (2012).

954   Vessel segmentation analysis of ischemic stroke images acquired with photoacoustic

955   microscopy. *Proc.SPIE*, *8223*. https://doi.org/10.1117/12.911089

956   Sommer, C., Straehle, C., Kothe, U., & Hamprecht, F. A. (2011). Ilastik: Interactive learning

957   and segmentation toolkit. *Eighth IEEE International Symposium on Biomedical Imaging*,

958   230–233. https://doi.org/10.1109/ISBI.2011.5872394

959   Sternberg, S. R. (1983). Biomedical Image Processing. *Computer*, *16*(1), 22–34.

960   https://doi.org/10.1109/MC.1983.1654163

961   Stolz, B. J., Kaeppler, J., Markelc, B., Mech, F., Lipsmeier, F., Muschel, R. J., Byrne, H. M.,

962   & Harrington, H. A. (2020). Multiscale topology characterises dynamic tumour vascular

963   networks. In *arXiv*. https://doi.org/2008.08667

964   Trédan, O., Galmarini, C. M., Patel, K., & Tannock, I. F. (2007). Drug Resistance and the

965   Solid Tumor Microenvironment. *JNCI: Journal of the National Cancer Institute*, *99*(19),

966   1441–1454. https://doi.org/10.1093/jnci/djm135

967   Treeby, B. E., & Cox, B. T. (2010). k-Wave: MATLAB toolbox for the simulation and

968   reconstruction of photoacoustic wave fields. *Journal of Biomedical Optics*, *15*(2), 1–12.

969   https://doi.org/10.1117/1.3360308

970   Tsai, W.-H. (1985). Moment-preserving thresolding: A new approach. *Computer Vision,*

971   *Graphics, and Image Processing*, *29*(3), 377–393.

972   https://doi.org/https://doi.org/10.1016/0734-189X(85)90133-1

973   Ul Haq, I., Nagaoka, R., Makino, T., Tabata, T., & Saijo, Y. (2016). 3D Gabor wavelet based

974   vessel filtering of photoacoustic images. In *Conference proceedings: ... Annual*

975   *International Conference of the IEEE Engineering in Medicine and Biology Society.*

976   *IEEE Engineering in Medicine and Biology Society. Conference* (Vol. 2016).

977   https://doi.org/10.1109/EMBC.2016.7591576

978   Wang, L. V, & Yao, J. (2016). A practical guide to photoacoustic tomography in the life

979   sciences. *Nature Methods*, *13*(8), 627–638. https://doi.org/10.1038/nmeth.3925

980   Zhao, F., Chen, Y., Hou, Y., & He, X. (2019). Segmentation of blood vessels using rule-

981   based and machine-learning-based methods: a review. *Multimedia Systems*, *25*(2),

982   109–118. https://doi.org/10.1007/s00530-017-0580-7

983   Zhao, H., Liu, C., Li, K., Chen, N., Zhang, K., Wang, L., Lin, R., Gong, X., Song, L., & Liu, Z.

984   (2019). Multiscale vascular enhancement filter applied to in vivo morphologic and

985   functional photoacoustic imaging of rat ocular vasculature. *IEEE Photonics Journal*,

986   *11*(6). https://doi.org/10.1109/JPHOT.2019.2948955

987

988

989 **Supplementary information for:**

990 Quantification of vascular networks in photoacoustic mesoscopy

991

992 **Authors:**

993 Emma L. Brown[1,2,#], Thierry L. Lefebvre[1,2,#], Paul W. Sweeney[1,2,#], Bernadette J. Stolz[3], Janek

994 Gröhl[1,2], Lina Hacker[1,2], Ziqiang Huang[2] [+], Dominique-Laurent Couturier[2], Heather A.

995 Harrington[3], Helen M. Byrne[3] and Sarah E. Bohndiek[1,2] *.

996

997 [1] Department of Physics, University of Cambridge, JJ Thomson Avenue, Cambridge, CB3

998 0HE, U.K.

999

1000 [2] Cancer Research UK Cambridge Institute, University of Cambridge, Robinson Way,

1001 Cambridge, CB2 0RE, U.K.

1002

1003 [3] Mathematical Institute, University of Oxford, Woodstock Road, Oxford, OX2 6GG, U.K.

1004

1005 + Now at: European Molecular Biology Laboratory Imaging Centre, Heidelberg, Germany.

1006

1007 # These authors contributed equally

1008 * Corresponding author: email seb53@cam.ac.uk, telephone +44 1223 337267.

1009

1010 **Keywords:** photoacoustic imaging, vasculature, segmentation, topology

1011

1012

1013

1014

Supplementary Materials and Methods

1. String phantom preparation

The string phantom used in this study was prepared by mixing 1.5 g agarose (Fluka Analytical, 05039-500G) in 97.3 mL deionised water in a glass media bottle and heated in a microwave until the solution turned clear. After cooling down the solution to 60°C, 2.08 mL of pre-warmed intralipid was added to generate a reduced scattering coefficient of 5.0 cm^{-1} according to a previously characterised recipe(Joseph et al., 2017). The mixture was poured into a 3D-printed phantom mould, which was designed in Autodesk Fusion 360 (San Rafael, CA, USA) and printed using an Anet A6 Printer with polylactic acid (PLA PRO 1.75mm Fluorescent Yellow PLA 3D Printer Filament, 832-0254, RS Components, UK) as a base material. **Supplementary Figure 2** shows the phantom mould with and without agar.

2. 3D CNN for ROI delineation

2.1. Preparation of training data

Image volumes consist of a series of 8-bit grayscale *Tiffs* (no compression) of 600 x 600 pixels in the XY-plane and a stack of 700 images in Z, with anisotropic voxels of size 20 x 20 x 4 μm^3. Our dataset has a total of 166 PAI volumes, each paired with a corresponding binary semi-manually-annotated volume, where a voxel value of 0 and 255 indicates the background or tumour ROIs, respectively. The annotated volumes were generated by an experienced user, who first identified the top and bottom image containing the tumour in Z. Within these upper and lower bounds, ROIs were manually drawn in the XY plane on approximately 4 image slices. Bound by these data, a convex hull was extrapolated to approximate the ROI in the remaining image slices.

