

RESEARCH ARTICLE OPEN ACCESS

An Optimized Framework of QSM Mask Generation Using Deep Learning: QSMmask-Net

Gawon Lee¹  | Woojin Jung² | Ken E. Sakaie³ | Se-Hong Oh^{1,3}

¹Department of Biomedical Engineering, Hankuk University of Foreign Studies, Yongin, Republic of Korea | ²AIRS Medical, Seoul, Republic of Korea | ³Department of Diagnostic Radiology, Diagnostics Institute, The Cleveland Clinic, Cleveland, Ohio, USA

Correspondence: Se-Hong Oh (ohs@hufs.ac.kr)

Received: 8 December 2024 | **Revised:** 10 April 2025 | **Accepted:** 28 April 2025

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (RS-2023-NR076864) and the Hankuk University of Foreign Studies Research Fund.

Keywords: brain mask for QSM | brain QSM | deep learning-based QSM processing | deep neural networks | magnetic susceptibility mapping | MRI | QSM | segmentation

ABSTRACT

Quantitative susceptibility mapping (QSM) provides the spatial distribution of magnetic susceptibility within tissues through sequential steps: phase unwrapping and echo combination, mask generation, background field removal, and dipole inversion. Accurate mask generation is crucial, as masks excluding regions outside the brain and without holes are necessary to minimize errors and streaking artifacts during QSM reconstruction. Variations in susceptibility values can arise from different mask generation methods, highlighting the importance of optimizing mask creation. In this study, we propose QSMmask-net, a deep neural network-based method for generating precise QSM masks. QSMmask-net achieved the highest Dice score compared to other mask generation methods. Mean susceptibility values using QSMmask-net masks showed the lowest differences from manual masks (ground truth) in simulations and healthy controls (no significant difference, $p > 0.05$). Linear regression analysis confirmed a strong correlation with manual masks for hemorrhagic lesions (slope = 0.9814 ± 0.007 , intercept = 0.0031 ± 0.001 , $R^2 = 0.9992$, $p < 0.05$). We have demonstrated that mask generation methods can affect the susceptibility value estimations. QSMmask-net reduces the labor required for mask generation while providing mask quality comparable to manual methods. The proposed method enables users without specialized expertise to create optimized masks, potentially broadening QSM applicability efficiently.

1 | Introduction

Quantitative susceptibility mapping (QSM) is an advanced magnetic resonance imaging (MRI) technique that provides a quantitative spatial distribution of magnetic susceptibility value. Iron and myelin are the primary sources of QSM tissue contrast in the human brain. Iron exhibits paramagnetic susceptibility, resulting in positive susceptibility values, whereas myelin lipids

possess diamagnetic susceptibility, causing negative values in myelinated white matter.

QSM has become promising for clinical applications, including Alzheimer's disease, multiple sclerosis, Parkinson's disease, and microhemorrhagic disease [1–8]. In Alzheimer's disease, QSM reliably indicates cognitive decline [9], demonstrating increased susceptibility in the hippocampus, temporal, and frontal lobes.

Abbreviations: COSMOS, calculation of susceptibility through multiple orientation sampling; iLSQR, iterative LSQR; LSQR, least-squares; STI, susceptibility tensor imaging; V-SHARP, variable-kernel SHARP.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](#), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *NMR in Biomedicine* published by John Wiley & Sons Ltd.

In multiple sclerosis, QSM detects higher paramagnetic susceptibility in the basal ganglia compared to controls [10, 11]. Similarly, QSM sensitively identifies increased iron content in the substantia nigra pars compacta in Parkinson's disease [12]. Furthermore, QSM has been found to perform better than conventional MRI in identifying hemorrhagic lesions [13, 14] and shows high sensitivity to blood–brain barrier disruptions [15, 16].

Although QSM provides valuable information regarding iron and myelin deposition, broader application faces challenges, particularly in reliably compensating for background fields [17–20]. This compensation is contingent on the use of a precise and reliable tissue mask [21, 22], but generating an accurate mask without laborious manual intervention is challenging. Stewart et al. [23] proposed a basic brain segmentation method utilizing two-step process including thresholding and hole filling. Although this method provides a robust segmentation method, there remains a need for an automated method capable of more precisely removing background fields. Deep neural networks have been proposed as a potential solution for segmenting Alzheimer's disease-related regions of interest (ROIs) [24], tumors [25], and white matter hyperintensities [26], but existing algorithms are not specifically optimized for QSM mask generation.

In this study, we developed and evaluated QSMmask-net, a deep learning-based approach for mask generation in QSM, aiming

to simplify processing and enhance applicability. QSMmask-net was trained to generate masks as accurate as manually drawn masks, improving QSM accuracy. To validate its effectiveness, we compared QSM maps produced using QSMmask-net with those from conventional methods and state-of-the-art methods.

