A deep learning classification task for accurate brain navigation during functional ultrasound imaging

Authorship
Théo Lambert[1,2,3,4], Clément Brunner[1,2,3,4], Dries Kil[1,2,3,4], Roel Wuyts[3], Ellie D'Hondt[3] & Alan Urban[1,2,3,4,*]

Affiliations
[1] Neuro-Electronics Research Flanders, Leuven, Belgium
[2] VIB, Leuven, Belgium
[3] Imec, Leuven, Belgium
[4] Department of Neuroscience, Faculty of Medicine, KU Leuven, Leuven, Belgium

* Corresponding author.
E-mail address: alan.urban@nerf.be

Abstract
Functional ultrasound imaging is a breakthrough technology for imaging brain activity at high spatiotemporal resolution. As it monitors hemodynamic activity, the resulting images are a non-standard representation of local anatomy. This leads to difficulties when determining the exact anatomical location of the recorded image, which is necessary for correctly interpreting the data. Here we propose a convolutional neural network-based framework for accurately navigating the brain during functional ultrasound imaging solely based on vascular landmarks. Our approach uses an image classification task to identify a suitable set of reference positions, from which the anatomical position of an image can be inferred with a precision of 102 ± 98 µm. Further analysis revealed that the predictions are driven by deep brain areas. The robustness of our approach was validated using an ischemic stroke model. It confirms that functional ultrasound imaging information is sufficient for positioning even when local blood flow is disrupted, as observed in many brain pathologies.

Introduction

The brain is often identified as the most complex organ of the human body. It consists of very intricate webs of interconnected neurons and is responsible for every thought, action, memory and feeling that we experience. To support their activity, neurons require a large amount oxygen and nutrients, which is dynamically provided by a complex vascular system[1–3].

As neurons do not have internal reserves of energy, their firing creates an energy demand, which causes a local increase in blood flow[4,5]. This mechanism linking neuronal firing to a local increase in cerebral blood flow is also known as neurovascular coupling[6,7]. Monitoring these local hemodynamic changes forms the basis of several functional brain imaging techniques[8]. The leading technology is blood-oxygen-level-dependent functional magnetic resonance imaging (BOLD-fMRI)[9], which is widely used in clinical studies[10]. However, in a preclinical context, the spatiotemporal resolution of this technique is limited and the required infrastructure make it expensive while implementing complex behavioral paradigms is challenging[8,11].

Overcoming these limitations, functional ultrasound imaging (fUSI) is a breakthrough technology combining large depth of field, high spatiotemporal resolution and affordability[8,12]. fUSI tracks hemodynamic changes in small vessels at several frames per second using ultrafast plane-wave illumination of the tissue[13–15]. Similar to BOLD-fMRI, the fluctuations in cerebral blood volume which are measured with fUSI are a proxy of local neuronal activity[16–19]. In particular, fUSI has recently been used in preclinical research to identify the brain regions involved in the optokinetic reflex[17], to map the networks activated by specific cell types in the superior colliculus[18] and to report functional activity in primates during active tasks and visual stimulation[20,21].

A peculiarity of the technology is that, unlike fMRI, a fUSI recording only provides information on the local vasculature - termed a micro-Doppler image - which is a non-standard anatomical representation. Navigating the brain using this information has proven to be challenging and requires expert knowledge of the vasculature. Indeed, for interpreting the recorded activity, scientists need to be able to map the hemodynamic changes to the corresponding anatomical regions. This is typically performed by registering the data to a reference atlas such as the Allen Brain[22] or Paxinos atlas[23]. To achieve such registration, it is necessary to identify the anatomical position at which the data is acquired precisely. This is usually done by looking for morphological similarities between the recorded image and the anatomical reference atlas, e.g., cortical thickness or anatomical structures such as the ventricles and vascular landmarks. This topic has been recently addressed in ref.[24], who developed an automated ultrasound-based neuro-navigation system to roughly identify their position in the brain. In their approach, they couple online registration of the recorded micro-Doppler volume to a pre-registered reference volume acquired at the start of every experiment.

Here we propose an alternative deep learning approach leveraging the power of Convolutional Neural Networks (CNN) for image classification. Solely based on the vasculature, it allows for accurately locating single micro-Doppler images without the need for a pre-registered reference. Instead of predicting real-valued positions by registering the recorded micro-Doppler image to a vascular template, we defined multiple target classes, each corresponding to an anatomical position. This set of key positions forms a reference grid on the brain from which one can infer the actual location of the recorded image.

We selected a DenseNet121-CNN and an additive chi-square kernel support vector machine with a histogram of oriented gradients features extractor (HOG-SVM) respectively as main and baseline models[25,26], based on a prior performance evaluation. Both were trained for classification on high-quality micro-Doppler images from *in vivo* acquisitions on a set of 51 rats accounting for the intrinsic morphological variability encountered in brain-wide navigation. First, we compared the two models in terms of effect of the spacing between consecutive classes on the classification accuracy. The spacing with the highest classification accuracy allowed us to identify a suitable set of key anatomical positions. Second, we calculated the expected resolution we can achieve based on these key positions. Third, we used the Gradient-weighted Class Activation Map (GradCAM) technique to locate the discriminative features underlying accurate classification[27]. Finally, the robustness of the model was evaluated in a rat stroke model where the cerebral blood flow is locally disrupted. Our approach has shown to be robust to such image degradation, with a very limited effect on the quality of the predictions.

Results

Micro-Doppler datasets. The work presented here was performed using a micro-Doppler dataset that was acquired from 51 rats. By stepping the ultrasound probe along the antero-posterior axis of the brain, images were acquired in different coronal planes, with an in-plane resolution of 100×110 µm and 300 µm slice thickness. The step size between consecutive images was set at 125 µm (Fig. 1a) from the anatomical reference point Bregma (B) +3.0 to -6.5 mm. Fig. 1.b displays example micro-Doppler images, on which major anatomical structures were annotated.

Anatomical reference grid: optimal spacing of key positions based on their classification accuracy. To determine which set of key positions is the most suited for forming a reference grid, five datasets were created by down-sampling the initial brain scans from the original step size (125 µm) to step values ranging from 250 to 750 µm (Fig. 1c). Each dataset can be identified by its step size subscripted with the corresponding number of positions, e.g., '500_{20}' stands for the dataset with a 500 µm step size therefore comprising 20 images per animal. This dataset is depicted in Supplementary Fig. 1, with major vessels labeled using a vascular atlas[28].

