

# Spectroscopic and Deep Learning-Based Approaches to Identify and Quantify Cerebral Microhemorrhages

**Christian Crouzet**

University of California, Irvine

**Gwangjin Jeong**

Dankook University

**Rachel Chae**

Massachusetts Institute of Technology

**Krystal LoPresti**

University of California, Irvine

**Cody Dunn**

University of California, Irvine

**Danny Xie**

University of California, Irvine

**Chiagoziem Agu**

Albany State University

**Chuo Fang**

University of California, Irvine

**Ane Nunes**

University of California, Irvine

**Wei Ling Lau**

University of California, Irvine

**Sehwan Kim**

Dankook University

**David Cribbs**

University of California, Irvine

**Mark Fisher**

University of California, Irvine

**Bernard Choi (✉ [choib@uci.edu](mailto:choib@uci.edu))**

University of California, Irvine

**Keywords:** CMHs, microhemorrhages, pathology

**Posted Date:** January 18th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-135361/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Title:** Spectroscopic and deep learning-based approaches to identify and quantify cerebral microhemorrhages

Christian Crouzet<sup>1</sup>, Gwangjin Jeong<sup>2</sup>, Rachel H Chae<sup>3</sup>, Krystal T LoPresti<sup>1,4</sup>, Cody E Dunn<sup>1,4</sup>, Danny F Xie<sup>1,4</sup>, Chiagoziem Agu<sup>1,5</sup>, Chuo Fang<sup>6</sup>, Ane CF Nunes<sup>7</sup>, Wei Ling Lau<sup>7</sup>, Sehwan Kim<sup>2</sup>, David H Cribbs<sup>8</sup>, Mark Fisher<sup>1,6</sup>, and Bernard Choi<sup>1,4,9,10</sup>

**Affiliations:**

1 Beckman Laser Institute and Medical Clinic, University of California, Irvine, CA, USA

2 Department of Biomedical Engineering, Beckman Laser Institute Korea, Dankook University, Cheonan, 31116, Republic of Korea

3 Massachusetts Institute of Technology, Cambridge, MA, USA

4 Department of Biomedical Engineering, University of California, Irvine, CA, USA

5 Albany State University, Albany, Georgia, USA

6 Neurology and Pathology & Laboratory Medicine, University of California, Irvine, CA, USA

7 Medicine, Division of Nephrology, University of California, Irvine, CA, USA

8 Institute for Memory Impairments and Neurological Disorders, University of California, Irvine, CA, USA

9 Department of Surgery, University of California, Irvine, CA, USA

10 Edwards Lifesciences Center for Advanced Cardiovascular Technology, University of California, Irvine, CA, USA

**Corresponding author name and e-mail:** Bernard Choi, choib@uci.edu

**Abstract:**

Cerebral microhemorrhages (CMHs) are associated with cerebrovascular disease, cognitive impairment, and normal aging. One method to study CMHs is to analyze histological sections (5-40  $\mu\text{m}$ ) stained with Prussian blue. Currently, users manually and subjectively identify and quantify Prussian blue-stained regions of interest, which is prone to inter-individual variability and can lead to significant delays in data analysis. To improve this labor-intensive process, we developed and compared three digital pathology approaches to identify and quantify CMHs from Prussian blue-stained brain sections: 1) ratiometric analysis of RGB pixel values, 2) phasor analysis of RGB images, and 3) deep learning using a mask region-based convolutional neural network. We applied these approaches to a preclinical mouse model of inflammation-induced CMHs. One-hundred CMHs were imaged using a 20x objective and RGB color camera. To determine the ground truth, four users independently annotated Prussian blue-labeled CMHs. The deep learning and ratiometric approaches performed better than the phasor analysis approach compared to the ground truth. The deep learning approach had the most precision of the three methods. The ratiometric approach has the most versatility and maintained accuracy, albeit with less precision. Our data suggest that implementing these methods to analyze CMH images can drastically increase the processing speed while maintaining precision and accuracy.

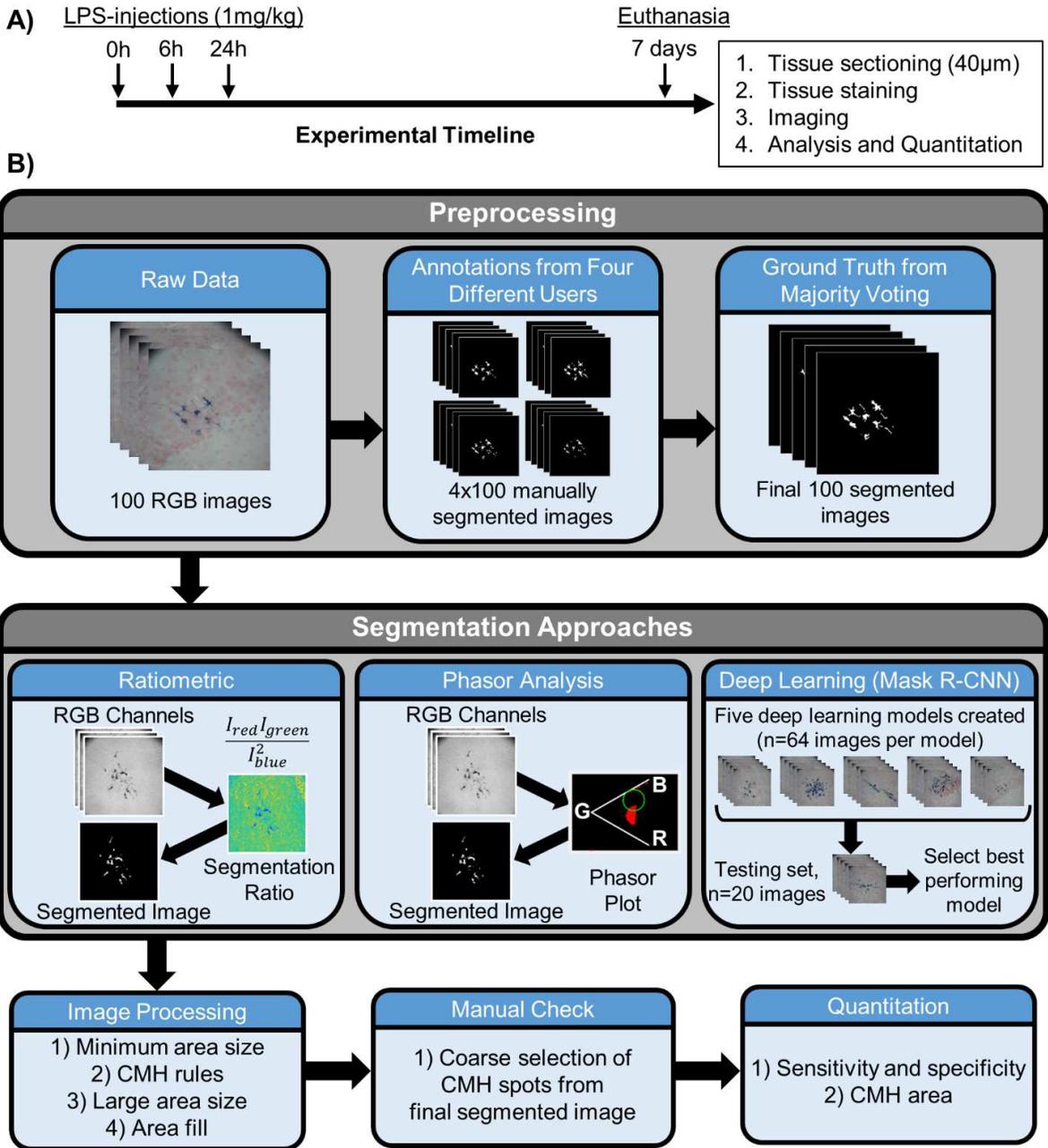
## **Introduction:**

Cerebral microhemorrhages (CMHs), which are detected as cerebral microbleeds on gradient-echo T2\*-weighted magnetic resonance imaging (MRI), are associated with cognitive impairment<sup>1</sup> and an increased risk of hemorrhagic and ischemic stroke<sup>2</sup>. CMHs are present in 20% of individuals over 60 years old and 40% of individuals over 80 years old<sup>3,4</sup>. In addition to normal aging<sup>5,6</sup>, CMHs are increased in several disease states, including cerebral amyloid angiopathy<sup>7,8</sup>, chronic kidney disease<sup>9</sup>, hypertension<sup>10,11</sup>, and cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL)<sup>12</sup>. Despite the large number of medical conditions associated with CMHs, there remains a fundamental gap in understanding their mechanistic formation. One method that facilitates the study of CMH formation is the histological analysis of fixed tissue sections (5-40  $\mu\text{m}$ ) stained with Perl's Prussian blue. Current methods that identify and quantify CMHs are typically subjective and require tedious microscopic inspection. They rely on manual selection of Prussian blue-stained regions or use manual thresholding<sup>13</sup>, which is time and labor intensive. Recent advances in digital pathology address these limitations by eliminating the need for laborious staining<sup>14</sup> and automating the quantitative analysis of the samples<sup>15,16</sup>.