2 | Materials and Methods

2.1 | Overall Pipeline

The overall pipeline of the proposed method is shown in Figure 1. In Step 1, the QSMmask-net was trained. The network generated a whole-brain mask by minimizing loss with the manual mask using gradient echo (GRE) magnitude images. In Step 2, we performed ROI analysis on the QSM maps of simulation data reconstructed using eight different masks: a “manual” mask, a “standard” mask, a “standard+” mask, “FSL + hole filling” mask, “HD-BET” mask [27], “SynthStrip” mask [28], “Consensus” mask [2], and the “QSMmask-net” mask from the QSMmask-net. Here, the standard mask refers to a mask generated using FSL BET [29], and the standard+ mask refers to a mask with pixel dilation applied to the standard mask. The FSL + hole filling mask refers to a mask generated by applying hole filling to the standard mask. The consensus mask refers to a mask generated using the recommended method from

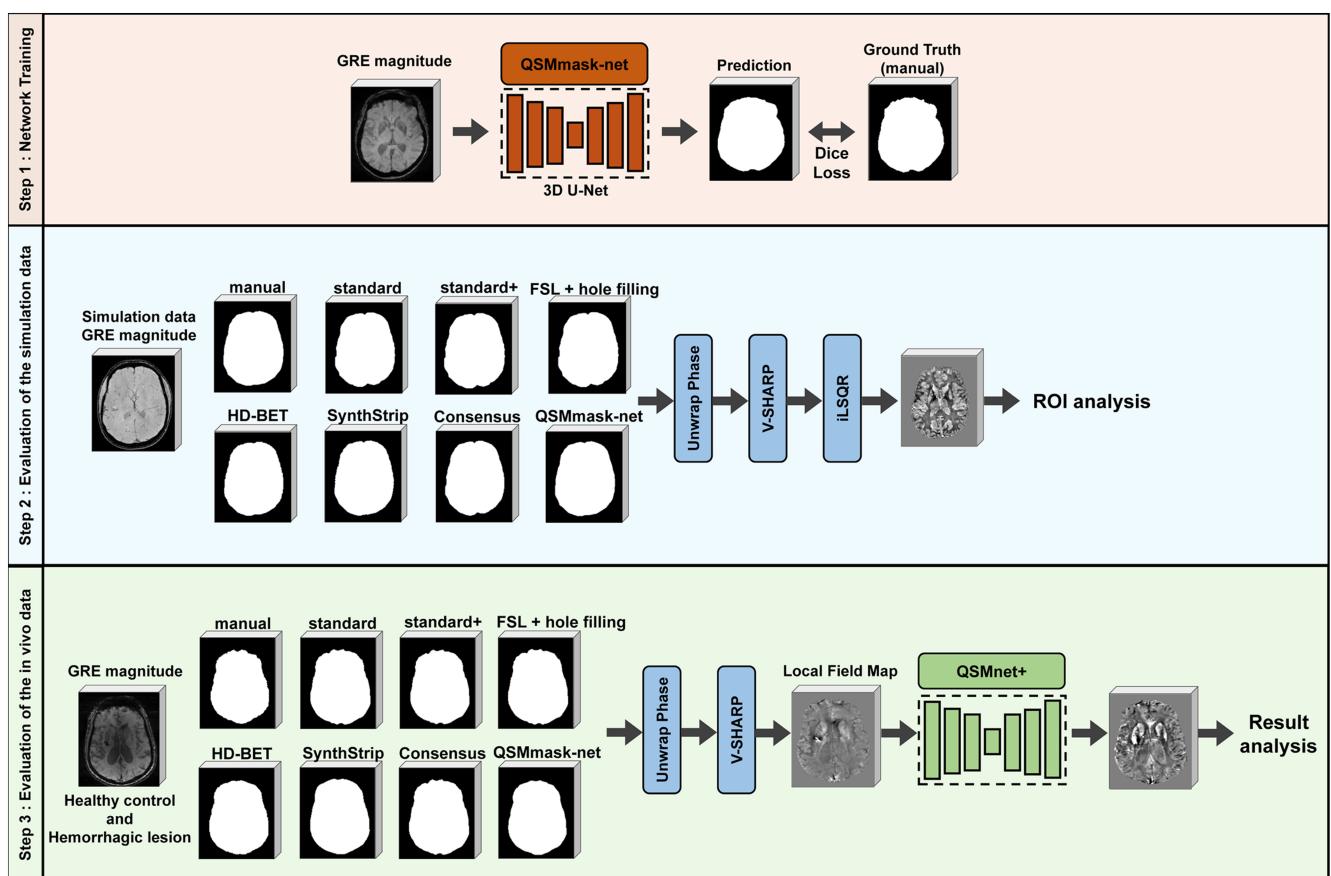


FIGURE 1 | Overall pipeline. In Step 1, the QSMmask-net model was trained using the magnitude of GRE images from the OASIS dataset. The ground truth was constructed using the manual mask method. In Steps 2 and 3, the QSMmask-net was validated by comparing representative ROIs and hemorrhagic lesions with the susceptibility values reconstructed using the manual, standard, standard+, FSL + hole filling, HD-BET, SynthStrip, Consensus, and QSMmask-net masks for the simulation data and in vivo data.

the QSM consensus paper [2]. Further details regarding these mask generation methods are provided in Section 2.4. In Step 3, we analyzed the QSM maps from four healthy controls and eight participants with hemorrhagic lesions. The source code and trained parameters can be accessed at <https://github.com/HUFS-AIMST/QSMmask-net>.

2.2 | Datasets for Network Training and Validation

Two datasets were used for the training and validation of the QSMmask-net. The first was the OASIS-3 dataset [30]. The second was created from *in vivo* scans of 18 healthy controls, acquired with three different scan parameter sets.

2.2.1 | OASIS-3 Dataset