After a preliminary performance evaluation (Supplementary Table 1, Materials and Methods – Model selection), a DenseNet121-CNN and HOG-SVM with additive chi-square kernel were selected respectively as main and baseline models[21,22]. Both were trained to classify images with respect to their anatomical position on a subset of 25 rats. The hyperparameters were tuned on a validation set of 13 rats, and their performance were assessed on a testing set of 13 animals.

For both models, we observed an increase in the classification accuracy along with the step size, ranging from respectively 56.2% / 58% (dataset 250_{39}) to respectively 98.2% / 95.9% (dataset 750_{13}), for DenseNet121-CNN and HOG-SVM (Fig. 1d, Supplementary Table 2). The validation and testing accuracies are broadly similar, and interestingly, the largest drop in performance as compared to the previous dataset occurs for dataset 250_{39} (-25.7% for DenseNet121 and -22.4% for HOG-SVM, Supplementary Table 2). This is the only dataset whose step size is below the technology resolution, i.e., 300 µm.

The testing accuracy is lower for the HOG-SVM compared to the DenseNet121-CNN, irrespective of the step size. However, the performance comparison using the McNemar[29] statistical test exhibited no statistically significant differences apart from dataset 500_{20}

(**p=0.0012, Supplementary Table 2). With regards to the extrema, both models achieve similar maximum class accuracy, while the DenseNet121-CNN provides respectively higher and lower minimum class accuracy than the HOG-SVM for datasets 625_{15} / 500_{20} and 375_{26} / 250_{39}. We selected dataset 500_{20} as the set of key positions as the 500 µm step size offers a good trade-off between the number of positions (20) and their identification confidence (DenseNet121-CNN: 93.1%; HOG-SVM: 85.0%).

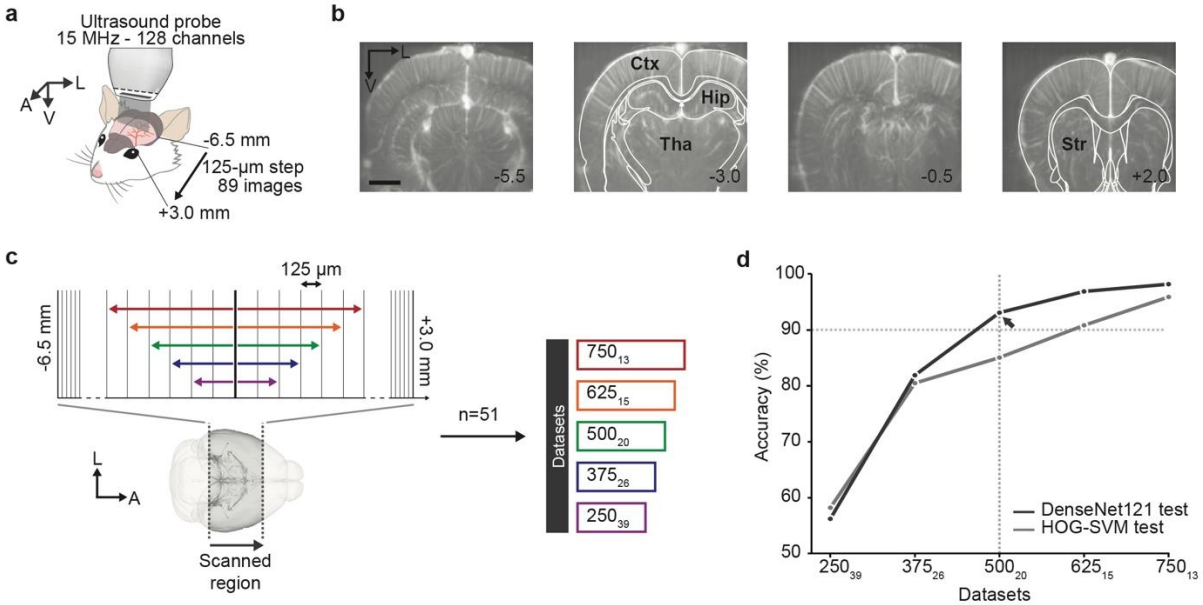


Fig. 1 Selection of the key positions based on classification accuracy. a, Schematic representation of the setup used for micro-Doppler imaging on rats. The ultrasonic probe was positioned along the antero-posterior axis using a motorized linear stage. The imaging was performed from B +4.0 to -7 mm with a 125 µm step size for a total number of 89 images. A: anterior, L: left, V: ventral. **b,** Example micro-Doppler images extracted from a single rat. A simplified version of the Paxinos brain atlas[23] is overlaid in white. Large anatomical structures are identified in black. The number in lower right corner corresponds to the Bregma position (mm) of the image. Ctx: cortex, Hip: hippocampus, Tha: thalamus, Str: striatum. Scale bar: 2 mm. **c,** Schematic representation of the datasets created for determining the best set of key positions. Left: each set of 2D scans is down-sampled with 5 different factors, corresponding to the increase in the step size between two consecutive images, illustrated with colored arrows. Right: Each dataset can be identified by its step size subscripted with the corresponding number of positions (<spatial step>$_{\text{number of locations}}$). Arrows color code: purple 250 µm, blue 375 µm, green 500 µm, orange 625 µm, red 750 µm. **d,** Testing classification accuracy of DenseNet121 and HOG-SVM models for each set of 2D scans ordered by their corresponding step size (n=13). The first set with accuracy above 90% is 500_{20}.

Resolution assessment. We further investigated the per-Bregma position classification accuracies (Fig. 2a, Supplementary Table 3), which are non-uniformly distributed and range from 80 to 100%. The DenseNet121-CNN outperforms the HOG-SVM at every position, and the anterior part of the brain exhibits in general lower accuracies for both. The DenseNet121-CNN-associated confusion matrix reveals that misclassifications map to neighboring classes (Fig. 2b) and were not concentrated in a subset of animals.

Based on this observation, we evaluated the actual resolution of our methodology. For each position, we created an additional dataset comprising of the position itself and images from the four closest anterior and posterior planes. The interval between those neighboring planes is 125 µm, which is the smallest step size available from the original scan. For instance, the dataset for the position B +1.0 mm consists of positions +0.500, +0.625, +0.750, +0.875, +1.000, +1.125, +1.250, +1.375, +1.500 mm. All data from the validation and test sets were aggregated together.