Here, we set out to develop and compare three digital pathology approaches to automatically identify and quantify CMHs from sections stained with Prussian blue: 1) ratiometric analysis of RGB pixel values, 2) phasor analysis of RGB images, and 3) deep learning. The ratiometric method applies a threshold to the ratio of the red and green pixel intensities to the blue intensity-squared. The phasor analysis method applies a mask to the RGB values in the phasor space<sup>17,18</sup>. The deep learning approach uses a mask region-based convolutional neural network (R-CNN) to segment the RGB images<sup>19</sup>. We found that each approach performs well at identifying CMHs and quantifying the CMH area. Furthermore, the deep learning and ratiometric approaches performed better than the phasor analysis approach. We found that the deep learning approach had the most precision, while the ratiometric approach had good accuracy and the most versatility, but less precision than the deep learning approach.

## **Materials and Methods:**

An overview of the experimental design, segmentation approaches, and processing steps are outlined in Figure 1.



**Figure 1: Overarching experimental timeline and data processing scheme.** (A) Experimental timeline to induce cerebral microhemorrhages (CMHs), followed by tissue sectioning, staining, imaging, analysis, and quantitation. (B) Data preprocessing steps include: 1) acquiring raw color images of each CMH, 2) manually annotating color images by four users, and 3) determining the ground truth through majority voting. Three segmentation approaches (ratiometric, phasor analysis, and deep learning) were developed and compared. After the segmentation approaches, each set of results went through a series of image processing, data quality check, and quantitation steps (see the *Image Processing* section for details regarding each approach).

### *Animal Model:*

All animal procedures were performed in accordance with a protocol approved by the University of California, Irvine's Institutional Animal Care and Use Committee. Male C57BL/6J mice (12 weeks old) were used (n=4). Our study design complied with ARRIVE 2.0 guidelines. However, since our overall goal with the animal research was to generate a set of brain histology slides with exogenously labeled CMHs, we did not include a control group or incorporate blinding or randomization practices. To induce CMH formation<sup>20</sup>, we administered three doses of lipopolysaccharide (LPS) (1mg/kg) to the mice at 0h, 6h, and 24h via intraperitoneal injection (Figure 1A). The mice fed and drank ad lib and received up to three daily doses of 1 mL saline subcutaneously. Mice were euthanized seven days after the first injection with an overdose of sodium pentobarbital. After euthanasia, the chest cavity was immediately opened, and cardiac perfusion of saline followed by buffered formalin was performed. The brain was extracted and then immersed in buffered formalin for 24 hours. The brain was then stored in PBS with 0.02% sodium azide<sup>21</sup> until the time of sectioning.

### *Sample Preparation:*

Brains were sectioned into 40- $\mu$ m coronal sections using a freezing microtome. Sections were stained with Prussian blue to detect hemosiderin, a marker of CMHs, using 10% hydrochloric acid and 5% potassium ferrocyanide trihydrate (Sigma, St. Louis, MO) for 30 min and rinsed in deionized water. Sections were then counterstained with Nuclear Fast Red (Sigma, St. Louis, MO) for 5 min, rinsed in deionized water, dehydrated, and mounted on a glass slide.

### *Data Acquisition and Preprocessing*

The stained sections were scanned for CMHs and imaged using a color camera (Chameleon3, FLIR, Wilsonville, OR) coupled to an inverted Nikon microscope. One hundred unique images of Prussian blue-labeled CMHs were acquired and used for subsequent data analysis. Four independent users manually selected CMHs. Majority voting (i.e., at least two of the four users defined a pixel as a CMH) was done to determine the ground truth segmented data.

## **Segmentation Approaches**

The key image processing steps for the segmentation approaches are shown in Supplementary Figure 1.

### *Ratiometric Approach*

The ratiometric approach was developed based on spectroscopic data acquired at 31 wavelengths (Supplementary Figures 2 and 3). We defined a segmentation ratio using the red, green, and blue pixel intensities to spatially discriminate Prussian blue-labeled CMH pixels from background pixels:  $segmentation\ ratio = \frac{I_{red}I_{green}}{I_{blue}^2}$ . Segmentation ratio values that were below the specified threshold were included in the segmented image. We applied multiple segmentation ratio threshold values to 1) obtain segmented data for image processing, 2) calculate a receiver operator characteristic (ROC) curve<sup>22</sup>, and 3) determine the optimal segmentation ratio threshold. Please see the *Data Quantification and Statistical Analysis* section for how the optimal segmentation ratio was determined.

### *Phasor Analysis Approach*

Phasor approaches in biomedical optics have been extensively used in fluorescence lifetime imaging microscopy (FLIM) to identify fluorophores in the collected images<sup>23,24</sup>. Recently, phasor approaches have been applied to spectroscopic data and RGB images from fluorescence microscopy and brightfield microscopy, respectively<sup>17,18</sup>. We applied a spectral phasor analysis technique to RGB images to isolate Prussian blue-labeled CMH pixels from background pixels. A discrete Fourier transform was performed at each pixel, which resulted in a phasor that consists of a real (H) and imaginary (S) component:  $H = \sum_{l=1}^3 I_l \cos \frac{2\pi l}{3}$  and  $S = \sum_{l=1}^3 I_l \sin \frac{2\pi l}{3}$  where  $I$  is the collected intensity at channel  $l$ , and  $l$  is the red, green, or blue channel. For each image, a plot of H vs. S was plotted to create a 2D phasor space.

To remap Prussian blue-labeled pixels from the phasor space plot to the original image, we found the bluest location in phasor space by minimizing  $x/y$ , where  $x$  is the row index and  $y$  is the column index of the phasor plot matrix. The center of a circular mask (green circle in Figure 1B) was created from this point, and pixels enclosed by the circular mask in phasor space were remapped to a segmented image. We varied the size of the circular masks to 1) obtain segmented data for image processing, 2) calculate a ROC curve, and 3) determine the optimal mask in phasor space. Please see the *Data Quantification and Statistical Analysis* section for how the optimal circular mask was determined.

### *Deep Learning Approach*

For data preprocessing, the 100 CMH images were randomly divided into a training ( $n=80$ ) and testing ( $n=20$ ) set. The original 1000x1000 image resolution of the training set was downsampled to a resolution of 512x512 to reduce memory requirements and increase image processing speed. We applied a run-coding scheme to the 100 acquired CMH images<sup>25</sup>. To avoid overfitting, we augmented the training set by implementing random flip, rotations (90~270 degree), pixel multiplication with random values (0.8~1.5), and image blurring with Gaussian kernels (random sigma: 0~5) methods.

To identify and segment CMHs, we adopted a mask region-based convolutional neural network (R-CNN) deep learning algorithm developed by Matterport Inc.<sup>26</sup>. The Resnet-101 feature pyramid network model was employed as a backbone for the Mask R-CNN. The model was initialized using the weights obtained from the pre-training of the Microsoft common objects in context (MS-COCO)<sup>27</sup> datasets through a transfer learning scheme<sup>28</sup>. Next, the network heads were trained for 20 epochs with a learning rate of 0.005. Then, all networks were trained for 60 epochs with a learning rate of 0.001. Finally, callback was set to stop training when the validation loss did not improve for ten epochs.

To assess consistency across the training data, we randomly divided the 80 training set images into five groups ( $n=16$ /group). Five models were created ( $n=64$  images/model) by leaving one group out per model. Each model was validated on the testing set. The Dice similarity coefficient<sup>29</sup> was calculated using the testing, training, and combined data sets. Across the five different models, the Dice similarity coefficient ranged from 0.772 to 0.810 using the combined

data sets, which shows consistency across the developed models. The model with the largest Dice similarity coefficient across the combined data set was used as the final deep learning model for further analysis. A ROC curve was calculated using the optimal deep learning model by applying different threshold values to the probability of each pixel being a CMH.

### *Image Processing*

We empirically designed and performed the following image processing steps for all segmented images (ground truth, ratiometric, phasor, and deep learning). First, all segmented areas less than  $5 \mu\text{m}^2$  were removed. Next, segmented areas less than  $50 \mu\text{m}^2$  that were greater than  $50 \mu\text{m}$  from the closest segmented spot were removed. For the ratiometric and phasor approaches, segmented areas larger than  $6,000 \mu\text{m}^2$  were removed. The phasor approach filled holes with an area of less than  $1.73 \mu\text{m}^2$ , which was empirically chosen. Finally, for the ratiometric and phasor approaches, a manual check was performed by drawing a coarse circle around all Prussian blue spots to remove pixels that were not CMHs.