Prior to training, image volumes and binary masks were downsampled to an isotropic volume of 256 x 256 x 256 voxels to fit into computer memory. Data was locally standardised and normalised to a pixel range between 0 and 1 and the volumes randomly partitioned into training, validation, and testing subsets. Here, ~5% of images were allocated for testing, with

1041   the remaining portion split 80:20 for training and validation respectively (8 / 126 / 32 image

1042   volumes, respectively).

1043   *2.2 Neural Network Architecture for ROI delineation*

1044   The 3D CNN is based on the U-Net architecture(Ronneberger et al., 2015) extended

1045   for volumetric delineation(Çiçek et al., 2016). The structure consists of an encoder, which

1046   extracts spatial features from a 3D image volume, and a decoder, which constructs a

1047   segmentation map from these features (**Supplementary Figure 10**). The network architecture

1048   consists of five convolutional layers. The encoder path contains two 3 x 3 x 3 convolutions

1049   followed by a rectified linear unit (ReLU) activation for faster convergence and accuracy(Çiçek

1050   et al., 2016). Each ReLU activation is followed by 2 x 2 x 2 max pooling with strides of two in

1051   each dimension. For the 3rd, 4th and 5th layers, dropout is applied to reduce segmentation bias

1052   and ensure segmentation is performed utilising high-level features that may not have been

1053   considered in our semi-manual ROI annotations.

1054   The decoder path consists of two 3 x 3 x 3 deconvolutions of strides of 2 in each

1055   dimension, followed by 3 x 3 x 3 convolutions, batch normalisation and ReLU activation. High-

1056   resolution features were provided via shortcut connections from the same layer in the encoder

1057   path. The final layer applied an additional 1 x 1 x 1 convolution followed by sigmoid activation

1058   to ensure the correct number of output channels and range of pixel values [0, 1]. The input

1059   layer is designed to take *n* grayscale (one channel) tumour volumes as input with a pre-defined

1060   volume (128 x 128 x 128 voxels in X, Y, Z-direction used here). The U-Net binary mask

1061   prediction contains an equal number of voxels as the input. The CNN was implemented in

1062   Keras(Chollet & Others, 2015) with the Tensorflow framework(Abadi et al., 2015). The model

1063   was trained and tested on a Dell Precision 7920 with a Dual Intel Xeon Gold 5120 CPU with

1064   128 GB RAM and a NVIDIA Quadro GV100 32 GB GPU.

1065   *2.3. Hyperparameter Optimisation*

1066   Hyperparameters were optimised and evaluated using Talos(*Autonomio Talos*, 2019),

1067   a fully-automated hyperparameter tuner for Keras. A random search optimisation strategy was

1068   deployed using the quantum random method. Here, a probabilistic reduction scheme was

1069   used to reduce the number of parameter permutations by removing poorly performing

1070   hyperparameter configurations from the remaining search space after a predefined interval.

1071   The number of filters used ranged from 16 in the 1st layer to 512 in the 5th. Dropout at a rate

1072   of 0.2 was applied in the 3rd, 4th and 5th layers. A Glorot uniform initialiser was used for all

1073   convolution and deconvolution layers. The model was trained using an Adam optimiser with

1074   learning and decay rates of 10^{-5} and 10^{-8}, respectively, and the dice coefficient (F1)(Crum et

1075   al., 2006) used as the loss function.

1076   *2.4. U-Net Training & Predictions*

1077   Training was performed with a batch size of 3 image volumes for a total of 120 epochs

1078   (**Supplementary Figure 11A**). The fully-trained network achieved an accuracy of 88.3% and

1079   87.3% on the training and validation sets respectively (**Supplementary Figure 11B**).

1080   Following training and test, we applied the CNN to the entire set of volumes to compare

1081   predictions of ROI volume to the ground truth (**Supplementary Figure 11C**). Blood volumes

1082   were then calculated within the predicted ROIs using the AT method and compared against

1083   the user annotations (**Supplementary Figure 11D**). We found a significant correlation

1084   between user annotated and predicted data for both ROI volume (Spearman's rank

1085   correlation: $r = 0.821$, $p < 0.0001$) and blood volume ($r = 0.958$, $p < 0.0001$), indicating our

1086   CNN achieves sufficient performance against the experienced user to be applied for extracting

1087   tumours prior to testing the segmentation pipeline.

1088   *3. Signal-to-noise ratio characterisation*

1089   PAI quality pre-segmentation was quantified by measuring signal-to-noise ratio (SNR), defined

1090   as the mean of signal over the standard deviation of the background signal. For *in silico* and

1091   in phantom ground truth datasets, the mean of the signal was taken within the binary ground

1092   truth masks of the images and reported for different depths.

1093

1094

1095

1096 **Supplementary Tables**

1097 **Supplementary Table 1:** Descriptions of our statistical and topological descriptors.

Descriptor	Description
Connected Components, β_0	Number of 0-dimensional topological features, *i.e.* the number of subgraphs or clusters (vascular subnetworks). Values are normalised with respect to the total number of edges per segmented image volume.
Loops, β_1	Number of 1-dimensional topological features, *i.e.* the numbers of looping structures in vascular graph. Values are normalised with respect to the total number of edges per segmented image volume.
Sum-of-angles measure (SOAM)	The sum of angles between tangents to the curve taken at regular intervals normalised against vessel length, *i.e.* the average change in angle per unit length.
Chord-to-length ratio (CLR)	The ratio between the Euclidean distance connecting the two ends of a blood vessel and the length of the blood vessel, *e.g.* a straight vessel has a CLR equal to 1.

1098

1099 **Supplementary Table 2:** Training and testing dataset split for random forest-based

1100 segmentation in ilastik.

Data	Ground truth labels	Training	Testing
In silico	Original binary labels of L-net branches and surrounding background	30 L-nets	30 L-nets (data in **Figures 2** and **3**)
In vitro	Manual labelling of all XY slices containing strings and of surrounding background	2 string phantom scans	5 string phantom scans (data in **Figures 4** and **5**)
In vivo	Manual labelling of 10 XY slices per image at distributed depths and of surrounding background	20 PDX tumour scans	14 PDX tumour scans (data in **Figures 6-9**)

1101

1102 **Supplementary Table 3.** Mean computation time in seconds for each segmentation method

1103 on *in silico, in vitro,* and *in vivo* data. Note: Segmentations were performed on a dual Intel

1104 Xeon E5-2623 v4 2.60 GHz quad-core processor and 64.0 GB of RAM.

Data	AT	AT+VF	RF	RF+VF
In silico	7.4	198.1	191.1	381.8
In vitro	12.9	219.6	1280.0	1486.7
In vivo	38.4	1215.5	1500.7	2677.8

1105

1106

1107 **Supplementary Table 4.** Absolute number of connected components for each L-Net skeleton

1108 generated from the ground truth and each segmentation method. Network names are

1109 organised based on number of recursive L-Net iterations and index, for example, 'LNet_i4_0'

1110 is the zeroth network of those with 4 iterations. Note, the number of known branching points

1111 is equal to number of iterations minus 3.