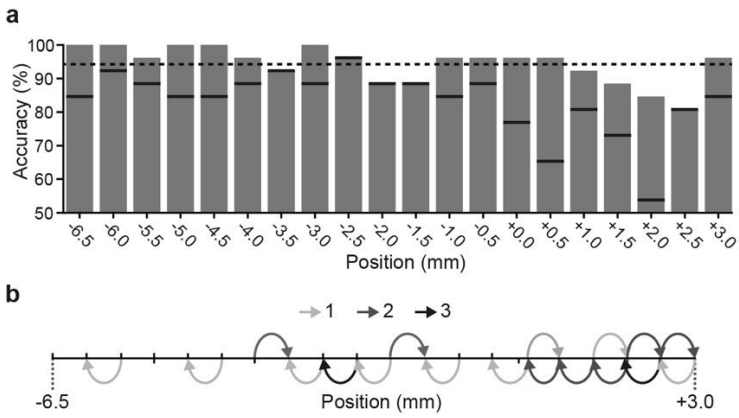


Fig. 2 Per-key position analysis of the DenseNet121 and HOG-SVM prediction. a, Per-Bregma position display of the DenseNet121 model accuracies for the dataset 500_{20} evaluated on 26 animals, including validation and testing sets. The horizontal dashed line represents the mean classification accuracies of DenseNet121. For each position, the black bar represents the HOG-SVM model accuracy. **b,** Graphical representation of the DenseNet121 misclassifications. Each arrow goes from the true position to the predicted position. The light grey to black intensity represents the number of rats for which cross-sections were misclassified.

We processed each position with the DenseNet121-CNN model trained on the dataset 500_{20} and collected the output probabilities. Then we identified the position with the highest probability and computed its absolute deviation from the target position. The mean and standard deviation of this absolute deviation are presented position-wise in Table 1.

Bregma position (mm)	Mean (µm)	Standard deviation (µm)
-6.5	NA	NA
-6.0	135	104
-5.5	87	103
-5.0	87	76
-4.5	77	78
-4.0	96	111
-3.5	48	78
-3.0	77	61
-2.5	87	76
-2.0	87	58
-1.5	77	78
-1.0	115	115
-0.5	135	77
0.0	135	77
0.5	115	104
1.0	106	96
1.5	115	104
2.0	106	108
2.5	163	142
3.0	NA	NA
Average	**102**	**98**

Table 1. Absolute deviation from the target position at each key position. Mean and standard deviation of the absolute deviation from the target position. For a given key position T, such absolute deviation corresponds to the positioning error made when taking the maximum probability of class T in a set comprising the target position and the four closest anterior and posterior planes.

On average, the set of key positions in the dataset 500_{20} are identified with a tolerance of 102 ± 98 µm. Though the positions with highest deviations do not always correspond to the lowest classification accuracy, most of the highest mean and standard deviation are in the anterior part of the dataset.

Spatial localization of the discriminative patterns. To further estimate the reliability of the key positions, we searched for the discriminative features using Gradient-weighted Class Activation Maps (GradCAM)[27], a visualization technique that highlights the image areas contributing the most to the network's inference toward a given prediction[27]. We registered the 2D scans with a digital version of the rat Paxinos atlas[23] to adjust for potential differences in probe positioning and to allow for inter-animal comparison. For each position, the GradCAM results were averaged across animals and thresholded to alleviate the effect of interpolating these low resolution heatmaps (Fig. 3a).

The averaged maps obtained for the set of keys positions were overlaid on the corresponding set of registered micro-Doppler images (Fig. 3b). According to the outcome, a single part in the image is driving the classification irrespective of the location in the brain, at the exception of B -2.0 mm which exhibits two small, connected areas. The identification of the local vasculature associated with the heat maps' location reveal that the branches of several large vessels play a major role in the classification process (Supplementary Fig. 1) including the thalamo-perforating arteries diverging from the posterior cerebral artery (PCA), the thalamostriate veins and branches (tlv), and the patterns produced by neighboring vessels such as the great cerebral vein of Galen and the longitudinal hippocampal veins. Furthermore, we observed that the classification of the four most anterior cross-sections (B +1.5, +2.0, +2.5, +3.0 mm) mostly relies on the anterior cerebral artery (ACA), the azygos pericallosal arteries (APCA) and the thalamostriate veins/arteries. Most of these vessels supply brain regions located in subcortical regions, such as the thalamus, the hippocampus and the striatum as shown in Fig. 1c.