### *Data Quantification and Statistical Analysis*

ROC curves were calculated after the first two image processing steps. To obtain the final segmented image for each segmentation approach and subsequent data quantification, we performed one of the following steps:

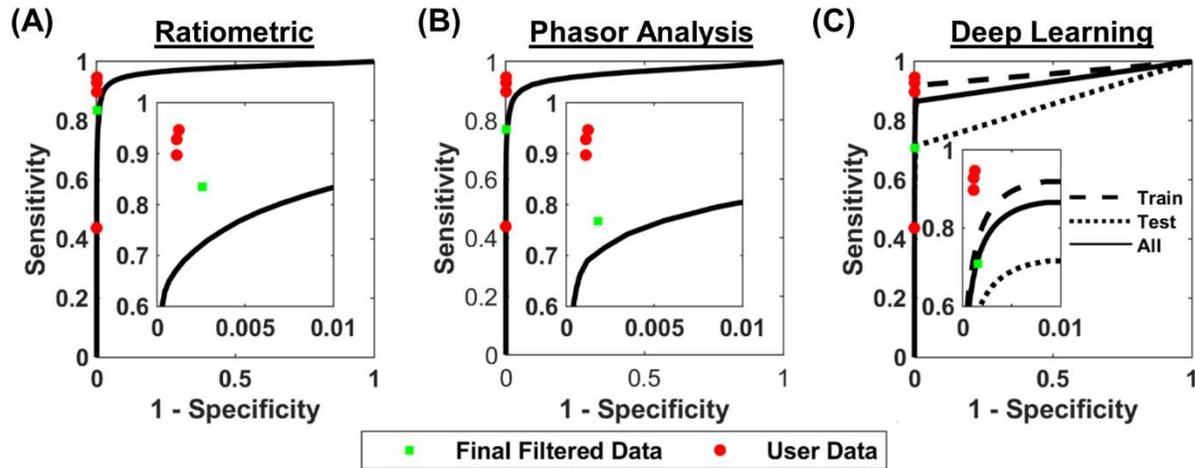
- 1) Ratiometric approach: following all image processing steps, we found the median segmentation ratio threshold across all RGB images that resulted in the minimum percent error between the ground truth area and segmented area;
- 2) Phasor approach: after implementing all image processing steps, we found the mean radius of the circular mask in phasor space that, when remapped to the original image, resulted in the minimum percent error between the ground truth area and segmented area;
- 3) Deep learning approach: the final deep learning model had the highest Dice similarity coefficient across all images.

The final sensitivity and specificity were calculated for each approach using their final segmented image. To assess the agreement between the segmented areas and ground truth area, we calculated the intraclass correlation coefficient (ICC). We used a two-way random effects model with absolute agreement to calculate the ICC<sup>30</sup>. To further assess the agreement between the segmented areas and the ground truth, we generated Bland-Altman plots with 95% confidence limits of agreement<sup>31</sup>.

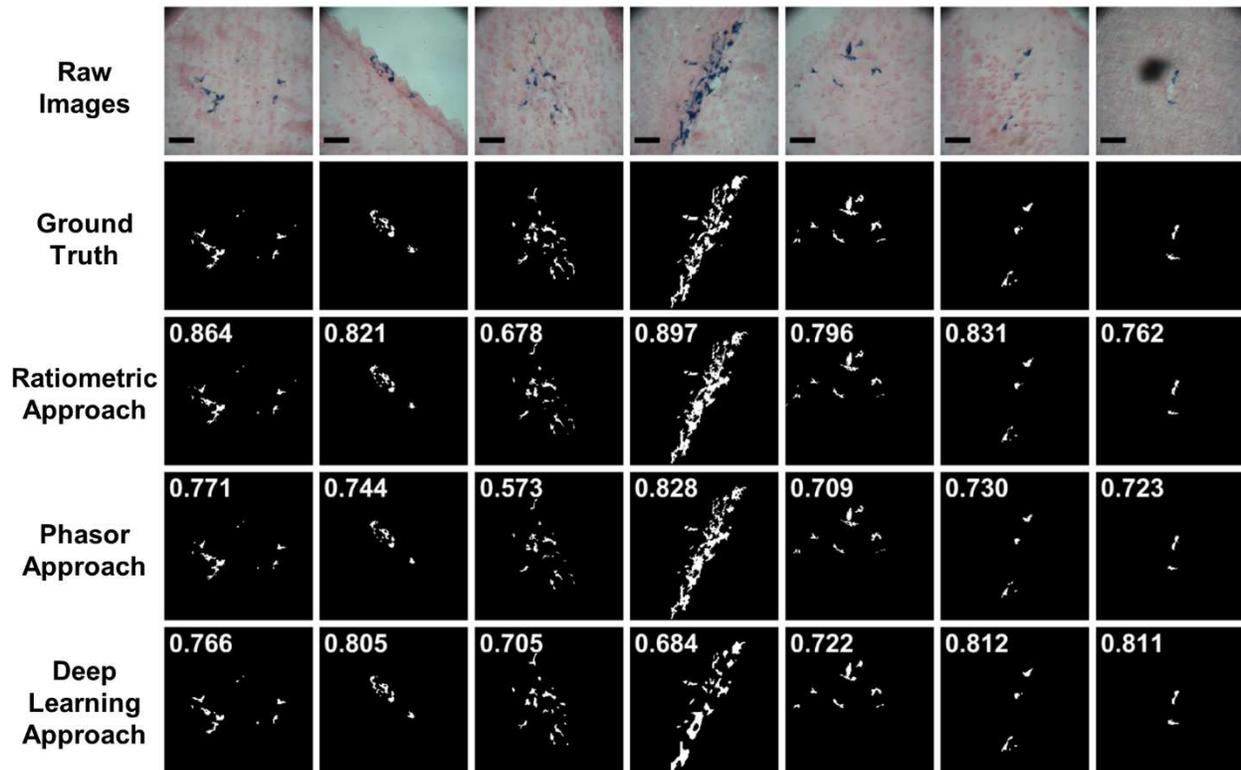
### **Results:**

*Identifying CMH pixels:* To assess the agreement between the regions of interest (ROIs) selected by the four independent users, we calculated their intraclass correlation coefficient (ICC) from the calculated CMH areas. The ICC was 0.954 (95% confidence interval: 0.897-0.976), which demonstrates excellent agreement between the four independent users<sup>30</sup>. To assess the ability of each segmentation approach to identify CMH pixels, we compared the ground truth data to the segmented CMH data from each approach. Figure 2 shows ROC curves for the (A) ratiometric, (B) phasor analysis, and (C) deep learning approaches compared to the ground truth, and the final post-processed sensitivity and specificity. The ratiometric ROC

curve has an AUC of 0.973, and a final post-processed sensitivity and specificity (green square) of 0.835 and 0.997, respectively (Figure 2A). The phasor analysis ROC curve has an AUC of 0.960 and a final post-processed sensitivity and specificity (green square) of 0.768 and 0.998, respectively (Figure 2B). Using the images that the deep learning model was tested and trained on resulted in ROC curves with AUCs of 0.958 and 0.856, respectively. With a combination of the training and test set images, the deep learning model ROC curve has an AUC of 0.932 and a final post-processed sensitivity and specificity (green square) of 0.708 and 0.998, respectively (Figure 2C). Qualitative inspection shows good agreement among seven representative RGB images, ground truth data, and segmented images from each approach (Figure 3).

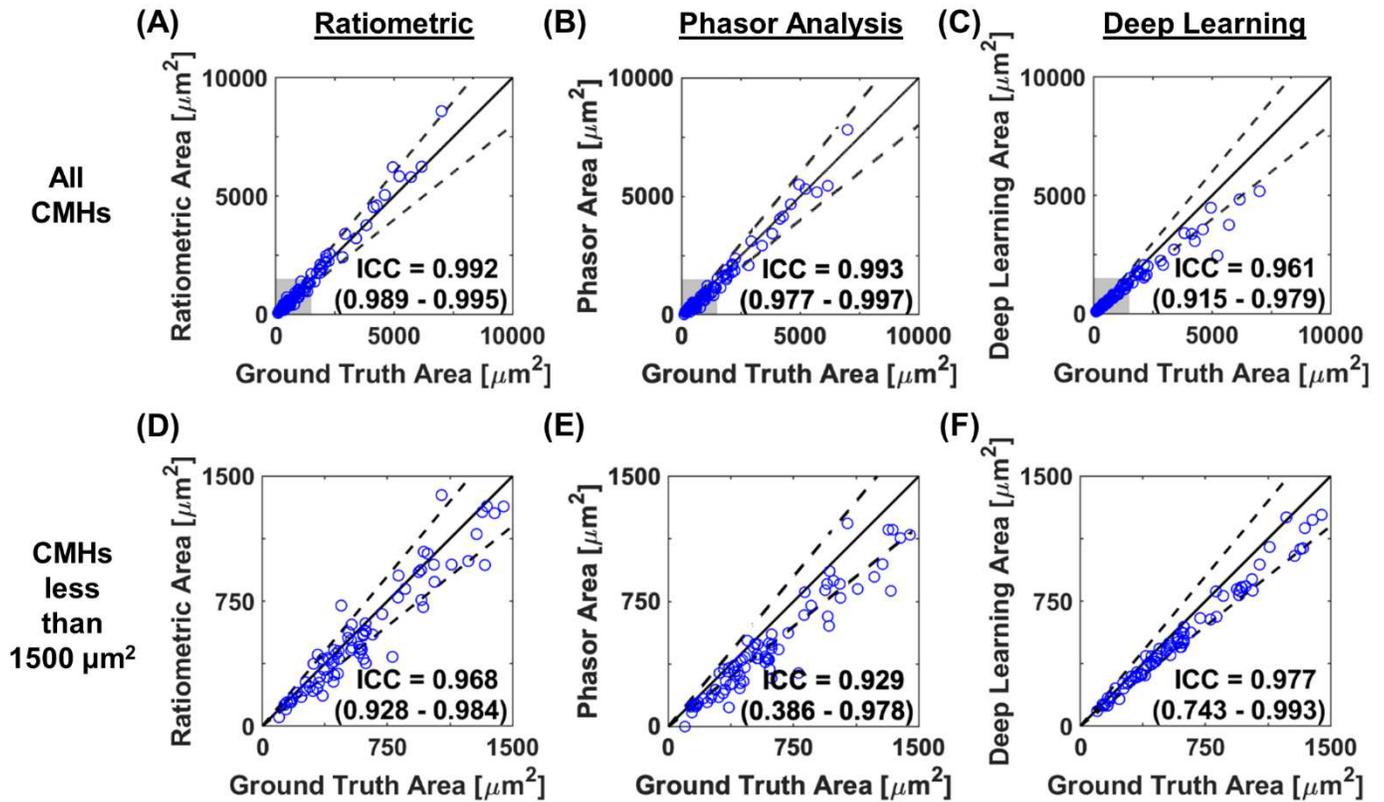


**Figure 2: Receiver operating characteristics (ROC) curve for each approach to localize CMHs.** (A) The ratiometric approach had an AUC of 0.973 and a final post-processed sensitivity and specificity of 0.835 and 0.997, respectively. (B) The phasor analysis approach had an AUC of 0.960 and a final post-processed sensitivity and specificity of 0.768 and 0.998, respectively. (C) The deep learning approach had a ROC curve for the training (dashed line), testing (dotted line), and combined (solid) data sets with AUCs of 0.958, 0.856, and 0.932, respectively. The final post-processed sensitivity and specificity were 0.708 and 0.998, respectively. The red circles represent the sensitivity and specificity for each independent user (relative to the ground truth), and the green squares represent the post-processed sensitivity and specificity for each approach.