The OASIS-3 dataset included 132 participants (age range, 73 ± 8 years; 71 males, 61 females). The dataset was randomly sampled and included 90 healthy controls, 15 participants with mild cognitive impairment, 2 participants with dementia, and 25 participants with Alzheimer's disease. The MRI scans were performed on a 3T Tim Trio scanner (Siemens Healthineers, Erlangen, Germany). We used the magnitude of GRE images obtained from susceptibility-weighted imaging (SWI) scans to generate masks. The OASIS-3 dataset was acquired with a field of view (FOV) = $192 \times 256 \times 80$ and a voxel size = $0.9 \times 0.9 \times 2 \text{ mm}^3$. The voxel size was resampled to a uniform $1 \times 1 \times 1 \text{ mm}^3$ for training purposes.

2.2.2 | Healthy Control Training Dataset

The healthy control training dataset consisted of data from 18 subjects (age range, 64 ± 6 years; 7 males, 11 females). This protocol was approved by the institutional review board. All the participants were provided written consent. Images were acquired on a 3T Prisma Fit scanner (Siemens Healthineers, Erlangen, Germany). The sequences included in this protocol were as follows:

- SWI scan: FOV = $240 \times 288 \times 176 \text{ mm}^3$, voxel size = $0.8 \times 0.8 \times 0.8 \text{ mm}^3$, repetition time (TR) = 46 ms, echo time (TE) = 13 ms, bandwidth = 260 Hz/pixel, and flip angle = 20°
- 3D GRE scan, single-echo acquisition: FOV = $256 \times 224 \times 176 \text{ mm}^3$, voxel size = $1 \times 1 \times 1 \text{ mm}^3$, TR = 33 ms, TE = 25 ms, bandwidth = 100 Hz/pixel, and flip angle = 15°
- 3D GRE scan, two-echo acquisition: FOV = $256 \times 224 \times 176 \text{ mm}^3$, voxel size = $1 \times 1 \times 1 \text{ mm}^3$, TR = 27 ms, TE1 = 16 ms, TE2 = 22 ms, bandwidth = 260 Hz/pixel, and flip angle = 20°

The SWI images were resampled to match the voxel size to $1 \times 1 \times 1 \text{ mm}^3$. The resampling was achieved by using the Fourier transform, cropping in the x-y plane, zero-padding in the z-direction, and applying the inverse Fourier transform. The number of participants for each scan was as follows: 1 participant for the SWI scan; 3 participants for the 3D GRE scan, single-echo acquisition; and 14 participants for the 3D GRE scan, two-echo acquisition.