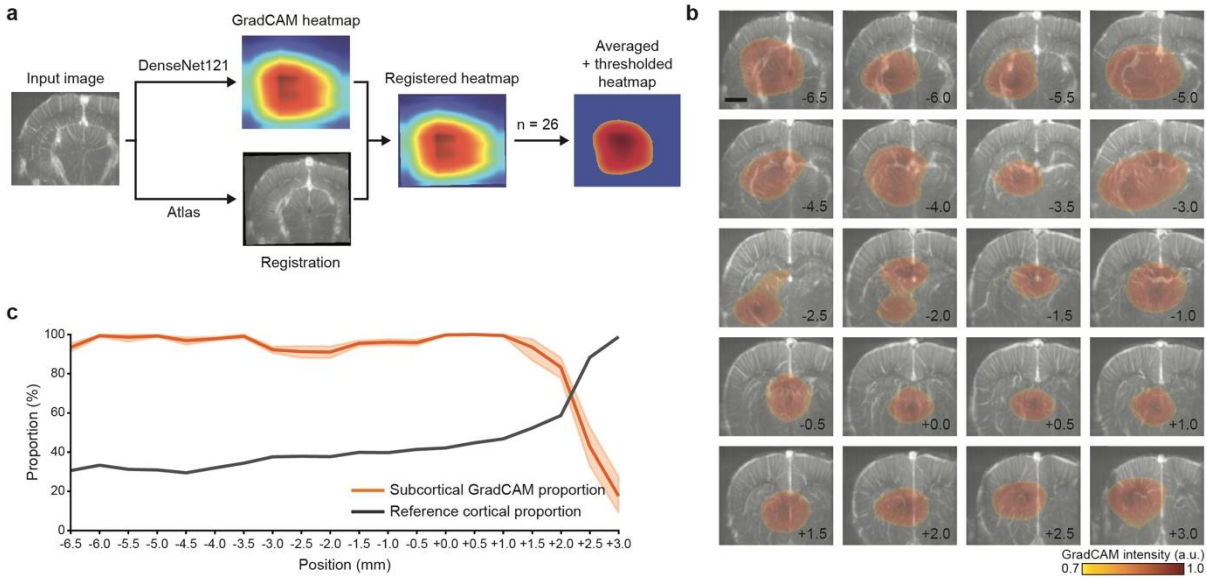


Fig. 3 Visualization of image areas driving DenseNet121's predictions using Gradient-weighted Class Activation Map (GradCAM) on the set of key positions. a, Schematic representation of the workflow to produce the average heatmap. Each image is processed through the DenseNet121-CNN model to collect the GradCAM heatmaps, and in parallel registered for in-plane alignment. The 26 heatmaps, including validation and testing sets, are averaged and thresholded at 0.7. **b,** Display of the averaged heatmaps overlaid on registered reference micro-Doppler images at each position. The color scale indicates the GradCAM intensity (arbitary unit, a.u.). Number corresponds to the Bregma position (mm) of the image. Scale bar: 2 mm. **c,** For each position, the proportion of the GradCAM heatmap located in the subcortex is presented in orange. The error bar corresponds to the 95% confidence interval. The black curve displays the corresponding proportion of cortex and was computed from a reference individual.

Evaluating model robustness on a cortical stroke model. To further validate the reliability of subcortical vascular patterns for accurate position identification, we assessed the performance of both HOG-SVM and DenseNet121-CNN on a cortical stroke model. In these experiments, a subset of 28 rats were subjected to stroke by means of the permanent occlusion of the left middle cerebral artery provoking a significant decrease of signal (-60%) in the cortex of the left hemisphere. Whole-brain micro-Doppler scans were acquired before and 70 mins after stroke induction (Fig. 4a; see ref.[30] for details). From these 28 datasets, 14 were part of the training set, 7 of the validation set and 7 of the testing set.

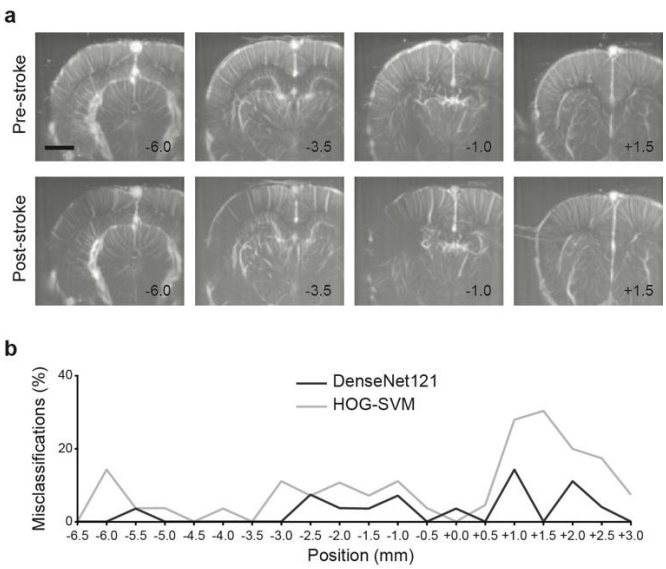


Fig. 4 Effect of a cortical stroke on the classification accuracy. a, Set of micro-Doppler images before (top row) and after stroke induction (bottom row). Number corresponds to the Bregma position (mm) of the image. Scale bar: 2 mm. **b,** Proportion of micro-Doppler images misclassified after the stroke induction at each key position, for DenseNet121 (black curve) or HOG-SVM (grey curve; n=28).

To evaluate the accuracy of our models in such pathological conditions, the positions of the post-stroke images were predicted with both DenseNet121-CNN and HOG-SVM without prior re-training. For each model, the proportion of accurately classified pre-stroke images whose position were incorrectly inferred after stroke was computed (Fig. 4b). Images originally from the training, validation and testing sets were aggregated together for this experiment as no substantial differences were observed. For the DenseNet121-CNN model, the overall proportion of misclassification is 2.7%. Only two positions exhibit a proportion higher than 10% (4/28 and 3/27 images respectively for B +1.0 and +2.0 mm) while for 75% of the positions only one image is misclassified. For the HOG-SVM model, the overall proportion of misclassifications is 8.5%. 8 positions exhibit a proportion higher than 10%, including 3 over 20%, and 45% of the positions have at most one misclassified image. For both models we observed that the stroke has a larger effect on the predictions in the anterior part of the brain (Fig. 4c).

Discussion

In this study, we proposed a novel framework for brain navigation in functional ultrasound imaging experiments. Our approach relies on CNN-based image classification to identify a set of anatomical positions that serve as a reference frame, from which the location of a micro-Doppler image can be inferred with high precision. We selected a DenseNet121-CNN as main model and an HOG-SVM as baseline model.

First, we defined a set of anatomical reference positions which act as the classes in our image classification task. By analyzing the effect of different step sizes on the classification accuracy we were able to select a set of key positions offering a trade-off between the identification confidence and the number of positions in the reference grid. We concluded that with a 500 µm step size, both models accurately classify each image to the corresponding anatomical position (>90%). The effect of the step size on the accuracy above 375 µm is linear-like for HOG-SVM while exponential-like for DenseNet121. This can explain the statistically significant difference in performance on datasets 500_{20} and 625_{15}. Such decrease can be attributed to the similarity in vasculature across neighboring planes at the technology resolution. This hypothesis is supported by the drop in the minimum class accuracy along with the step size for both models. We then computed on the set of key positions an estimated positioning error of 102 ± 98 µm for DenseNet121, which is smaller than the micro-Doppler image thickness (300 µm) and therefore sufficient.