**Figure 3: Representative examples (n=7) of raw RGB images, ground truth data, and three segmentation approaches.** The sensitivity is shown in the upper left of each image. The specificity of each image is greater than 0.990. The scale bar in the raw images is 50  $\mu\text{m}$ .

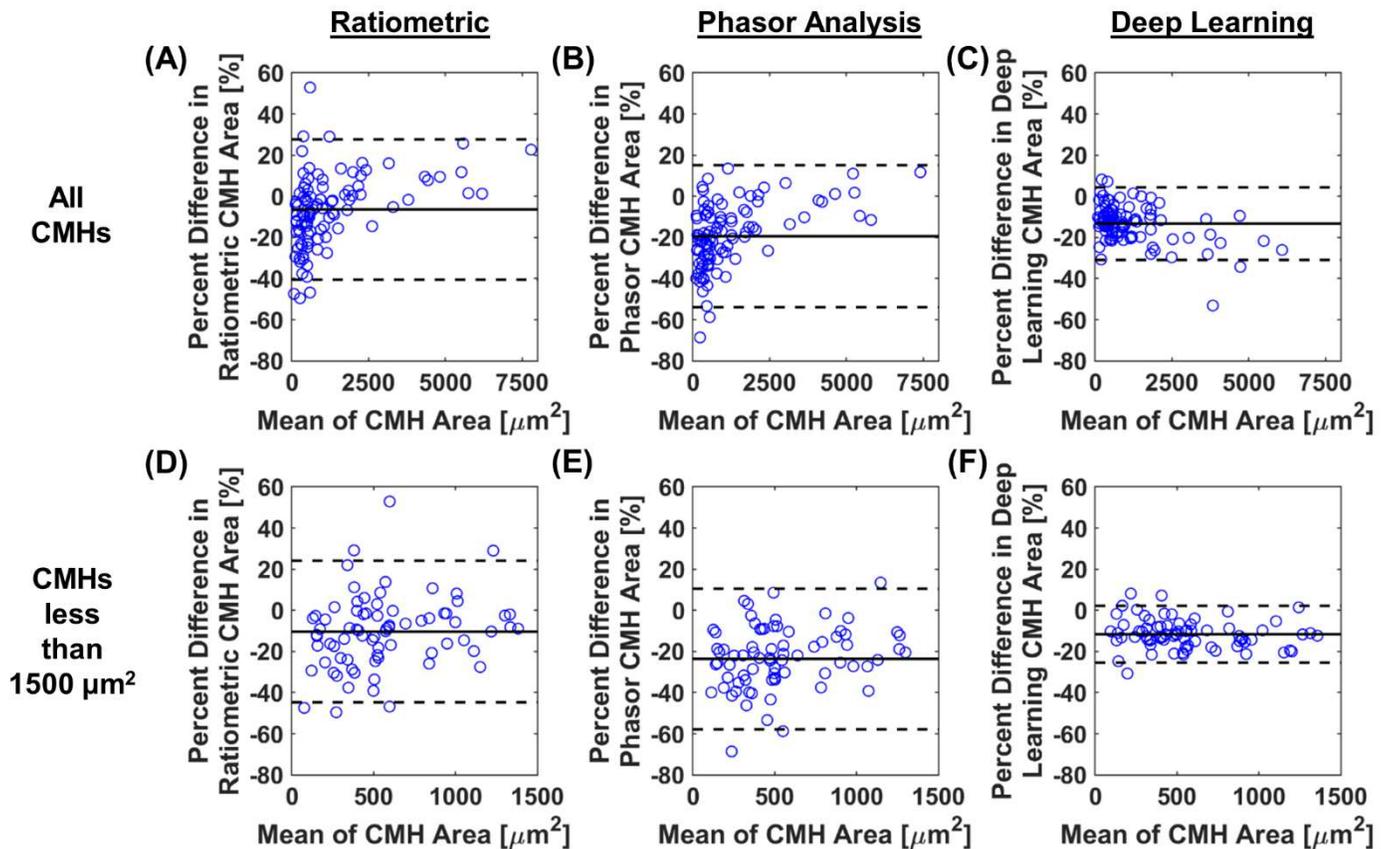
*Quantifying CMH Area:* To assess the agreement of the CMH area between each segmentation approach and the ground truth, we calculated the ICC. We examined the ICC using all CMH areas (n=100), and CMH areas less than 1,500  $\mu\text{m}^2$  (n=75). Using all calculated areas, the ICCs from the ratiometric, phasor, and deep learning approaches were 0.992 (95% confidence interval: 0.989-0.995), 0.993 (95% confidence interval: 0.977-0.997), and 0.961 (95% confidence interval: 0.915-0.979), respectively (Figure 4A-C). Using areas less than 1,500  $\mu\text{m}^2$ , the ICCs from the ratiometric, phasor, and deep learning approaches were 0.968 (95% confidence interval: 0.928-0.984), 0.929 (95% confidence interval: 0.386-0.978), 0.977 (95% confidence interval: 0.743-0.993), respectively (Figure 4D-F). These results demonstrate excellent agreement between each developed approach and the ground truth (majority voting data). However, the ratiometric approach had the most consistent results compared to the phasor and deep learning approaches based on the 95% confidence interval.



**Figure 4: CMH area quantification between segmentation approaches and ground truth data with intraclass correlation coefficient (ICC).** (A-C) Comparison between the ground truth and the segmentation approaches (A) ratiometric, (B) phasor analysis, and (C) deep learning for all calculated areas. (D-E) Magnified data of the shaded gray box shown in A-C. Comparison between the ground truth and the segmentation approaches (D) ratiometric, (E) phasor analysis, and (F) deep learning for areas less than 1,500  $\mu\text{m}^2$ . Intraclass correlation coefficients (ICC) and 95% confidence intervals are located in the lower right of each figure. In each graph (A-F), the solid black line is unity, and the dashed lines are  $\pm 20\%$  error.

To further assess the agreement between the segmented areas from each approach and the ground truth, we generated Bland-Altman plots. We examined the Bland-Altman plots using all CMH areas ( $n=100$ ), and CMH areas less than 1,500  $\mu\text{m}^2$  ( $n=75$ ). For each graph in Figure 5, the y-axis is the percent difference between the areas from the selected segmentation approach (ratiometric, phasor, or deep learning) and the ground truth. The x-axis is the mean CMH area between the selected segmentation approach and ground truth. Each segmentation approach underestimates the calculated areas (negative percent difference), as is shown by the horizontal solid black line. This underestimation could be attributed to the users from the manual method not drawing on the exact border of the CMH, but around the CMH. The mean difference from the ratiometric approach (-6.4% and -10.3% for all CMH areas and less than 1,500  $\mu\text{m}^2$ , respectively) was the closest to zero compared to the phasor (-19.4% and -23.6% for all CMH areas and less than 1,500  $\mu\text{m}^2$ , respectively) and deep learning (-13.3% and -11.6% for all CMH areas and less than 1,500  $\mu\text{m}^2$ , respectively) approaches. The proximity to zero suggests that the ratiometric approach may be most closely tied to the average CMH area. However, the deep

learning approach had the smallest range for the upper and lower limits of agreement (35.2% and 27.6% for all CMH areas and less than 1,500  $\mu\text{m}^2$ , respectively) compared to the ratiometric (68.1% and 68.8% for all CMH areas and less than 1,500  $\mu\text{m}^2$ , respectively) and phasor (69.1% and 68.3% for all CMH areas and less than 1,500  $\mu\text{m}^2$ , respectively) approaches. These results collectively suggest that the deep learning approach had the highest precision among the three developed approaches.



**Figure 5: CMH area quantification between the segmentation approaches and ground truth data with the Bland-Altman plot.** (A-C) Comparison between the ground truth and the segmentation approaches (A) ratiometric, (B) phasor analysis, and (C) deep learning for all calculated areas. (D-E) Comparison between the ground truth and the segmentation approaches (D) ratiometric, (E) phasor analysis, and (F) deep learning for areas less than 1,500  $\mu\text{m}^2$ . In each graph (A-F), the y-axis is the percent difference between the areas from the selected segmentation approach (ratiometric, phasor, or deep learning) and the ground truth, and the x-axis is the mean area between the selected segmentation approach and ground truth. The solid black line is the mean difference and the dashed black lines are the 95% upper and lower limits of agreement.