2.3 | Datasets for Network Evaluation

We utilized data from four healthy controls (age range, 64 ± 5 years; 2 males, 2 females) ("healthy control evaluation dataset") and data from eight hemorrhage patients ("hemorrhage patient dataset") to assess the QSMmask-net. These protocols were approved by the institutional review board. The healthy control evaluation dataset was acquired with the same scan parameters as the 3D GRE scan, two-echo acquisition. The hemorrhage patient dataset was provided with written consent. Seven participants underwent scanning on a Siemens Skyra 3T scanner (scan parameters: 3D single-echo GRE with $\text{FOV} = 192 \times 192 \times 80 \text{ mm}^3$, voxel size = $1 \times 1 \times 1 \text{ mm}^3$, TR = 33 ms, TE = 25 ms, bandwidth = 100 Hz/pixel, and flip angle = 15°). The remaining participant was scanned on a Philips 3T Ingenia scanner (Best, The Netherlands; scan parameters: $\text{FOV} = 220 \times 220 \times 144 \text{ mm}^3$, voxel size = $0.5 \times 0.5 \times 2 \text{ mm}^3$, TR = 36 ms, TE = 25 ms, bandwidth = 255 Hz/pixel, and flip angle = 17°). The Philips datasets were resampled to achieve a $1 \times 1 \times 1 \text{ mm}^3$ resolution. The resampling was performed by applying zero-padding in the z-direction and cropping in the x-y plane direction in the Fourier domain.

2.4 | Mask Generation

Eight sets of masks were generated, and the results obtained with each of these masks were compared. The first set was generated by the QSMmask-net. The standard mask was generated using FSL BET with a threshold of 0.5. To address holes in the standard mask, we applied the dilation and hole filling methods to the standard mask, respectively. The standard+ mask was generated by applying pixel dilation to the standard mask using a disk-shaped kernel with a radius of 1. Dilation was applied to include the sagittal sinus; this is difficult to accomplish when using FSL BET to achieve complete segmentation. The FSL+hole filling mask was generated by applying "binary_fill_holes" function from Simple ITK [31–33] to the standard mask, utilizing 8-connectivity for defining connected components. For comparison with state-of-the-art methods, HD-BET and SynthStrip masks were generated from 3D GRE scans. The Consensus mask was generated following the recommendations outlined in the QSM consensus paper [2]. Finally, the manual mask was generated by manually segmenting all of the data from the OASIS-3 and healthy control training datasets. This manual dataset served as the ground truth. Three trained raters autonomously created manual masks by modifying precalculated standard masks. Intraclass correlation coefficients (ICC) were calculated to assess inter-rater reliability.

The modification rule stipulated the inclusion of two voxels beyond the boundary of the brain tissue and one voxel at the junction of the frontal and parietal lobes. This modification rule was empirically defined by observing the difference between an input mask (a mask generated from magnitude images using FSL BET or other algorithms) for V-SHARP [19] and a newly generated mask by V-SHARP. V-SHARP has been reported to generate an eroded mask from the input mask [34, 35]. Thus, the number of voxels used to dilate the mask was empirically defined to reduce the overeroded voxels and to exclude nonbrain tissue close to the background. The final manual mask was generated by calculating the intersection of the masks generated by the three raters. The mean time required to produce a manual mask was 1.5 h.

2.5 | Training the QSMmask-Net Model

The QSMmask-net model was trained on an RTX 2080-Ti with a GPU-RAM capacity of 12GB. From the OASIS dataset and the healthy control training dataset, we used data from 122 participants for training and randomly selected data from 10 participants in the OASIS dataset for validation. A total of $128 \times 128 \times 64$ patches and a 3D U-Net architecture were used [36]. The batch size was set to 4. The number of epochs was set to 200, and early stopping was used to halt the network once it had achieved a satisfactory level of convergence. The network parameters were updated using the Dice loss and Adam optimizer [37] with a learning rate of 0.001. In addition, data augmentation was applied using scale intensity, random flip, random crop, random affine transform, Gibbs noise, and Gaussian smoothing implemented in a MONAI library [38]. The final binary segmentation mask was generated by applying a threshold of 0.5 to the predicted likelihood map.

2.6 | Evaluation of the QSMmask-Net

2.6.1 | ROI Analysis of QSM Reconstruction Challenge

2.0 Simulation Data

We conducted an ROI analysis on simulation data from the QSM reconstruction challenge 2.0 [39]. We analyzed the susceptibility values of 10 different ROIs, including the caudate, globus pallidus, putamen, red nucleus, dentate nucleus, substantia nigra, subthalamic nucleus, thalamus, white matter, gray matter, and cerebrospinal fluid (CSF). We used ROI masks provided by simulation data. The QSM map was then reconstructed following these steps: (1) Laplacian-based phase unwrapping [40], (2) background field removal using STI-Suite V-SHARP, and (3) QSM reconstruction performed using STI-Suite iLSQR [41].

We analyzed the mean susceptibility values of the ROIs in both the ground truth susceptibility map from simulation data and the QSM map using the eight different masks. In addition, a paired *t*-test was performed. The significant level was defined as 0.05.