Name	Ground Truth	AT	AT+VF	RF	RF+VF
LNet_i4_0	1	17	4	5	3
LNet_i4_1	1	12	1	5	4
LNet_i4_2	1	8	3	3	2
LNet_i4_3	1	23	3	24	111
LNet_i4_4	1	15	1	2	1
LNet_i6_0	1	12	3	10	16
LNet_i6_1	1	21	2	4	6
LNet_i6_2	1	13	3	4	8
LNet_i6_3	1	1	2	3	2
LNet_i6_4	1	21	9	8	4
LNet_i8_0	1	20	9	16	16
LNet_i8_1	1	26	5	16	9
LNet_i8_2	1	12	9	12	5
LNet_i8_3	1	12	7	13	6
LNet_i8_4	1	7	2	14	11
LNet_i10_0	1	30	14	29	21
LNet_i10_1	1	30	12	24	19
LNet_i10_2	1	18	9	24	14
LNet_i10_3	1	40	20	33	34
LNet_i10_4	1	21	21	28	27
LNet_i12_0	1	76	16	49	43
LNet_i12_1	1	68	23	52	52
LNet_i12_2	1	58	24	40	37

LNet_i12_3	1	49	19	83	72
LNet_i12_4	1	58	26	68	53
LNet_i14_0	1	88	39	155	103
LNet_i14_1	1	81	45	112	100
LNet_i14_2	1	69	36	93	79
LNet_i14_3	1	91	46	87	89
LNet_i14_4	1	104	34	116	74

1112

1113

1114

1115

1116

1117

1118

1119

1120

1121

1122

1123

1124

1125

1126

1127

1128 **Supplementary Table 5.** Absolute number of loops for each L-Net skeleton generated from

1129 the ground truth and each segmentation method. Network names are organised based on

1130 number of recursive L-Net iterations and index, for example, 'LNet_i4_0' is the zeroth network

1131 of those with 4 iterations. Note, the number of known branching points is equal to number of

1132 iterations minus 3.

Name	Ground Truth	AT	AT+VF	RF	RF+VF
LNet_i4_0	0	46	0	2	2
LNet_i4_1	0	27	0	4	16
LNet_i4_2	0	42	0	0	8
LNet_i4_3	1	45	13	67	86
LNet_i4_4	0	41	0	1	11
LNet_i6_0	0	54	1	12	27
LNet_i6_1	1	28	0	13	11
LNet_i6_2	1	63	7	22	72
LNet_i6_3	0	6	0	0	0
LNet_i6_4	2	47	0	9	8
LNet_i8_0	4	68	1	33	22
LNet_i8_1	2	21	0	4	1
LNet_i8_2	2	53	0	2	11
LNet_i8_3	1	86	0	8	13
LNet_i8_4	1	40	0	1	0
LNet_i10_0	4	0	0	0	0
LNet_i10_1	20	14	0	0	1
LNet_i10_2	9	20	0	0	0
LNet_i10_3	9	33	0	14	12
LNet_i10_4	11	7	0	1	1
LNet_i12_0	73	9	4	24	12
LNet_i12_1	123	37	5	25	25
LNet_i12_2	106	32	7	35	36

54

LNet_i12_3	30	4	1	1	1
LNet_i12_4	62	2	1	8	13
LNet_i14_0	353	16	9	18	13
LNet_i14_1	426	43	15	58	43
LNet_i14_2	395	19	9	19	20
LNet_i14_3	376	74	15	58	51
LNet_i14_4	304	29	12	20	29

1133

1134

1135 **Supplementary Table 6.** The number of edges for each L-Net skeleton generated from the

1136 ground truth and each segmentation method. Network names are organised based on number

1137 of recursive L-Net iterations and index, for example, 'LNet_i4_0' is the zeroth network of those

1138 with 4 iterations. Note, the number of known branching points is equal to number of iterations

1139 minus 3.

1140

Name	Ground Truth	AT	AT+VF	RF	RF+VF
LNet_i4_0	3	160	4	19	10
LNet_i4_1	3	122	3	24	59
LNet_i4_2	3	144	3	5	32
LNet_i4_3	7	177	49	247	420
LNet_i4_4	3	136	3	9	38
LNet_i6_0	15	217	18	63	122
LNet_i6_1	18	134	12	55	60
LNet_i6_2	19	222	34	94	236
LNet_i6_3	15	37	14	15	14
LNet_i6_4	21	167	11	50	43
LNet_i8_0	65	267	34	172	130
LNet_i8_1	65	139	19	77	57
LNet_i8_2	69	227	35	68	85
LNet_i8_3	62	315	33	85	92
LNet_i8_4	62	148	18	44	37
LNet_i10_0	238	114	46	113	101
LNet_i10_1	251	161	54	116	111
LNet_i10_2	196	164	43	100	86
LNet_i10_3	179	241	64	197	183
LNet_i10_4	253	138	69	151	138
LNet_i12_0	687	211	97	332	282
LNet_i12_1	669	283	103	296	296

LNet_i12_2	698	310	130	388	349
LNet_i12_3	794	188	84	182	197
LNet_i12_4	662	150	79	249	246
LNet_i14_0	2359	317	175	524	396
LNet_i14_1	2226	463	262	642	591
LNet_i14_2	2037	261	158	461	427
LNet_i14_3	1707	459	174	506	461
LNet_i14_4	1936	345	153	462	460

1141

1142

1143 **Supplementary Table 7.** The number of nodes for each L-Net skeleton generated from the

1144 ground truth and each segmentation method. Network names are organised based on number

1145 of recursive L-Net iterations and index, for example, 'LNet_i4_0' is the zeroth network of those

1146 with 4 iterations. Note, the number of known branching points is equal to number of iterations

1147 minus 3.