## **Discussion:**

Histological sections remain the gold standard for assessing pathology from *ex vivo* tissue samples<sup>14,32</sup>. Clinical pathologists and researchers rely on data from these samples to guide patient management and gain mechanistic insights. However, tissue processing is laborious, requires extensive expertise, and is subjective<sup>14</sup>. Furthermore, gaining quantitative data from these samples is time exhaustive and subjective. Specifically, there are typically four steps required to characterize Prussian blue-stained CMHs. First, the brain is sectioned and stained with Prussian blue. Second, an individual visually inspects brain sections to identify and count CMHs with a brightfield microscope. Third, images of the identified CMHs are digitally acquired. Finally, a user manually outlines and adjusts RGB thresholds<sup>13</sup> to quantify Prussian blue positivity or CMH area<sup>20</sup>. Each step is prone to error, variability, and subjectivity, and demands extensive personnel time.

In the present study, we set out to develop a digital processing approach to decrease the subjectivity and time involved in CMH quantitation. We developed and compared three approaches to identify and quantify CMHs stained with Prussian blue: 1) ratiometric, 2) phasor analysis, and 3) deep learning (Figure 1). Our ratiometric approach used a combination of red, green, and blue pixel intensity values. Automated histological studies using different wavelengths of light have primarily been applied in cancer applications<sup>33,34</sup>. One breast cancer study used multispectral (16-wavelengths) and RGB data to spectrally unmix breast cancer markers<sup>33</sup>. The multispectral approach performed better than the RGB approach but is substantially more costly as it uses an expensive camera with an embedded liquid crystal tunable filter. A major advantage of our approach is that it can use the simple RGB values to study CMHs, which decreases the complexity and cost of the imaging system. Our second approach, the phasor analysis method, transforms the RGB image into phasor space through Fourier analysis. Phasor analysis in biomedical optics was first pioneered using FLIM data to identify fluorophores in images<sup>23</sup>. A proof-of-concept study applied spectral phasor analysis to brightfield, RGB data from a lymph node with metastatic cancer<sup>18</sup>. We leveraged off these advances in spectral phasor analysis of brightfield microscopy to separate Prussian blue-stained CMH pixels from background pixels. Although the phasor approach was the worst-performing approach, it may have advantages with new data sets from different microscopes, as it is based on global image analysis and requires limited or no reference information<sup>17</sup>. Our last approach, which uses deep learning, applies a mask region-based convolutional neural network (R-CNN). Many digital pathology approaches have used mask R-CNNs in applications of nuclei segmentation<sup>12</sup> and cancer<sup>35-37</sup>. Similar to our study, these studies spatially identified specific features (e.g., nuclei). Our study is the first to apply a mask R-CNN to identify and quantify CMHs in histological images, where it performed well, similar to prior reports in other applications.

Our data demonstrate that each approach performs well at identifying CMHs and quantifying CMH areas (Figures 2-4). A summary of the different approaches is shown in Table 1. The deep learning and ratiometric approaches performed better than the phasor analysis approach. The deep learning approach was the most precise method but had less versatility as it is not easily

modifiable. This approach may require a lengthy re-training process when quantifying CMHs from different microscopes and image sets. The ratiometric approach had less precision than the deep learning approach. However, this approach accurately quantifies the CMH area and provides the user with more versatility to alter the threshold robustly. Therefore, depending on the needs of the user and the type of data set, the deep learning or ratiometric approach will be the optimal method to detect and quantify CMHs. Our data suggest that implementing these methods to analyze CMH data can drastically increase the processing speed while maintaining precision and accuracy.

**Table 1: Comparison between manual method and three segmentation approaches for quantitation of cerebral microhemorrhages.**

	Analysis Process	Average time commitment per image	Modifiable	Sensitivity, Specificity	Intraclass correlation coefficient (95% confidence interval)	Absolute area difference ( $\mu\text{m}^2$ ) (interquartile range)	Percent area difference (%) (interquartile range)
Manual Approach	Subjective	10 minutes					
Ratiometric Approach	Semi-automated	15 seconds	Easy	0.835, 0.997	0.992 (0.989-0.995)	75.2 (33.3-172.3)	11.0 (4.7-21.2)
Phasor Approach	Semi-automated	30 seconds	Moderate	0.768, 0.998	0.993 (0.977 -0.997)	124.5 (55.2-255.2)	18.8 (9.3-29.4)
Deep Learning Approach	Automated	3 seconds	Difficult	0.708, 0.998	0.961 (0.915-0.979)	71.1 (33.6-167.7)	12.7 (8.8-18.9)

Our study does contain limitations that can be addressed with future work. Although the deep learning approach had the lowest sensitivity and ICC, the training data set was relatively small. Increasing the data set size could improve the deep learning results, especially for the larger sized CMHs, where the deep learning approach underestimated the CMH area. Second, the CMHs in our study were induced with LPS. Further testing of these approaches in other CMH models, such as chronic kidney disease<sup>9</sup>, hypertension<sup>7</sup>, and cerebral amyloid angiopathy<sup>38-40</sup>, in addition to human brain samples, are needed. Third, the algorithms have not been tested on different microscopes to assess if a calibration factor is needed to account for differences in camera spectral sensitivity and excitation spectrum. Future work should test and assess the ability of the algorithms on different microscopes and with the use of multiple wavelengths of laser light on optically-cleared thick tissue sections to obtain three-dimensional (3D) information<sup>41,42</sup>. Fourth, the brain sections were visually scanned by users to find and acquire a digital image for each CMH. Using a slide scanner in conjunction with one of the developed approaches could drastically decrease the personnel time required to identify and quantify CMHs. Fifth, although Prussian blue is commonly used to detect CMHs, it does not detect newly formed CMHs, nor distinguish between hemoglobin-derived (hemosiderin) and nonhemoglobin-derived iron<sup>13</sup>. Combining our Prussian blue-developed approaches with future algorithms that

use fresh CMHs stained with hematoxylin and eosin (H&E)<sup>43</sup> may help detect non-acute and acute CMHs. Finally, in this work, we only focused on the analysis of CMH images labeled with exogenous Prussian blue. Future work that virtually stains samples through machine learning of autofluorescent and absorption features may minimize inter-batch variability due to staining, which could revolutionize the histological pipeline for examining CMHs by removing the manual tissue staining step<sup>14,44,45</sup>.

In conclusion, we developed and compared three segmentation approaches: 1) ratiometric, 2) phasor analysis, and 3) deep learning to identify and quantify CMHs stained with Prussian blue. The developed approaches can substantially reduce the considerable personnel and time commitment required to quantify CMHs in individual histological sections. The deep learning and ratiometric approaches performed better than the phasor analysis approach. We found that the deep learning approach had the most precision; however, the ratiometric approach had good accuracy and the most versatility, but less precision than the deep learning approach. Our data suggest that implementing these methods to analyze CMH images can drastically increase the processing speed, while maintaining precision and accuracy.

### **Acknowledgements**

We thank all members of the Choi, Fisher, Cribbs, Lau and Kim labs for their helpful discussions on this project. This work was supported in part by the Arnold and Mabel Beckman Foundation, the National Institutes of Health (5TL1TR001415-06 (to C.C.), P41EB015890, R01NS020989, R21AG066000, and R21NS111984) and by Leading Foreign Research Institute Recruitment Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (MSIT) (NRF-2018K1A4A3A02060572).

### **Author Contributions**

C.C. and B.C. developed the project concept, experimental design, and proposed automated algorithms. K.T.P. and C.A. collected data. K.T.P., C.E.D., D.F.X., and C.C. manually selected CMH ROIs. C.C., G.J., S.K., and R.H.C., developed the automated algorithms and wrote the main body of the manuscript. C.F., A.C.F.N., W.L.L., D.H.C., and M.F. provided expertise on CMHs and informed development of image processing steps. All authors reviewed the manuscript.