2.6.2 | Comparison of the Dice Score

To evaluate the segmentation accuracy of the QSMmask-net, we compared the mean Dice score of standard, standard+, FSL + hole filling, HD-BET, SynthStrip, Consensus, and QSMmask-net masks with manual masks in the evaluation datasets of four healthy subjects and eight hemorrhage patients. For the evaluation dataset of eight hemorrhage patients, the Consensus mask was excluded because the patient dataset consists of single-echo scans.

2.6.3 | Qualitative Comparison of the Effect of Masks on QSM Maps

To evaluate the effect of masks on QSM maps from four healthy controls, QSM maps were reconstructed following these steps: (1) Laplacian-based phase unwrapping, (2) generation of a local field map using STI-Suite V-SHARP, and (3) QSM map reconstruction using QSMnet+, a deep neural network that produces

QSM maps of quality comparable to COSMOS, employing only a single-head orientation and facilitating practically real-time processing [42]. We conducted a qualitative assessment of specific regions, such as the area between the frontal and temporal lobes. Because this ROI included overeroded voxels after the application of V-SHARP, we focused on whether the QSMmask-net generated a mask with minimally eroded voxels in the ROI to accurately reconstruct QSM.

2.6.4 | Comparison of Susceptibility Values of Representative ROIs in Healthy Subjects

We also compared the mean susceptibility values of the caudate, globus pallidus, putamen, red nucleus, substantia nigra, and thalamus. These ROIs were manually segmented by a single rater using ITK-SNAP [43]. The segmentations were performed on the GRE magnitude images. We performed paired *t*-tests to compare the mean susceptibility values obtained using manual segmentation with those obtained using the standard, standard+, and QSMmask-net masks. The significant level was defined as 0.05.

2.6.5 | Comparison of Susceptibility Values of Hemorrhagic Lesions

The QSM reconstruction pipeline for participants with hemorrhage lesions was identical to the QSM reconstruction pipeline used for four healthy controls. An experienced rater manually drew the ROIs of hemorrhagic lesions based on the GRE magnitude images. The mean and standard deviation of susceptibility values were then measured on the QSM maps created using the eight masks. To assess the clinical efficacy of the QSMmask-net, a linear regression analysis was also conducted on the susceptibility values of hemorrhagic lesions in eight participants.

3 | Results

3.1 | Inter-Rater Reliability Assessment of Manual Masks

ICC was calculated to evaluate inter-rater reliability among three trained raters for each dataset. Table 1 shows mean ICC values for the OASIS dataset ($ICC = 0.9982$), the healthy control training dataset ($ICC = 0.9976$), the healthy control evaluation dataset ($ICC = 0.9963$), and the hemorrhage patient dataset ($ICC = 0.9980$).

3.2 | ROI Analysis of QSM Reconstruction Challenge 2.0 Simulation Data

Susceptibility values obtained with the standard mask and the FSL + hole filling mask were significantly different from those obtained with the manual mask in the globus pallidus, putamen, dentate nucleus, gray matter, and CSF (Table 2). Values obtained with the standard+ mask differed significantly from those obtained with the manual mask in the globus pallidus, putamen, dentate nucleus, and CSF. Significant differences in susceptibility values were observed between the HD-BET and manual

TABLE 1 | Intraclass correlation (ICC) scores of manual masks.

Dataset		Mean ICC
Training/validation	OASIS dataset (N=132)	0.9982
	Healthy control training dataset (N=18)	0.9976
Inference	Healthy control evaluation dataset (N=4)	0.9963
	Hemorrhage patient dataset (N=8)	0.9980

Note: The mean ICC score represents the average of three pairwise ICC scores calculated between raters: ICC_1 (Rater 1 and Rater 2), ICC_2 (Rater 2 and Rater 3), and ICC_3 (Rater 1 and Rater 3).

mask in the globus pallidus, putamen, red nucleus, and CSF. Similarly, the SynthStrip mask resulted in significant differences from the manual mask in the globus pallidus, putamen, red nucleus, white matter, and CSF. The Consensus mask significantly differed from the manual mask in the globus pallidus, putamen, red nucleus, dentate nucleus, gray matter, and CSF. In contrast, susceptibility values obtained with the QSMmask-net mask demonstrated no significant difference from those obtained with the manual mask.

3.3 | Quantitative Comparison of the Masks

Table 3 shows the mean Dice coefficients comparing manual masks with each method for the healthy control evaluation dataset and the hemorrhage patient dataset, presented as mean \pm standard deviation. The QSMmask-net shows the highest Dice score for both groups and the lowest standard deviation for the hemorrhage patient dataset.