Name	Ground Truth	AT	AT+VF	RF	RF+VF
LNet_i4_0	4	131	8	22	11
LNet_i4_1	4	107	4	25	47
LNet_i4_2	4	110	6	8	26
LNet_i4_3	7	155	39	204	445
LNet_i4_4	4	110	4	10	28
LNet_i6_0	16	175	20	61	111
LNet_i6_1	18	127	14	46	55
LNet_i6_2	19	172	30	76	172
LNet_i6_3	16	32	16	18	16
LNet_i6_4	20	141	20	49	39
LNet_i8_0	62	219	42	155	124
LNet_i8_1	64	144	24	89	65
LNet_i8_2	68	186	44	78	79
LNet_i8_3	62	241	40	90	85
LNet_i8_4	62	115	20	57	48
LNet_i10_0	235	144	60	142	122
LNet_i10_1	232	177	66	140	129
LNet_i10_2	188	162	52	124	100
LNet_i10_3	171	248	84	216	205
LNet_i10_4	243	152	90	178	164
LNet_i12_0	615	278	109	357	313
LNet_i12_1	547	314	121	323	323
LNet_i12_2	593	336	147	393	350

58

LNet_i12_3	765	233	102	264	268
LNet_i12_4	601	206	104	309	286
LNet_i14_0	2007	389	205	661	486
LNet_i14_1	1801	501	292	696	648
LNet_i14_2	1643	311	185	535	486
LNet_i14_3	1332	476	205	535	499
LNet_i14_4	1633	420	175	558	505

1148

1149

1150   **Supplementary Figures**

1151
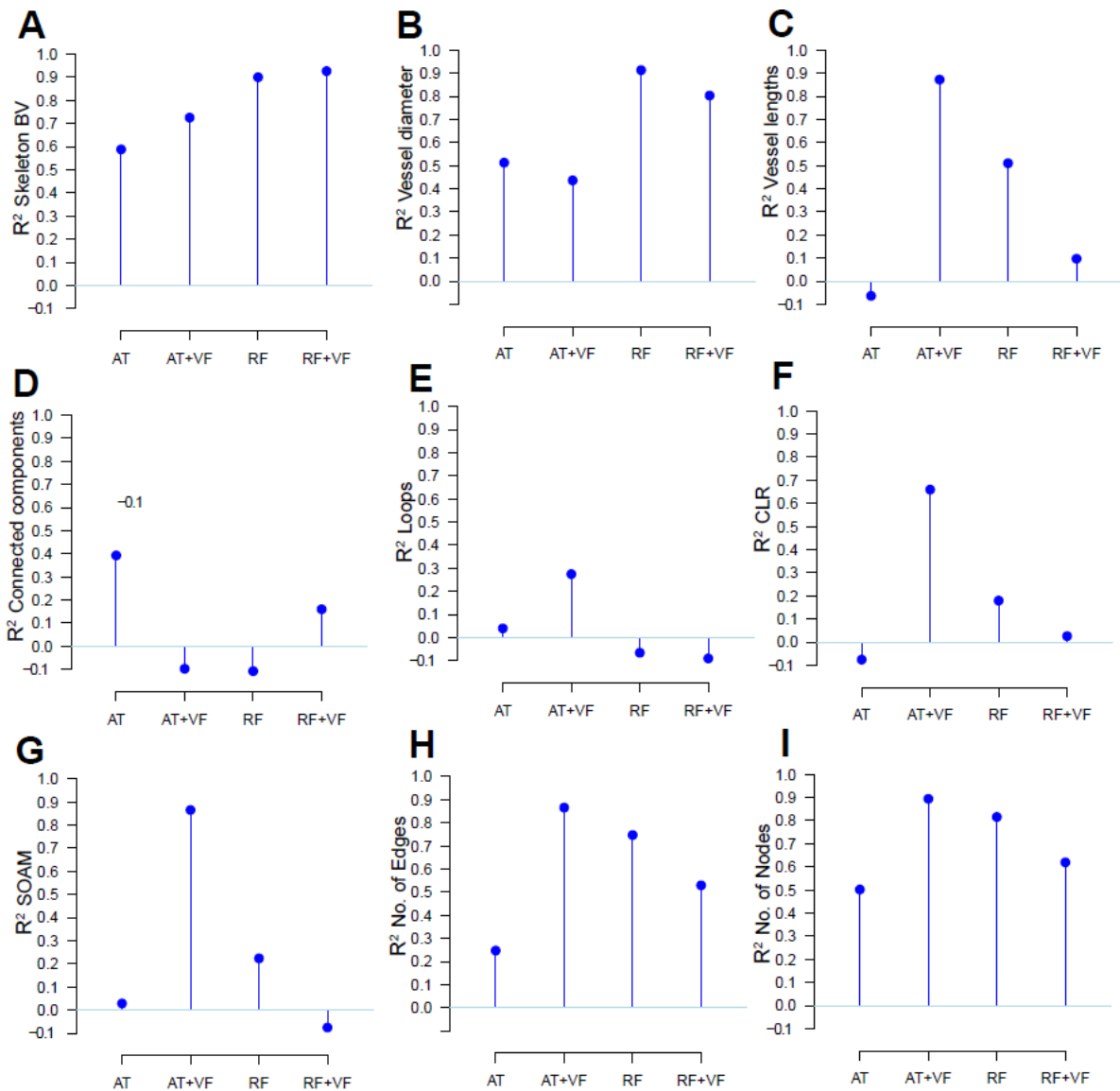
1152

1153   **Supplementary Figure 1. Random forest classifier segments PAI networks with high**

1154   **accuracy while autothresholding with vesselness filtering preserves network structure.**

1155   Bar plot for R^2 values calculated to compare the strength of relationship between the

1156   segmented networks (AT, AT+VF, RF or RF+VF) and ground-truth L-nets for the following

1157   metrics: (A) Normalised skeleton blood volume (BV), (B) Vessel diameters, μm, (C) Vessel

1158   lengths, μm, (D) Connected components, (E) Loops, (F) chord-to-length ratio (CLR), (G) sum-

1159   of-angle measure (SOAM), (H) Number of Edges and (I) Number of Nodes.