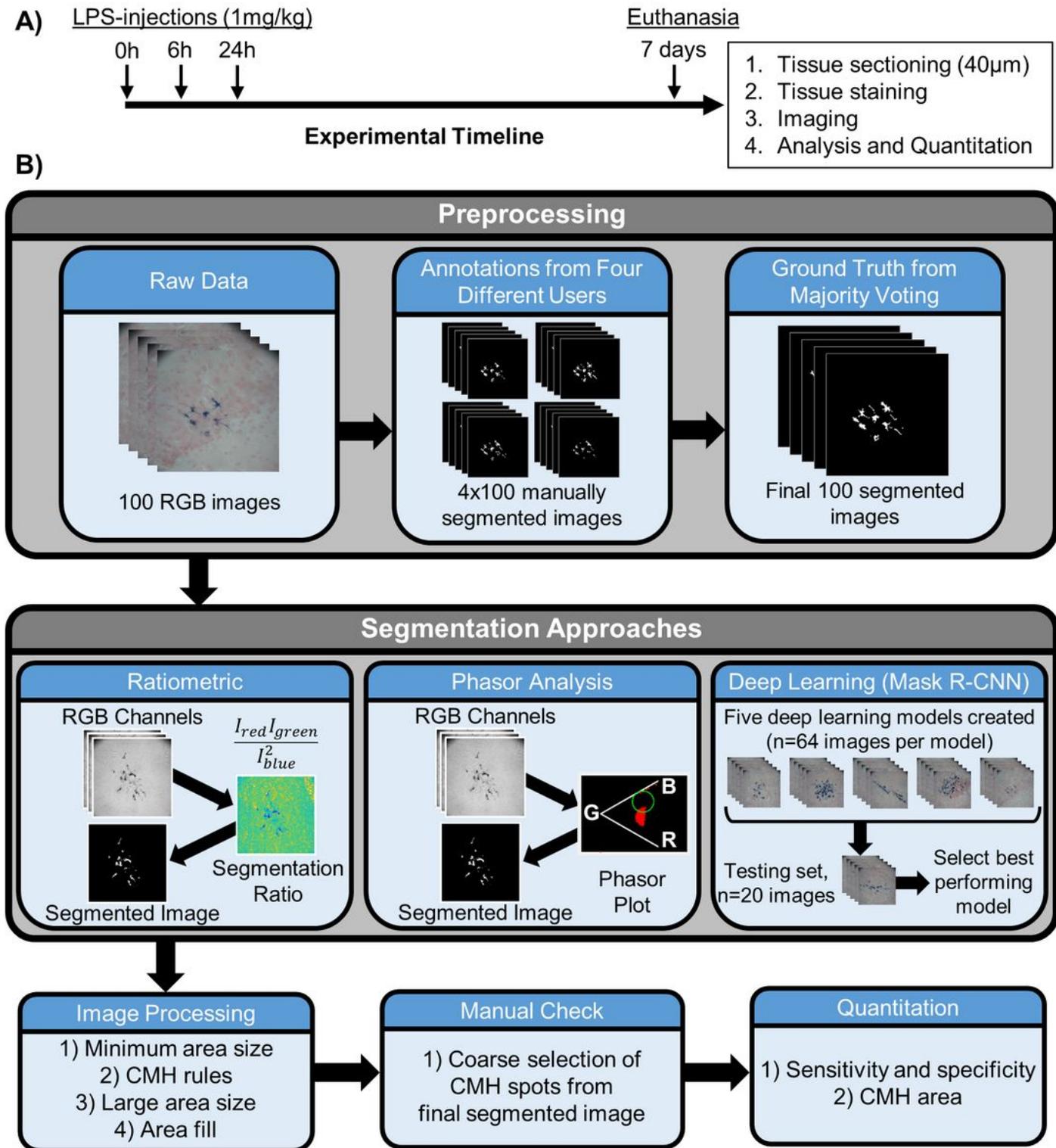
### **References**

1. Werring, D. J. *et al.* Cognitive dysfunction in patients with cerebral microbleeds on T2\*-weighted gradient-echo MRI. *Brain* **127**, 2265–2275 (2004).
2. Bokura, H. *et al.* Microbleeds are associated with subsequent hemorrhagic and ischemic stroke in healthy elderly individuals. *Stroke* **42**, 1867–1871 (2011).
3. Poels, M. M. F. *et al.* Prevalence and risk factors of cerebral microbleeds: An update of the rotterdam scan study. *Stroke* **41**, (2010).
4. Vernooij, M. W. *et al.* Prevalence and risk factors of cerebral microbleeds: The Rotterdam Scan Study. *Neurology* **70**, 1208–1214 (2008).
5. Fisher, M., French, S., Ji, P. & Kim, R. C. Cerebral Microbleeds in the Elderly A Pathological Analysis. *Stroke* **41**, 2782–2785 (2010).
6. Sumbria, R. K. *et al.* Aging exacerbates development of cerebral microbleeds in a mouse model. *J. Neuroinflammation* **15**, 1–7 (2018).

7. Vasilevko, V. *et al.* Aging and cerebrovascular dysfunction: Contribution of hypertension, cerebral amyloid angiopathy, and immunotherapy. *Ann. N. Y. Acad. Sci.* **1207**, 58–70 (2010).
8. Fisher, M. *et al.* Therapeutic modulation of cerebral microhemorrhage in a mouse model of cerebral amyloid angiopathy. *Stroke* **42**, 3300–3303 (2011).
9. Lau, W. L. *et al.* Chronic Kidney Disease Increases Cerebral Microbleeds in Mouse and Man. *Transl. Stroke Res.* **11**, 122–134 (2020).
10. Nyúl-tóth, Á. *et al.* Increases in hypertension-induced cerebral microhemorrhages exacerbate gait dysfunction in a mouse model of Alzheimer's disease. *GeroScience* **42**, 1685–1698 (2020).
11. Tarantini, S. *et al.* Insulin-like growth factor 1 deficiency exacerbates hypertension-induced cerebral microhemorrhages in mice, mimicking the aging phenotype. *Aging Cell* **16**, 469–479 (2017).
12. Lee, J. S. *et al.* Cerebral Microbleeds, Hypertension, and Intracerebral Hemorrhage in Cerebral Autosomal-Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy. *Front. Neurol.* **8**, 1–6 (2017).
13. Wadi, L. C. *et al.* Mechanisms of Cerebral Microbleeds. *J. Neuropathol. Exp. Neurol.* **79**, 1–7 (2020).
14. Rivenson, Y. *et al.* Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nat. Biomed. Eng.* **3**, 466–477 (2019).
15. Aeffner, F. *et al.* Introduction to Digital Image Analysis in Whole-slide Imaging: A White Paper from the Digital Pathology Association. *J. Pathol. Inform.* **10**, (2019).
16. Horai, Y., Kakimoto, T., Takemoto, K. & Tanaka, M. Quantitative analysis of histopathological findings using image processing software. *J. Toxicol. Pathol.* **30**, 351–358 (2017).
17. Fereidouni, F., Bader, A. N. & Gerritsen, H. C. Spectral phasor analysis allows rapid and reliable unmixing of fluorescence microscopy spectral images. *Opt. Express* **20**, 12729 (2012).
18. Fereidouni, F., Grif, C., Todd, A. & Levenson, R. Multispectral analysis tools can increase utility of RGB color images in histology. *J. Opt.* **20**, (2018).
19. Jiao, C., Su, K., Xie, W. & Ye, Z. Burn image segmentation based on Mask Regions with Convolutional Neural Network deep learning framework: more accurate and more convenient. *Burn. Trauma* **7**, 1–14 (2019).
20. Sumbria, R. K. *et al.* A murine model of inflammation-induced cerebral microbleeds. *J. Neuroinflammation* **13**, 1–12 (2016).
21. Loren, M., Crouzet, C., Bahani, A., Vasilevko, V. & Choi, B. Optical clearing potential of immersion-based agents applied to thick mouse brain sections. *PLoS One* **14**, 1–14 (2019).
22. Hanley, J. A. & McNeil, B. J. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology* **143**, 29–36 (1982).
23. Digman, M. A., Caiolfa, V. R., Zamai, M. & Gratton, E. The Phasor Approach to Fluorescence Lifetime Imaging Analysis. *Biophys. J.* **94**, L14–L16 (2008).
24. Stringari, C. *et al.* Phasor approach to fluorescence lifetime microscopy distinguishes different metabolic states of germ cells in a live tissue. *Proc. Natl. Acad. Sci.* **108**, 13582–13587 (2011).
25. Shapiro, L. & Stockman, G. *Computer Vision*. Pearson (2001).
26. Abdulla, W. Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow. (2017).
27. Lin, T.-Y. *et al.* Microsoft COCO: Common Objects in Context. *Eccv* 740–755 (2014).
28. Yosinski, J., Clune, J., Bengio, Y. & Lipson, H. How transferable are features in deep neural networks? *Adv. Neural Inf. Process. Syst.* **4**, 3320–3328 (2014).

29. Zou, K. H. *et al.* Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index. *Acad. Radiol.* **11**, 178–189 (2004).
30. Koo, T. K. & Li, M. Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **15**, 155–163 (2016).
31. Giavarina, D. Understanding Bland Altman analysis. *Biochem. Medica* **25**, 141–151 (2015).
32. Abbasi, S. *et al.* All-optical Reflection-mode Microscopic Histology of Unstained Human Tissues. *Sci. Rep.* **9**, 1–11 (2019).
33. Liu, W., Wang, L., Chen, J., Yuan, J. & Xiang, Q. Application of multispectral imaging in quantitative immunohistochemistry study of breast cancer : a comparative study. *Tumor Biol.* 5013–5024 (2016). doi:10.1007/s13277-015-4327-9
34. Hossain, A. *et al.* Automated approach for estimation of grade groups for prostate cancer based on histological image feature analysis. *Prostate* **80**, 291–302 (2020).
35. Gertych, A. *et al.* Convolutional neural networks can accurately distinguish four histologic growth patterns of lung adenocarcinoma in digital slides. *Sci. Rep.* **9**, 1–12 (2019).
36. Li, W. *et al.* Path R-CNN for Prostate Cancer Diagnosis and Gleason Grading of Histological Images. *IEEE Trans. Med. Imaging* **38**, 945–954 (2019).
37. Kurc, T., Bakas, S., Ren, X., Bagari, A. & Momeni, A. Segmentation and Classification in Digital Pathology for Glioma Research: Challenges and Deep Learning Approaches. *Front. Neurosci.* **14**, 1–15 (2020).
38. Petrushina, I. *et al.* Neurobiology of Disease Characterization and preclinical evaluation of the cGMP grade DNA based vaccine, AV-1959D to enter the first-in-human clinical trials. *Neurobiol. Dis.* **139**, 104823 (2020).
39. Veluw, S. J. Van *et al.* In vivo characterization of spontaneous microhemorrhage formation in mice with cerebral amyloid angiopathy. *J. Cereb. Blood Flow Metab.* **41**, 82–91 (2021).
40. Veluw, S. J. Van *et al.* Reduced vascular amyloid burden at microhemorrhage sites in cerebral amyloid angiopathy. *Acta Neuropathol.* **133**, 409–415 (2017).
41. Lo, P., Crouzet, C., Vasilevko, V. & Choi, B. Visualization of microbleeds with optical histology in mouse model of cerebral amyloid angiopathy. *Microvasc. Res.* **105**, (2016).
42. Glaser, A. K. *et al.* Multi-immersion open-top light-sheet microscope for high-throughput imaging of cleared tissues. *Nat. Commun.* **10**, 1–8 (2019).
43. Liu, S. *et al.* Comparative Analysis of H&E and Prussian Blue Staining in a Mouse Model of Cerebral Microbleeds. *J. Histochem. Cytochem.* **62**, 767–773 (2014).
44. Qorbani, A. *et al.* Microscopy with ultraviolet surface excitation (MUSE): A novel approach to real-time inexpensive slide-free dermatopathology. *J. Cutan. Pathol.* 498–503 (2018). doi:10.1111/cup.13255
45. Rivenson, Y. *et al.* PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light Sci. Appl.* **8**, (2019).