3.4 | Qualitative Comparison of the Masks

Figure 2 demonstrates variations caused by the mask generation approach, which included overdilated voxels and unwanted holes. For the comparison of overdilated voxels, there was no noticeable difference between the HD-BET, Consensus, QSMmask-net, and manual masks. However, the standard, standard+, FSL + hole filling, and SynthStrip masks demonstrated differences from the manual mask, particularly in the frontal lobe. These masks were slightly more overdilated compared to the QSMmask-net mask. Moreover, the standard mask demonstrated unwanted holes inside the brain tissue. These variations in the masks caused discrepancies in susceptibility values.

3.5 | Quantitative Comparison of the Effect of Masks on QSM Maps

Figure 3 presents a qualitative comparison of the QSM maps from 2 healthy participants reconstructed using manual, standard,

standard+, FSL + hole filling, HD-BET, SynthStrip, Consensus, and QSMmask-net masks, with the manual mask used as the reference. In the results from Subject 1, images using the standard mask included a large hole in the center of the brain. Although the dilation process partially addressed this issue, holes still remained in images using the standard+ mask. Because of the effect of these holes, QSM maps from the standard and standard+ masks demonstrated decreased signal intensity in the cerebral vein. In contrast, such holes were not observed in the FSL + hole filling, HD-BET, SynthStrip, Consensus, or QSMmask-net masks. In Subject 2, overdilated voxels were seen as highlighted in blue arrows within the blue box. With the standard and standard+ masks, nonbrain areas were prominently present around the periphery of the brain mask. As a result, the QSM maps obtained with the standard and standard+ masks exhibited artifacts caused by residual background fields at the boundary of the brain. Conversely, the FSL + hole filling, HD-BET, SynthStrip, Consensus, and QSMmask-net masks substantially reduced these artifacts compared to the standard and standard+ masks.

3.6 | Comparison of Susceptibility Values of Representative ROIs in Healthy Control Evaluation Dataset

In the healthy control evaluation dataset, no significant differences were observed across all ROIs when comparing the susceptibility values obtained using the manual mask to those obtained using the FSL + hole filling, HD-BET, Consensus, and QSMmask-net masks (Table 4). However, the susceptibility values obtained using the standard mask and the standard+ mask differed significantly from those obtained using the manual mask across all ROIs. Additionally, a significant difference in susceptibility values in the thalamus was observed between the SynthStrip and manual mask.

3.7 | Comparison of Susceptibility Values of Hemorrhagic Lesions

Figure 4 presents the linear regression results for a representative participant with a hemorrhagic lesion. The linearity of the susceptibility values reconstructed using the QSMmask-net mask (slope = 0.9937, intercept = 0.0004, R^2 = 0.998, p < 0.05) and the HD-BET mask (slope = 0.9911, intercept = 0.0018, R^2 = 0.9985, p < 0.05) was closer to the line of the susceptibility values reconstructed using the manual mask than that of the standard mask (slope = 0.7555, intercept = 0.0346, R^2 = 0.7645, p < 0.05). When compared with the standard, the linearity improved with the standard+ mask (slope = 0.9359, intercept = 0.0020, R^2 = 0.9968, p < 0.05), FSL + hole filling mask (slope = 0.8194, intercept = 0.0039, R^2 = 0.9939, p < 0.05), and SynthStrip mask (slope = 0.9887, intercept = 0.0019, R^2 = 0.9981, p < 0.05).

Figure 5 presents the mean susceptibility values of hemorrhagic lesions in eight patients obtained with the manual, standard, standard+, FSL + hole filling, HD-BET, SynthStrip, and QSMmask-net masks. The QSM map using the QSMmask-net mask and the FSL + hole filling mask exhibited the smallest mean discrepancy when compared with the other mask generation techniques.

TABLE 2 | Mean (standard deviation) susceptibility values for simulation data (ppm).