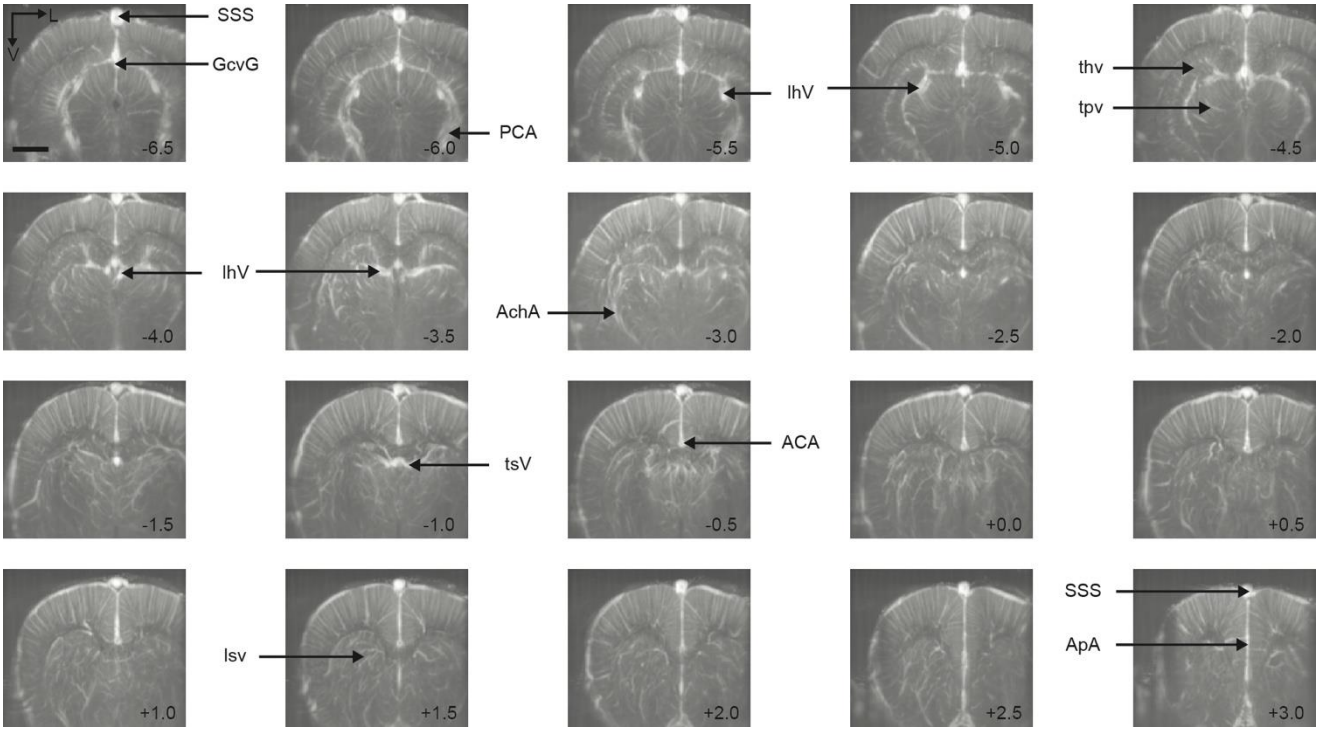
Further analysis using the GradCAM visualization technique revealed that the classification was mainly driven by highly consistent vascular structures located in the subcortex. This subcortical prevalence can partially be explained by the in-plane resolution of the current fUSI technology: the voxel size of 100×110×300 μm^3 is likely insufficient to highlight the vascular differences between cortical areas, where the penetrating arterioles have a diameter between 50 and 100 µm[1]. Another potential explanation comes from the CNNs sensitivity to texture differences when pretrained on ImageNet. This might explain why the cortical curvature and thickness variation across anatomical locations are not decisive factors in our approach. Finally, we validated the CNN's predictions in a rat stroke model where normal blood flow is disrupted. Analysis revealed that the number of misclassified images was marginal compared to the pre-stroke dataset (2% for DenseNet121-CNN, 8% for HOG-SVM), thus confirming the robustness of the inference.

The automated positioning of micro-Doppler images is a problem that has only recently been tackled. At the time of writing there is only one publication on this topic, where the recorded micro-Doppler image is automatically registered to a pre-aligned reference volume[31]. The main drawback of this method lies in the acquisition of a reference micro-Doppler volume at the start of every experiment, which is manually registered to a brain atlas. Our approach, on the other hand, does not require an animal-specific reference and therefore offers more flexibility in the experimental design. Furthermore, due to the large training dataset, our model is less sensitive to differences in brain size and shape, including in pathological conditions such as stroke. This will be strengthened even further as more fUSI data becomes available. Finally, the CNNs' computational efficiency allows for real-time image identification and can therefore be seamlessly integrated in an experimental workflow.

Although our work is validated on a rat model, we expect that the presented approach is universal and can easily be adapted to different use cases, including applications in other animal models such as non-human primates, pigeons or ferrets[20,21,32,33]. Early results support the applicability of our methodology to mice datasets (data not shown) and elicited similar results. Future work should also focus on extending the model to different probe orientations (e.g. sagittal) or different types of ultrasound transducers such as the recently developed volumetric fUSI system[34,35], which acquires dozens of planes simultaneously at the cost of a lower spatial resolution. Additionally, micro-Doppler imaging has also been successfully applied to humans in neurosurgery[36–38] and non-invasively in newborns by imaging through the fontanel[39]. In those contexts, an accurate positioning methodology would be of great value but comes with new challenges, such as the limited depth of imaging and the large differences in vessel scales. The increasing adoption of fUSI and related data diversity will allow further generalization of this approach.

To conclude, we believe that our methodology will constitute a valuable tool for the neuroscientific community in the coming years, as it will allow non-expert users to exploit the full potential of the fUSI technology.

Supplementary Materials



Supplementary Fig. 1, Set of micro-Doppler images extracted from the dataset 500_{20} covering a large part of the brain for one rat. Major vessels are identified in black. Number corresponds to the Bregma position (mm) of the image. ACA: anterior cerebral artery, AchA: anterior choroidal artery, ApA: azygos pericallosal artery, GcvG: great cerebral vein of Galen, lhV: longitudinal hippocampal vein, lsv: lenticulostriate vessels, thv: transverse hippocampal vessels, tsV: thalamostriate vein, tpv: thalamoperforating vessels, SSS: superior sagittal sinus, Ctx: cortex, Hip: hippocampus, Tha: thalamus, Str: striatum, A: anterior, L: left, V: ventral. Scale bar: 2 mm.

Model	ResNet50	DenseNet121 (main)	HOG-SVM chi2 (baseline)	HOG-SVM rbf	SIFT-SVM rbf	PCA-SVM rbf
Validation accuracy (%)	82.5	85.2	71.6	65.1	29.3	53.5
Testing accuracy (%)	77.2	81.9	80.4	68.3	24.5	64.5

Supplementary Table 1. Performance evaluation of a set of classical models on dataset 375_{26}.