# Figures



**Figure 1**

Overarching experimental timeline and data processing scheme. (A) Experimental timeline to induce cerebral microhemorrhages (CMHs), followed by tissue sectioning, staining, imaging, analysis, and quantitation. (B) Data preprocessing steps include: 1) acquiring raw color images of each CMH, 2)

manually annotating color images by four users, and 3) determining the ground truth through majority voting. Three segmentation approaches (ratiometric, phasor analysis, and deep learning) were developed and compared. After the segmentation approaches, each set of results went through a series of image processing, data quality check, and quantitation steps (see the Image Processing section for details regarding each approach).

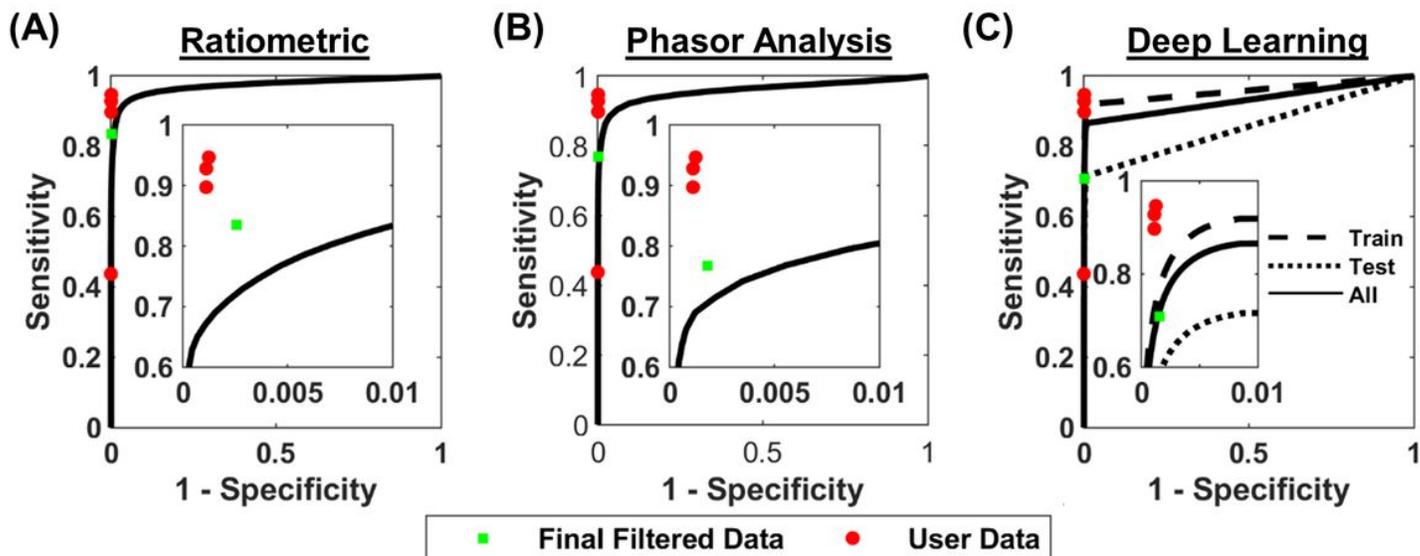
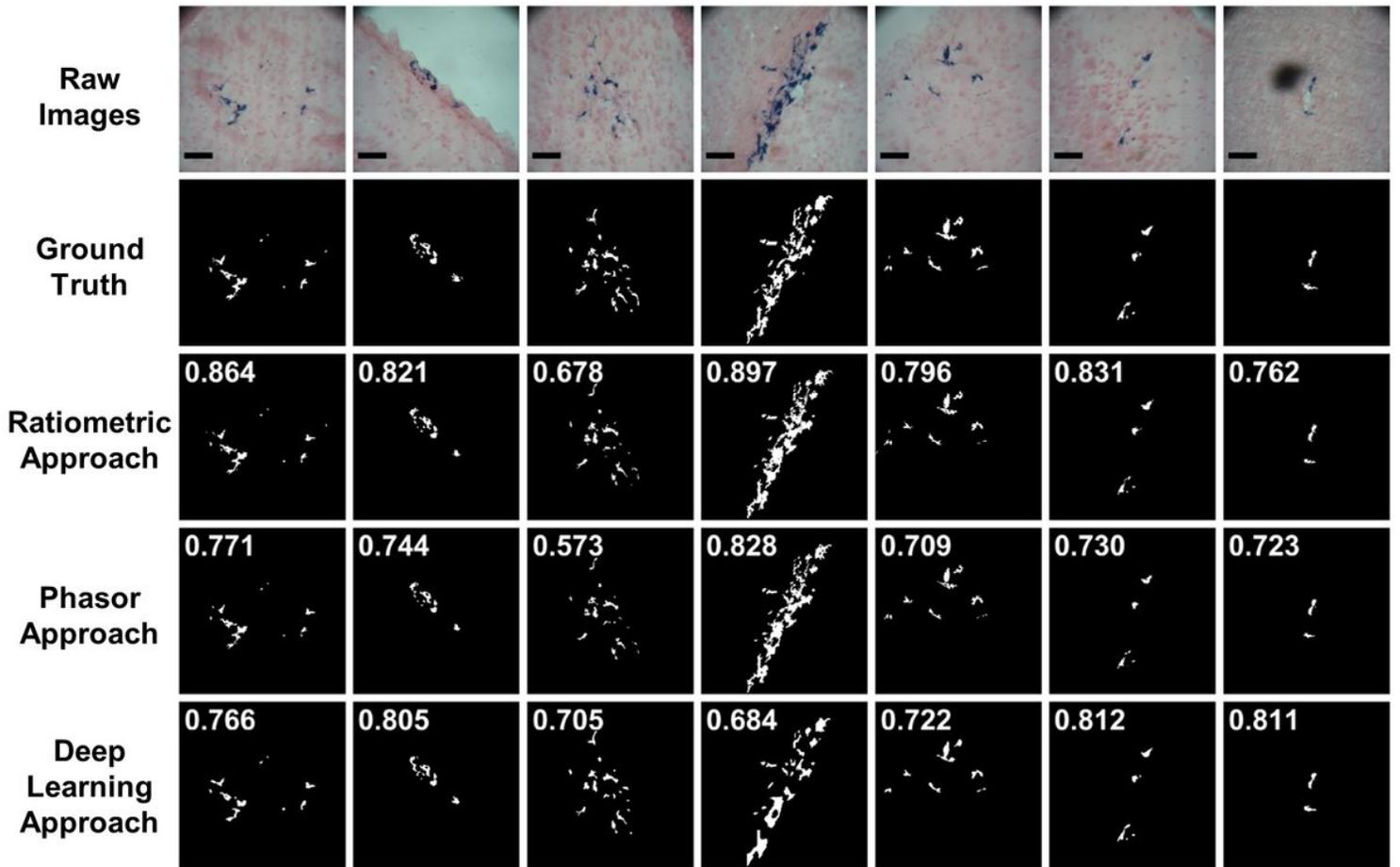


Figure 2

Receiver operating characteristics (ROC) curve for each approach to localize CMHs. (A) The ratiometric approach had an AUC of 0.973 and a final post-processed sensitivity and specificity of 0.835 and 0.997, respectively. (B) The phasor analysis approach had an AUC of 0.960 and a final post-processed sensitivity and specificity of 0.768 and 0.998, respectively. (C) The deep learning approach had a ROC curve for the training (dashed line), testing (dotted line), and combined (solid) data sets with AUCs of 0.958, 0.856, and 0.932, respectively. The final post-processed sensitivity and specificity were 0.708 and 0.998, respectively. The red circles represent the sensitivity and specificity for each independent user (relative to the ground truth), and the green squares represent the post-processed sensitivity and specificity for each approach.



**Figure 3**

Representative examples (n=7) of raw RGB images, ground truth data, and three segmentation approaches. The sensitivity is shown in the upper left of each image. The specificity of each image is greater than 0.990. The scale bar in the raw images is 50  $\mu\text{m}$ .