ROI	Mask						QSM/mask-net
	Manual	Standard	Standard+	FSL + hole filling	HD-BET	SynthStrip	
Caudate	0.0298 (0.0176) <i>p</i> = 1.0	0.0302 (0.0175) <i>p</i> = 1.0	0.0315 (0.0171) <i>p</i> = 1.0	0.0302 (0.0176) <i>p</i> = 1.0	0.0306 (0.0175) <i>p</i> < 0.05	0.0305 (0.0177) <i>p</i> < 0.05	0.0302 (0.0176) <i>p</i> = 1.0
Globus pallidus	0.0832 (0.0248) <i>p</i> < 0.05	0.0798 (0.0245) <i>p</i> < 0.05	0.0799 (0.0246) <i>p</i> < 0.05	0.0798 (0.0245) <i>p</i> < 0.05	0.0819 (0.0245) <i>p</i> < 0.05	0.0809 (0.0246) <i>p</i> < 0.05	0.0799 (0.0245) <i>p</i> < 0.05
Putamen	0.0121 (0.0207) <i>p</i> < 0.05	0.0112 (0.0193) <i>p</i> < 0.05	0.0117 (0.0194) <i>p</i> < 0.05	0.0112 (0.0194) <i>p</i> < 0.05	0.0116 (0.0203) <i>p</i> < 0.05	0.0117 (0.0195) <i>p</i> < 0.05	0.0112 (0.0193) <i>p</i> < 0.05
Red nucleus	0.0733 (0.0298) <i>p</i> = 0.1451	0.0730 (0.0296) <i>p</i> = 0.4303	0.0731 (0.0297) <i>p</i> = 0.1194	0.0729 (0.0296) <i>p</i> = 0.1194	0.0722 (0.0296) <i>p</i> < 0.05	0.0699 (0.0293) <i>p</i> < 0.05	0.0727 (0.0295) <i>p</i> < 0.05
Dentate nucleus	0.1161 (0.0395) <i>p</i> < 0.05	0.1156 (0.0397) <i>p</i> < 0.05	0.1147 (0.0397) <i>p</i> < 0.05	0.1156 (0.0397) <i>p</i> < 0.05	0.1168 (0.0395) <i>p</i> = 1.0	0.1166 (0.0396) <i>p</i> = 1.0	0.1155 (0.0396) <i>p</i> < 0.05
Substantia nigra and subthalamic nucleus	0.0648 (0.0298) <i>p</i> = 1.0	0.0668 (0.0282) <i>p</i> = 1.0	0.0660 (0.0281) <i>p</i> = 1.0	0.0667 (0.0282) <i>p</i> = 1.0	0.0664 (0.0288) <i>p</i> = 1.0	0.0661 (0.0284) <i>p</i> = 1.0	0.0668 (0.0282) <i>p</i> = 1.0
Thalamus	0.003 (0.0155) <i>p</i> = 1.0	0.0053 (0.0144) <i>p</i> = 1.0	0.0049 (0.0141) <i>p</i> = 1.0	0.0053 (0.0144) <i>p</i> = 1.0	0.0038 (0.0149) <i>p</i> = 1.0	0.0043 (0.0146) <i>p</i> = 1.0	0.0054 (0.0144) <i>p</i> = 1.0
White matter	-0.0152 (0.0145) <i>p</i> = 1.0	-0.0153 (0.0135) <i>p</i> = 1.0	-0.0154 (0.0141) <i>p</i> = 1.0	-0.0153 (0.0135) <i>p</i> = 1.0	-0.0152 (0.0138) <i>p</i> = 1.0	-0.0156 (0.0143) <i>p</i> < 0.05	-0.0152 (0.0135) <i>p</i> = 1.0
Gray matter (Cortex)	0.0084 (0.0187) <i>p</i> < 0.05	0.0081 (0.0162) <i>p</i> = 1.0	0.0086 (0.0189) <i>p</i> = 1.0	0.0081 (0.0162) <i>p</i> < 0.05	0.0085 (0.0173) <i>p</i> = 1.0	0.0084 (0.0192) <i>p</i> = 0.1791	0.0080 (0.0155) <i>p</i> < 0.05
CSF	0.0063 (0.0352) <i>p</i> < 0.05	0.0039 (0.0209) <i>p</i> < 0.05	0.0051 (0.0328) <i>p</i> < 0.05	0.0039 (0.0209) <i>p</i> < 0.05	0.0047 (0.0249) <i>p</i> < 0.05	0.0056 (0.0386) <i>p</i> < 0.05	0.0038 (0.0204) <i>p</i> < 0.05

Note: *p* values shown are for comparison with manual mask.
Abbreviations: CSF, cerebrospinal fluid; ROI, region of interest.

TABLE 3 | Mean (standard deviation) Dice score for inference dataset.

	Standard	Standard+	FSL + hole filling	HD-BET	SynthStrip	Consensus	QSMmask-net
Healthy control evaluation dataset (<i>N</i> =4)	0.964 (0.006)	0.957 (0.001)	0.954 (0.011)	0.962 (0.007)	0.866 (0.051)	0.954 (0.011)	0.966 (0.006)
Hemorrhage patient dataset (<i>N</i> =8)	0.939 (0.028)	0.945 (0.024)	0.940 (0.028)	0.939 (0.021)	0.905 (0.053)	N/A	0.955 (0.003)

Abbreviation: N/A, not applicable.