1160
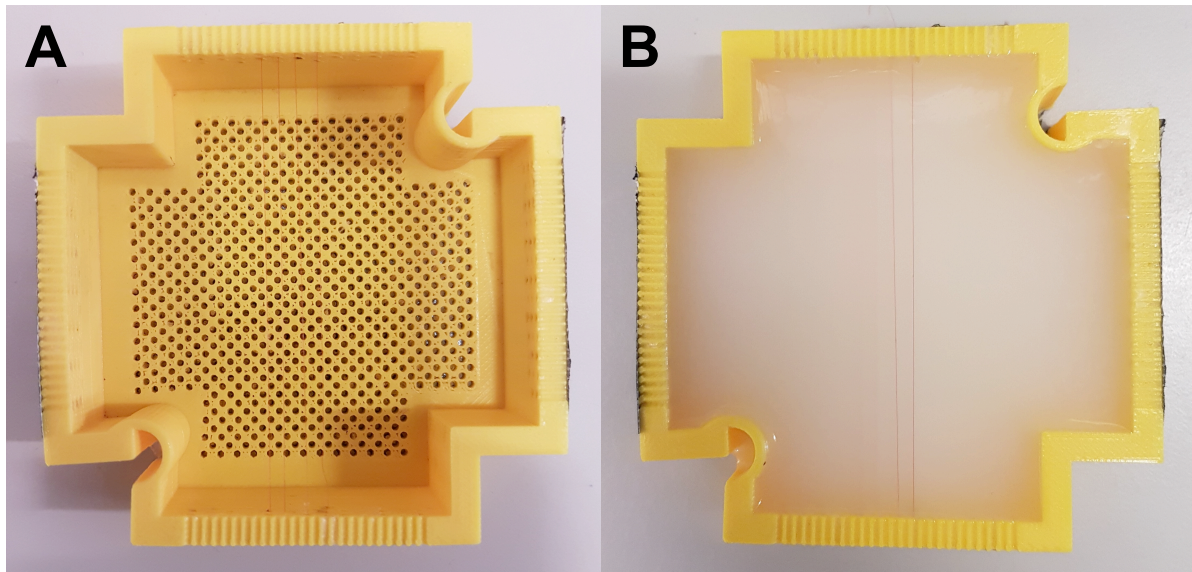
1161

1162 **Supplementary Figure 2. Photographs of the string phantom.** (A) 3D-printed mould (7.4 x

1163 7.4 cm, wall thickness: 4 mm) with the embedded strings and (B) with the agar gel. The top

1164 string was positioned at 0.5 mm from the agar surface, the middle one at 1 mm, and the bottom

1165 one at 2 mm depth.

1166
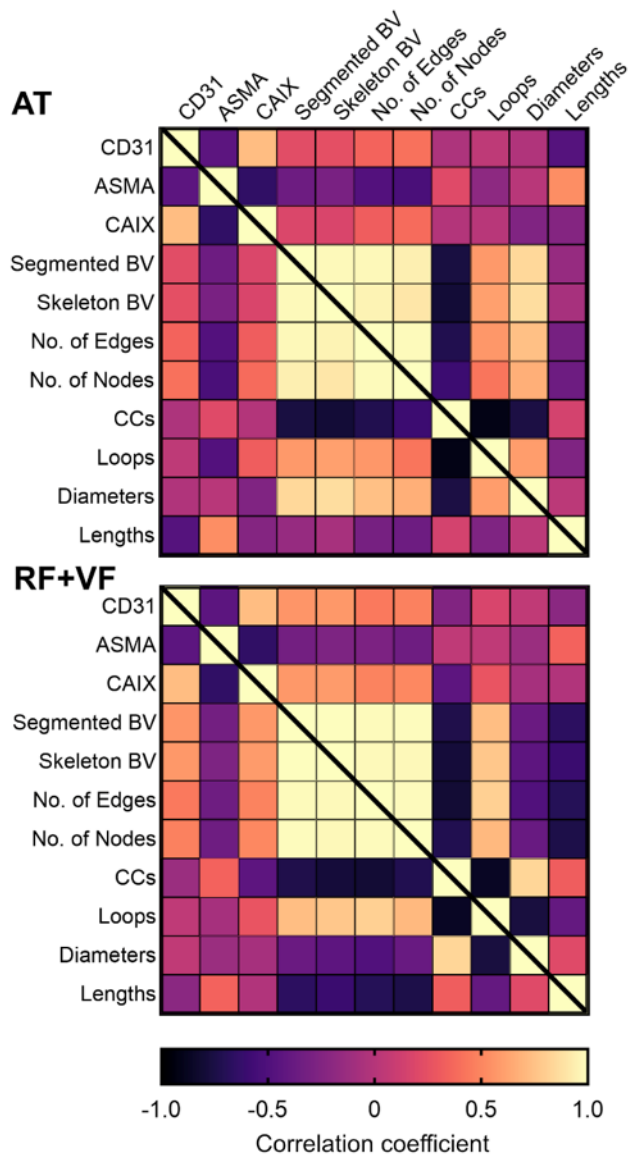
1167

Supplementary Figure 3. Correlation between blood volume and statistical and topological *in vivo* metrics with *ex vivo* IHC in AT and RF+VF segmented networks. Matrix of correlation coefficients for AT (top) and RF+VF (bottom) segmented networks. Pearson or spearman coefficients are used as appropriate, depending on data distribution. Note that none of the coefficients are significant for AT networks (p>0.05). For RF+VF, CD31 staining area and CAIX significantly correlated with segmented (p=0.04 and p=0.03 respectively) and skeletonised blood volume (p=0.03 for both).
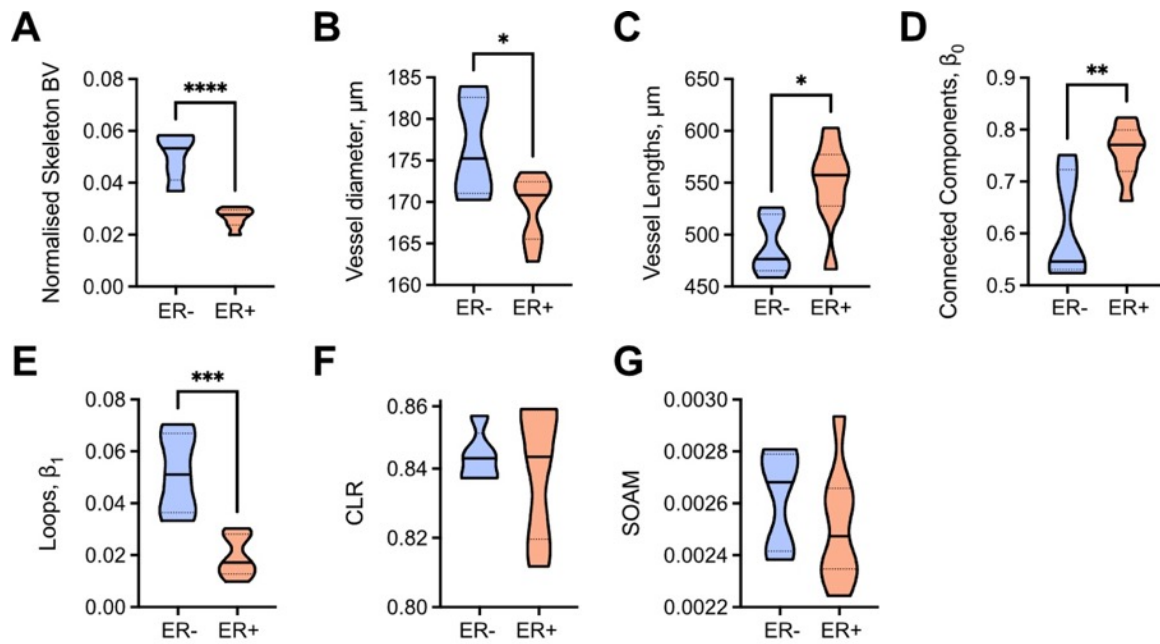
1175