Step size (µm)		250	375	500	625	750
Number of positions		**39**	**26**	**20**	**15**	**13**
DenseNet121	Validation accuracy (%)	64.3	85.2	95.4	100	100
	Test accuracy (%)	56.2	81.9	93.1	96.9	98.2
	Difference between test and validation accuracy	8.1	3.3	2.3	3.1	1.8
	Maximum class accuracy test (%)	84.6	100	100	100	100
	Minimum class accuracy test (%)	0.0	30.8	84.6	84.6	84.6
HOG-SVM	Validation accuracy (%)	54	71.6	81.9	88.7	92.9
	Test accuracy (%)	58	80.4	85	90.8	95.9
	Difference between test and val. Accuracy	-4	-8.8	-3.1	-2.1	-3
	Maximum class accuracy (%)	84.6	100	100	100	100
	Minimum class accuracy (%)	20.3	61.5	53.8	53.8	84.6
McNemar test	Test statistic	86.000	28.000	12.000	4.000	1.000
	P-value	0.552	0.450	0.002 (**)	0.012 (*)	0.219

Supplementary Table 2. DenseNet121 and HOG-SVM performance metrics on the datasets with different step sizes. *p-value<0.05, **p-value<0.01.

Bregma position (mm)	-6.5	-6.0	-5.5	-5.0	-4.5	-4.0	-3.5	-3.0	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5	3.0
Validation set																				
Accuracy (%)	100	100	100	100	100	100	100	100	100	92.3	84.6	92.3	92.3	92.3	100	100	92.3	84.6	76.9	100
Precision	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.93	0.92	0.92	0.86	1.00	1.00	0.93	0.93	0.92	0.92	0.91	0.87
Recall	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.92	0.85	0.92	0.92	0.92	1.00	1.00	0.92	0.85	0.77	1.00
F1-score	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0.92	0.88	0.89	0.96	0.96	0.96	0.96	0.92	0.88	0.83	0.93
Testing set																				
Accuracy (%)	100	100	92.3	100	100	92.3	84.6	100	92.3	84.6	92.3	100	100	100	92.3	84.6	84.6	84.6	84.6	92.3
Precision	1.00	0.93	1.00	1.00	0.93	1.00	1.00	0.81	0.86	1.00	1.00	0.93	1.00	0.93	0.86	0.92	0.92	0.79	0.85	1.00
Recall	1.00	1.00	0.92	1.00	1.00	0.92	0.85	1.00	0.92	0.85	0.92	1.00	1.00	1.00	0.92	0.85	0.85	0.85	0.85	0.92
F1-score	1.00	0.96	0.96	1.00	0.96	0.96	0.92	0.90	0.89	0.92	0.96	0.96	1.00	0.96	0.89	0.88	0.88	0.81	0.85	0.96

Supplementary Table 3. Per Bregma position classification metrics on dataset 500_{20} of DenseNet121.

restricted

Materials and Methods

Animals. Experimental procedures were approved by the Committee on Animal Care of the Catholic University of Leuven, in accordance with the national guidelines on the use of laboratory animals and the European Union Directive for animal experiments (2010/63/EU). Adult male Sprague-Dawley rats (n=51; Janvier Labs, France) with an initial weight between 200-300 g were housed in standard ventilated cages and kept in a 12:12 hrs reverse dark/light cycle environment at a temperature of 22 °C with *ad libitum* access to food and water.

Cranial window for brain-wide imaging and stroke induction. A cranial window extended from B +4.0 to −7.0 mm AP, laterally ±6.0 mm was performed in all rats under isoflurane anesthesia (Iso-Vet, Dechra, Belgium) with a mixture of 5% isoflurane in compressed dry air was used to induce anesthesia, subsequently reduced to 2.0-2.5% during surgery, and to 1.5% for imaging (see ref.[30] for details on surgical procedure). Xylocaine (0.5%, AstraZeneca, England) and Metacam (0.2mg/kg, Boehringer Ingelheim, Canada) were injected subcutaneously as pre-operative and post-operative analgesia; respectively. Intraperitoneal injection of 5% glucose solution was provided every 2hrs to prevent dehydration. 28 rats were subjected to stroke by the mean of permanent occlusion of the distal branch of the left middle cerebral artery as detailled in ref.[30].

2D scan micro-Doppler ultrasound imaging of brain vasculature. The data acquisition was performed using a functional ultrasound imaging scanner equipped with custom acquisition and processing software described in ref.[11]. In short, the scanner is composed of a linear ultrasonic transducer (15 MHz, 128 elements, Xtech15, Vermon, France) connected to 128-channel emission-reception electronics (Vantage, Verasonics, USA) that are both controlled by a high-performance computing workstation (fUSI-2, AUTC, Estonia). The transducer was motorized (T-LSM200A, Zaber Technologies Inc., Canada) to allow antero-posterior scanning of the brain. The acoustic coupling between the brain and the probe is ensured by a 2 mm layer of ultrasound gel (Aquasonic Clear, Parker Laboratories Inc, USA). Each coronal Doppler image is 12.8 mm width and 9 mm depth and is composed of 300 compound images acquired at 500 Hz. Each compound image is computed by adding nine plane-wave (4.5 kHz) with angles from -12° to 12° with a 3° step. The blood signal was extracted from 300 compound images using a single value decomposition filter and removing the 30 first singular vectors[40]. The Doppler image is computed as the mean intensity of the blood signal in these 300 frames that is an estimator of the cerebral blood volume[14,15]. This sequence enables a temporal resolution of 0.6 sec, an in-plane resolution of 100×110 µm, and an off-plane (thickness of the image) of 300 µm[11]. Finally, we performed a high-resolution 2D scan of the brain vasculature consisting of 89 coronal planes from B +4.0 to -7.0 mm spaced by 125 µm.

Registration of micro-Doppler images. The micro-Doppler 2D scans from all animals were aligned along the antero-posterior axis with respect to 2 reference cross-sections (B -3.0 and -1.0 mm) selected for their recognizable vascular patterns (Fig. 1c). This alignment for correcting potential shifts occurring either during surgery or imaging. Reference cross-sections were independently identified for every animal by two experts. Any disagreement was resolved post-hoc by consensus. Each micro-Doppler image is then identified by its anatomical position with respect to the Bregma reference point, e.g., B -3.0 mm.