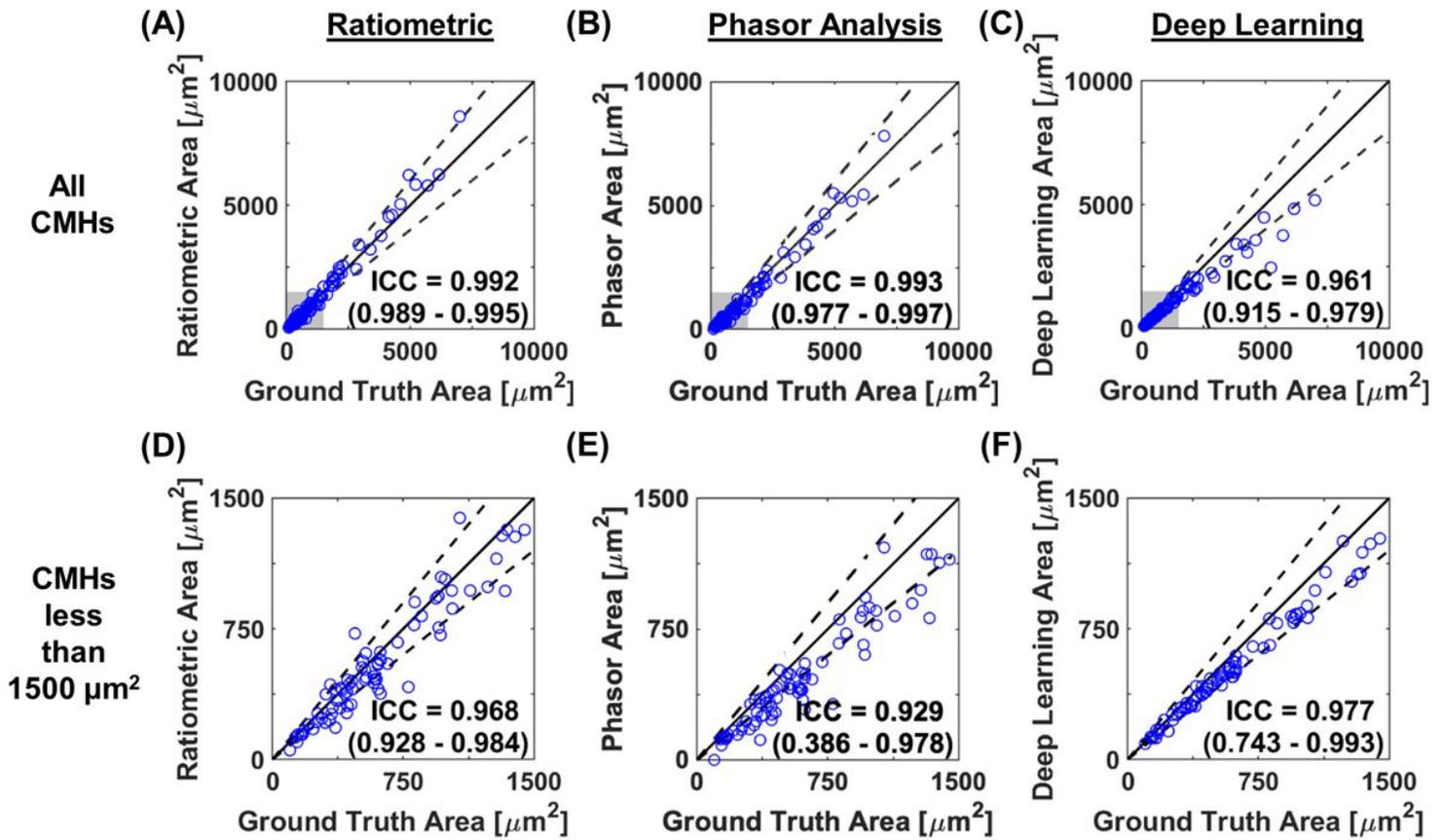
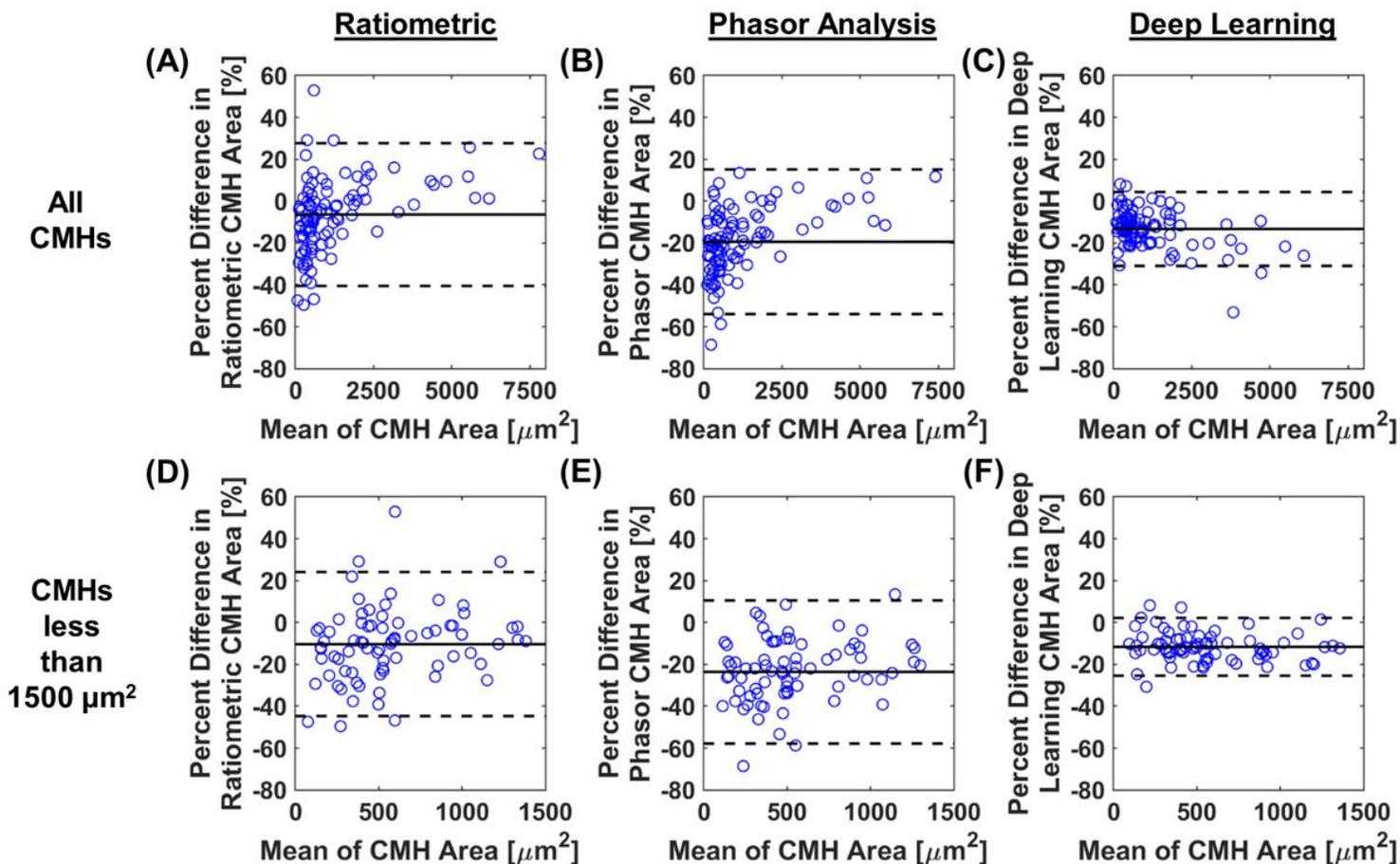


Figure 4

CMH area quantification between segmentation approaches and ground truth data with intraclass correlation coefficient (ICC). (A-C) Comparison between the ground truth and the segmentation approaches (A) ratiometric, (B) phasor analysis, and (C) deep learning for all calculated areas. (D-E) Magnified data of the shaded gray box shown in A-C. Comparison between the ground truth and the segmentation approaches (D) ratiometric, (E) phasor analysis, and (F) deep learning for areas less than  $1,500 \mu\text{m}^2$ . Intraclass correlation coefficients (ICC) and 95% confidence intervals are located in the lower right of each figure. In each graph (A-F), the solid black line is unity, and the dashed lines are  $\pm 20\%$  error.



**Figure 5**

CMH area quantification between the segmentation approaches and ground truth data with the Bland-Altman plot. (A-C) Comparison between the ground truth and the segmentation approaches (A) ratiometric, (B) phasor analysis, and (C) deep learning for all calculated areas. (D-E) Comparison between the ground truth and the segmentation approaches (D) ratiometric, (E) phasor analysis, and (F) deep learning for areas less than 1,500  $\mu\text{m}^2$ . In each graph (A-F), the y-axis is the percent difference between the areas from the selected segmentation approach (ratiometric, phasor, or deep learning) and the ground truth, and the x-axis is the mean area between the selected segmentation approach and ground truth. The solid black line is the mean difference and the dashed black lines are the 95% upper and lower limits of agreement.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [20210105091255CrouzetSciRepmanuscriptSupplementarymaterial.pdf](#)