Supplementary Figure 4. Statistical and topological analyses of AT+VF segmentation masks comparing ER- and ER+ tumours. (A-G) Abbreviations defined: blood volume (BV), chord-to-length ratio (CLR), sum-of-angle measure (SOAM). Data are represented by truncated violin plots with interquartile range (dotted black) and median (solid black). Comparisons between ER- and ER+ tumours made with unpaired t-test. *= $p<0.05$, **=$p<0.01$, ***=$p<0.001$, ****=$p<0.0001$.
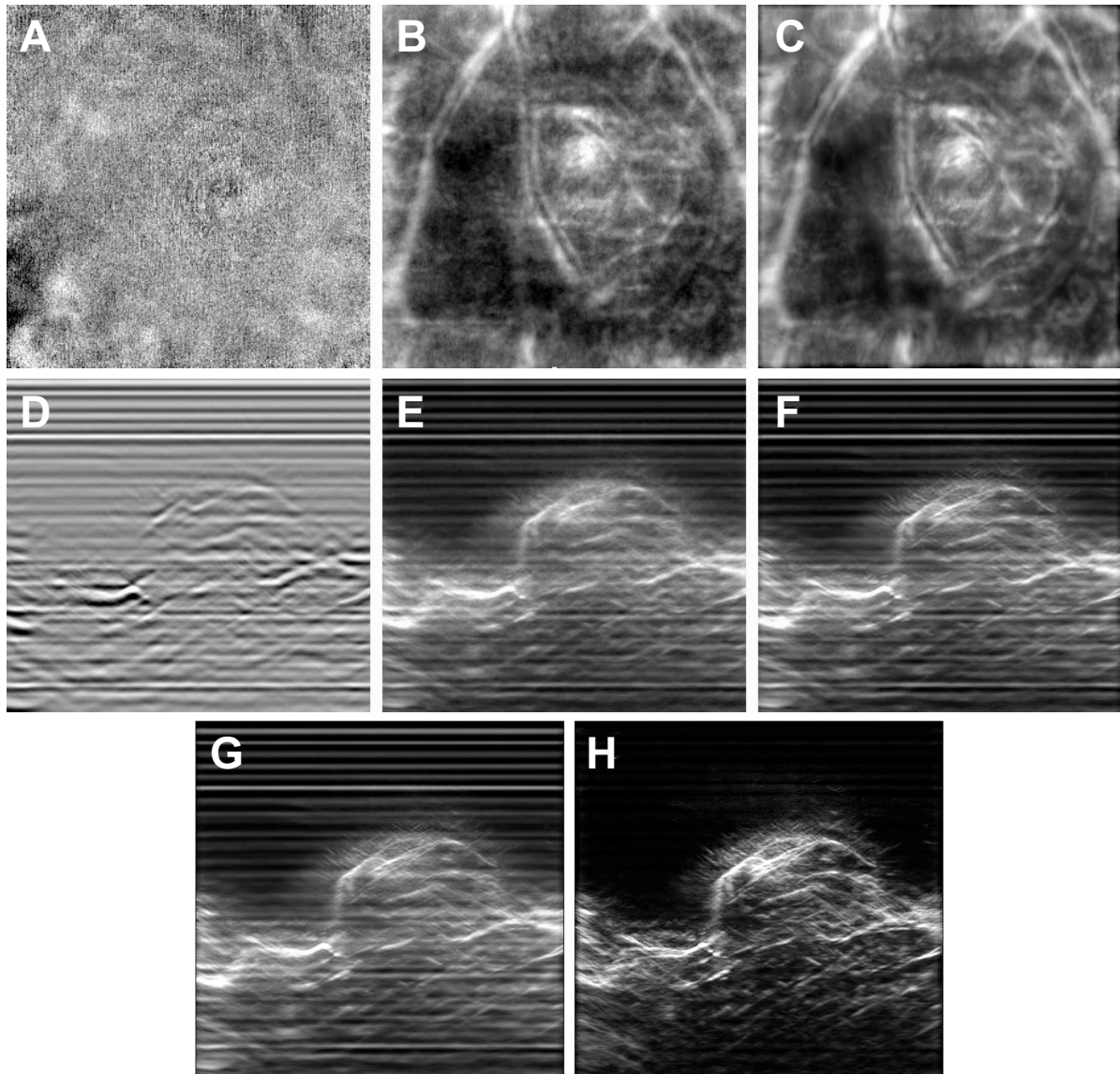
1182

1183
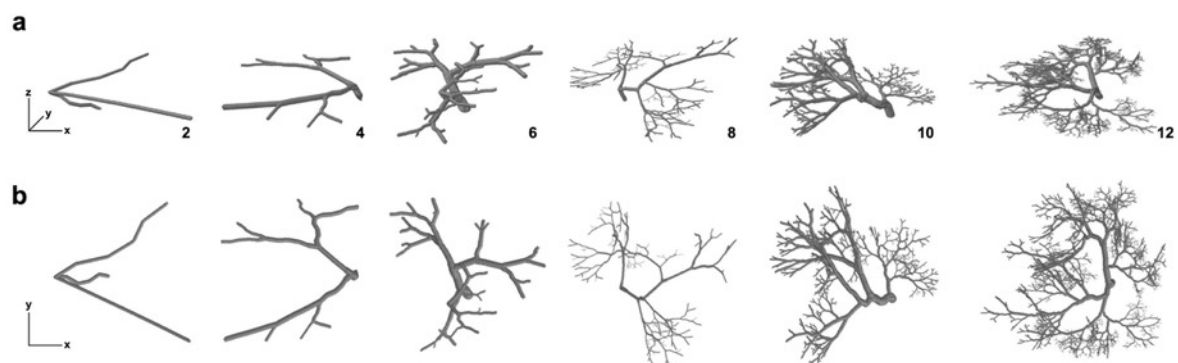
1184

Supplementary Figure 5. Generation of Lindenmayer System (L-System) vascular networks. (A) Segmented views of L-System vasculatures for an increasing number of branching generations (left to right; number of generations indicated). (B) Projected view in the (X,Y) plane of the architectures shown in (A).