Generation of datasets. Several datasets have been extracted from the initial scans using a down-sampling factor ranging from 2 to 5. This corresponds to an artificial increase in the step size between two consecutive cross-sections. To create the dataset associated with a given factor F, we extracted images from position B -3.0 mm with a step size of Fx125 µm, within the limits of the craniotomy (Fig. 1b). The 5 datasets stepped by [250, 375, 500, 625, 750] µm, respectively contain [39, 26, 20, 15, 13] different positions. We randomly selected 50% of the animals for training, 25% for tuning the hyperparameters (validation) and 25% for evaluating the final performances of the model (testing). We augmented the size of the training set with rotations of ±4°and ±8°.

Image preprocessing. To increase the contrast and reduce the intensity amplitude to a [0 1] interval, a correction factor (power of 0.25) has been applied to every pixel of all images in each dataset. The overall process has been implemented using MATLAB (R2018b, Mathworks, USA).

Model selection. To select the best model for the experiments, we evaluated 2 classical CNN architectures (ResNet50[41] and DenseNet121[25]), SVMs with different feature extraction methods (HOG[26], SIFT[42], PCA) and kernels[43]. These models were selected for their compatibility with datasets of relatively small sizes. Both ResNet50 and DenseNet121 were pretrained on ImageNet[44] as suggested in ref.[45]. The last layer of the network - the classifier - was replaced by a fully connected layer outputting n values, n being the number of anatomical locations, and passed through a softmax layer afterwards. Both CNN and SVM models were trained and evaluated on the dataset 375_{26}, corresponding to the smallest step size above the technology resolution along the antero-posterior axis and therefore the largest dataset without overlapping information. They were respectively implemented with the 'torchvision' (PyTorch, version 0.7.0). and 'scikit-learn' (version 0.23.1) Python packages.

Training and evaluation procedures for CNNs. For each of the datasets used in this work, images were resized to 224×320 pixels by bicubic interpolation, and their grey channel extended in RGB to fit the ImageNet format imposed by the pre-training. All the data were normalized with the mean and standard deviation of the full dataset. We augmented the size of the training set with rotations of ±4°and ±8°. The network's weights were optimized with the stochastic gradient descent algorithm using a cross-entropy loss function. The hyperparameters were selected through a random search and the final model performance was evaluated on the testing set. The overall procedure has been performed on a single machine, equipped with Xeon E5-2620 CPU (Intel, USA), 64 Gb RAM and 4 RTX2080 (8 GB) GPUs (Nvidia, USA).

Visualization of relevant features for image classification using GradCAM. We extracted the pixels in the input image driving the classification using the Gradient-weighted Class Activation Map (GradCAM) technique, following the recommendations from ref.[46] on the relevant visualization approaches. Briefly, this method aggregates the gradients associated with the prediction for each feature map in a given layer, to produce a coefficient measuring the contribution of each of the map to the network's prediction. Here, the gradients and feature maps were extracted at the last layer before the classifier. The output heatmaps were then resized by bilinear interpolation to the original image and thresholded at 0.7 to limit the effect of interpolation on the map.

GradCAM registration on stereotaxic atlases for anatomical regions extraction. We used a digital version of the rat Paxinos atlas[23,30] to extract the anatomical regions associated with the GradCAM. The input scan was taken as a volume and interpolated to fit the atlas resolution (50×50×50 μm^3 voxel size). A 3D rigid registration was performed using a MATLAB custom script[11,30]. This procedure has been applied to all the samples from the validation and testing sets by an expert. To extract the regions from the GradCAM heatmap, a volume (89 planes as the input data) was constructed from the heatmaps by zero-padding the missing sections before applying the transformation matrix.

Evaluation on the stroke dataset. We used 28 rats subjected to stroke (see above and ref.[30]). All rats were imaged in the original dataset, and 14/7/7 individuals were respectively present in training/validation/testing sets. The scans were registered and a dataset with 500 µm step size was created following the same procedure as for the previous experiment. The classes predictions were obtained by processing the images through the DenseNet121-CNN and HOG-SVM previously trained on dataset 500_{20} without re-training.

REFERENCES

1. Blinder, P. *et al.* The cortical angiome: an interconnected vascular network with noncolumnar patterns of blood flow. *Nat. Neurosci.* **16**, 889–897 (2013).
2. Kirst, C. *et al.* Mapping the fine-scale organization and plasticity of the brain vasculature. *Cell* **180**, 780-795.e25 (2020).
3. Todorov, M. I. *et al.* Machine learning analysis of whole mouse brain vasculature. *Nat. Methods* **17**, 442–449 (2020).
4. Attwell, D. *et al.* Glial and neuronal control of brain blood flow. *Nature* **468**, 232–243 (2010).
5. Rungta, R. L., Chaigneau, E., Osmanski, B.-F. & Charpak, S. Vascular compartmentalization of functional hyperemia from the synapse to the Pia. *Neuron* **99**, 362-375.e4 (2018).
6. Iadecola, C. The neurovascular unit coming of age: A journey through neurovascular coupling in health and disease. *Neuron* **96**, 17–42 (2017).
7. Kaplan, L., Chow, B. W. & Gu, C. Neuronal regulation of the blood-brain barrier and neurovascular coupling. *Nat. Rev. Neurosci.* **21**, 416–432 (2020).
8. Urban, A. *et al.* Understanding the neurovascular unit at multiple scales: Advantages and limitations of multi-photon and functional ultrasound imaging. *Adv. Drug Deliv. Rev.* **119**, 73–100 (2017).
9. Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878 (2008).
10. Matthews, P. M., Honey, G. D. & Bullmore, E. T. Applications of fMRI in translational medicine and clinical practice. *Nat. Rev. Neurosci.* **7**, 732–744 (2006).
11. Brunner, C. *et al.* Whole-brain functional ultrasound imaging in awake head-fixed mice. *Nat. Protoc.* **16**, 3547–3571 (2021).
12. Edelman, B. J. & Macé, E. Functional ultrasound brain imaging: Bridging networks, neurons, and behavior. *Curr. Opin. Biomed. Eng.* **18**, 100286 (2021).
13. Bercoff, J. *et al.* Ultrafast compound Doppler imaging: providing full blood flow characterization. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **58**, 134–147 (2011).
14. Macé, E. *et al.* Functional ultrasound imaging of the brain. *Nat. Methods* **8**, 662–664 (2011).
15. Mace, E. *et al.* Functional ultrasound imaging of the brain: theory and basic principles. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **60**, 492–506 (2013).
16. Urban, A. *et al.* Chronic assessment of cerebral hemodynamics during rat forepaw electrical stimulation using functional ultrasound imaging. *Neuroimage* **101**, 138–149 (2014).
17. Macé, É. *et al.* Whole-brain functional ultrasound imaging reveals brain modules for visuomotor integration. *Neuron* **100**, 1241-1251.e7 (2018).
18. Sans-Dublanc, A. *et al.* Optogenetic fUSI for brain-wide mapping of neural activity mediating collicular-dependent behaviors. *Neuron* **109**, 1888-1905.e10 (2021).
19. Nunez-Elizalde, A. O. *et al.* Neural basis of functional ultrasound signals. *bioRxiv* (2021) doi:10.1101/2021.03.31.437915.
20. Dizeux, A. *et al.* Functional ultrasound imaging of the brain reveals propagation of task-related brain activity in behaving primates. *Nat. Commun.* **10**, 1400 (2019).
21. Norman, S. L. *et al.* Single-trial decoding of movement intentions using functional ultrasound neuroimaging. *Neuron* **109**, 1554-1566.e4 (2021).
22. Wang, Q. *et al.* The Allen mouse brain common coordinate framework: A 3D reference atlas. *Cell* **181**, 936-953.e20 (2020).
23. Paxinos, G. & Watson, C. *Rat brain in stereotaxic coordinates*. (Academic Press, 2007).
24. Nouhoum, M. *et al.* A functional ultrasound brain GPS for automatic vascular-based neuronavigation. *Sci. Rep.* **11**, 15197 (2021).

25. Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2017). doi:10.1109/cvpr.2017.243.
26. Dalal, N. & Triggs, B. Histograms of oriented gradients for human detection. in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (IEEE, 2005). doi:10.1109/cvpr.2005.177.
27. Selvaraju, R. R. *et al.* Grad-CAM: Visual explanations from deep networks via gradient-based localization. in *2017 IEEE International Conference on Computer Vision (ICCV)* (IEEE, 2017). doi:10.1109/iccv.2017.74.
28. Dorr, A., Sled, J. G. & Kabani, N. Three-dimensional cerebral vasculature of the CBA mouse brain: a magnetic resonance imaging and micro computed tomography study. *Neuroimage* **35**, 1409–1423 (2007).
29. McNEMAR, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **12**, 153–157 (1947).
30. Brunner, C. *et al.* Brain-wide continuous functional ultrasound imaging for real-time monitoring of hemodynamics during ischemic stroke. (2022) doi:10.1101/2022.01.19.476904.
31. Nouhoum, M. *et al.* Fully-automatic ultrasound-based neuro-navigation : The functional ultrasound brain GPS. *Research Square* (2021) doi:10.21203/rs.3.rs-382732/v1.
32. Rau, R. *et al.* 3D functional ultrasound imaging of pigeons. *Neuroimage* **183**, 469–477 (2018).
33. Demene, C. *et al.* Functional Ultrasound Imaging of the thalamo-cortical auditory tract in awake ferrets using ultrafast Doppler imaging. in *2016 IEEE International Ultrasonics Symposium (IUS)* (IEEE, 2016). doi:10.1109/ultsym.2016.7728659.
34. Brunner, C. *et al.* A platform for brain-wide volumetric functional ultrasound imaging and analysis of circuit dynamics in awake mice. *Neuron* **108**, 861-875.e7 (2020).
35. Rabut, C. *et al.* 4D functional ultrasound imaging of whole-brain activity in rodents. *Nat. Methods* **16**, 994–997 (2019).
36. Imbault, M. *et al.* Functional ultrasound imaging of the human brain activity: An intraoperative pilot study for cortical functional mapping. in *2016 IEEE International Ultrasonics Symposium (IUS)* (IEEE, 2016). doi:10.1109/ultsym.2016.7728505.
37. Imbault, M., Chauvet, D., Gennisson, J.-L., Capelle, L. & Tanter, M. Intraoperative Functional Ultrasound Imaging of Human Brain Activity. *Sci. Rep.* **7**, 7304 (2017).
38. Soloukey, S. *et al.* Functional Ultrasound (fUS) During Awake Brain Surgery: The Clinical Potential of Intra-Operative Functional and Vascular Brain Mapping. *Front. Neurosci.* **13**, (2020).
39. Demene, C. *et al.* Functional ultrasound imaging of the brain activity in human neonates. in *2016 IEEE International Ultrasonics Symposium (IUS)* (IEEE, 2016). doi:10.1109/ultsym.2016.7728657.
40. Urban, A. *et al.* Functional Ultrasound Imaging of Cerebral Capillaries in Rodents and Humans. *Jacobs Journal of Molecular and Translational Medicine* **1**, (2015).
41. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 770–778 (IEEE, 2016).
42. Lowe, D. G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**, 91–110 (2004).
43. Zhang, J., Marszałek, M., Lazebnik, S. & Schmid, C. Local features and kernels for classification of texture and object categories: A comprehensive study. *Int. J. Comput. Vis.* **73**, 213–238 (2007).
44. Deng, J. *et al.* ImageNet: A large-scale hierarchical image database. in *2009 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2009). doi:10.1109/cvpr.2009.5206848.

45. Tajbakhsh, N. *et al.* Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Trans. Med. Imaging* **35**, 1299–1312 (2016).
46. Adebayo, J. *et al.* Sanity Checks for Saliency Maps. *arXiv [cs.CV]* (2018).

Acknowledgements

This work was supported by grants from the Leducq foundation (Network 15CVD02), from FWO (MEDI-RESCU$_2$, number AKUL/17/049), from VIB TechWatch (fUSI-MICE) and from internal NERF funding (3D-fUSI project). This research also received funding from the Flemish regional government (AI Research Program). We also thank the NERF animal caretakers including I. Eyckmans, F. Ooms and S. Luijten for their help with management of the rats.

Author contributions

	Lambert	Brunner	Kil	Wuyts	D'Hondt	Urban
Concept	X					X
Methodology	X					X
Software						X
Imaging		X				
Data Analysis - Interpretation	X	X	X	X		X
Manuscript	X	X	X	X	X	X
Proof reading	X	X	X	X	X	X
Supervision				X	X	X
Funding					X	X

Declaration of interests

A.U. is the founder and a shareholder of AUTC company commercializing functional ultrasound imaging solutions for preclinical and clinical research.