1189

1190

Supplementary Figure 6. RSOM data pre-processing in MATLAB. Mean Intensity Projection 2D view of an example RSOM tumour dataset along Z axis (A-C) and Y axis (D-F) axis. From left to right: raw data (A,D), high-pass filtered data (B,E), Wiener filtered data (C,F). The images are processed sequentially through this pipeline, using high-pass filtering to remove echo noises and low-pass adaptive Wiener filtering to further remove stochastic noise in the datasets. (G) Image after MATLAB pre-processing. (H) Image after background correction with rolling ball subtraction in Fiji. The periodical horizontal line artefacts are mostly removed after background correction. All images are 6 x 6 mm.

1199
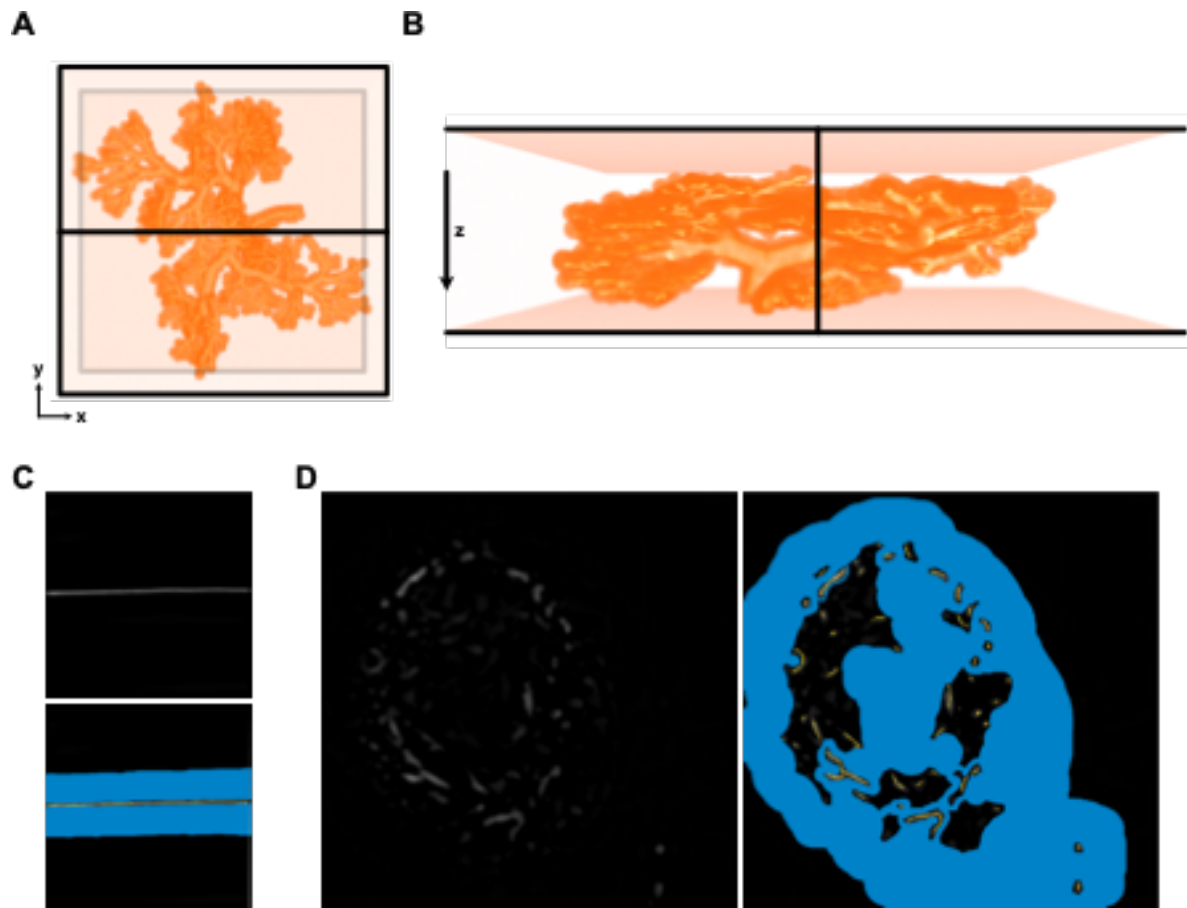
1200

Supplementary Figure 7. Labelling of photoacoustic data for random forest classifier training with ilastik. (A,B) Labels for the full vascular architecture of a given L-net were used for training of ilastik. The region of the L-net within 10 voxels of the vessels was labelled as background (dark orange) in addition to a three voxel thick planes (shown in black). The first was 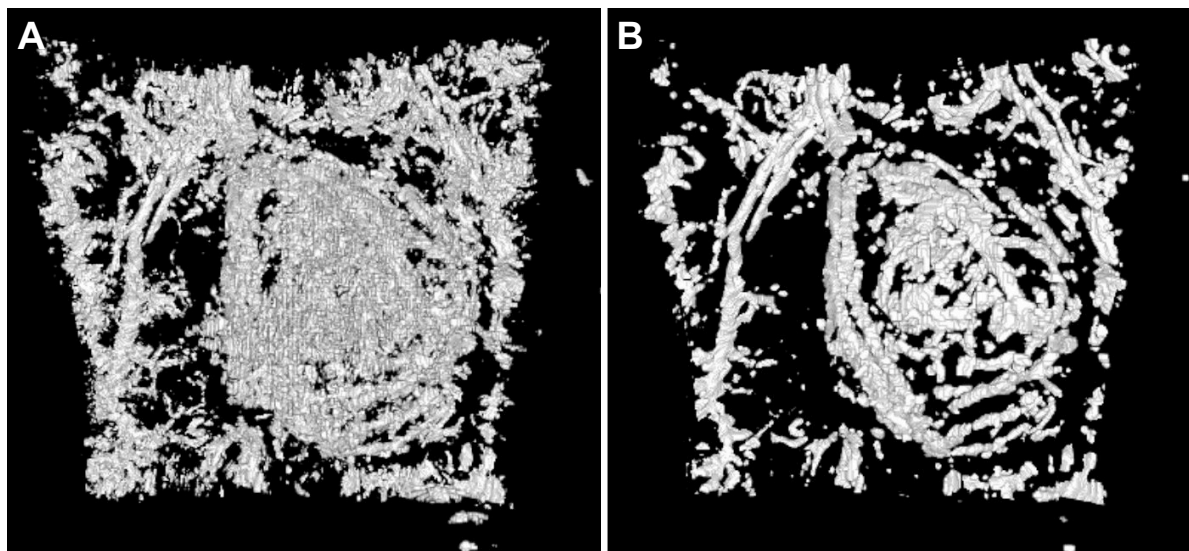located parallel to the z-axis, with the remaining two perpendicular at the top and bottom of the image volume. (C) Labelling of string volumes and (D) of PDX tumour vessels for ilastik training. For (C) and (D) background was labelled as blue and vessels labelled as yellow on 2D slices throughout the 3D volume stack. All images are 6 x 6 mm.
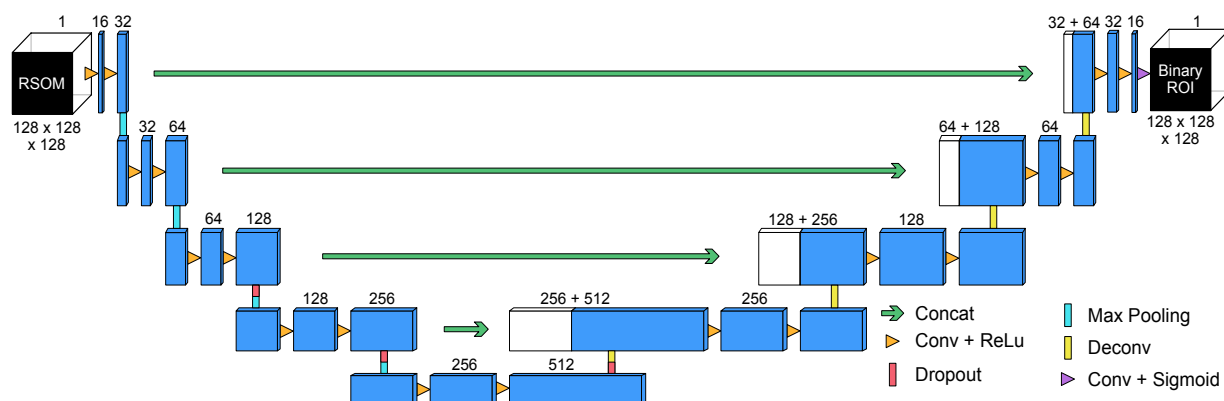
1210

1211 **Supplementary Figure 8. Median filtering of segmented RSOM images.** A 3D rendering

1212 of the exemplar RSOM dataset (6 x 6 x 2.5 mm in X, Y and Z dimensions) used in

1213 Supplementary Figure 6 is shown. (A) Autothresholded dataset. (B) Autothresholded dataset

1214 after 3D Median filtering, to remove impulse noise.

1215

1216

1217

1218

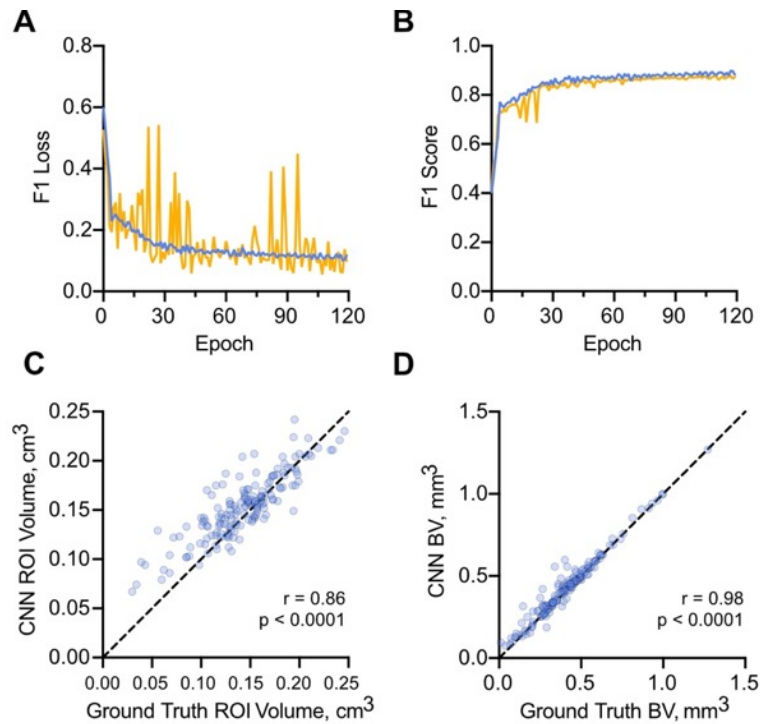1219 **Supplementary Figure 9**. **3D U-Net architecture.** The blue boxes indicate feature maps with

1220 the number of channels denoted above. The input and output image volumes consist of 128

1221 x 128 x 128 voxels. Concat = concatenation, Conv = convolution, ReLu = rectified linear unit,

1222 Deconv = deconvolution.

1223

1224

1225

1226 **Supplementary Figure 10. U-Net training metrics and predictions from the fully-trained**

1227 **architecture.** Training metrics: (A) F1 loss and (B) F1 score for the training (blue) and

1228 validation (orange) datasets. (C) Region-of-interest volumes calculated from the ground truth

1229 (GT) versus the U-Net mask. (D) Computed blood volumes using the ground truth and U-Net

1230 ROI estimations from (C). Note, the lines in (C) and (D) indicate a 1-to-1 relationship, and

1231 blood volumes in (B) were calculated using our auto-thresholding segmentation method.

1232