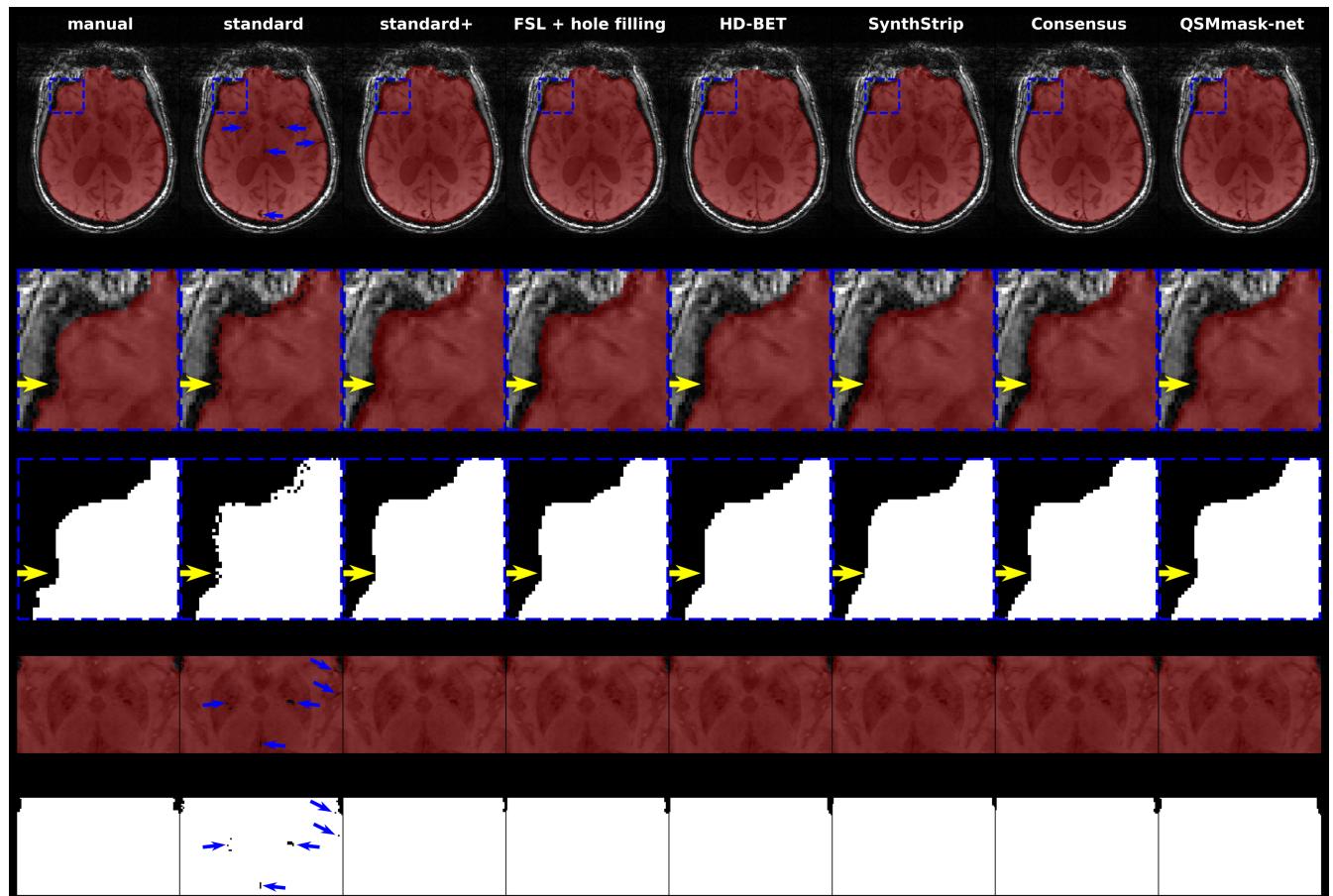


FIGURE 2 | Comparison of masks. (First row) GRE magnitude image with overlaid masks. Blue boxes indicate the differences caused by over-dilation, and blue arrows point to unwanted holes. (Second row) Zoomed-in mask-overlaid view of the differences resulting from over-dilation (yellow arrows). (Third row) Zoomed-in mask view highlighting differences resulting from over-dilation. (Fourth row) Zoomed-in mask-overlaid view of the differences resulting from unwanted holes (blue arrows). The standard mask exhibits unwanted holes within the brain tissue. (Fifth row) Zoomed-in mask view of the differences resulting from unwanted holes.

Table 5 shows a linear regression analysis comparing the hemorrhagic lesion susceptibility values of eight participants, obtained using the standard, standard+, FSL + hole filling, HD-BET, SynthStrip, and QSMmask-net masks with the results obtained using a manual mask. Although the regression slope for the SynthStrip mask was closest to 1, the QSMmask-net mask demonstrated superior linearity ($R^2=0.9992$) compared to the SynthStrip mask ($R^2=0.9925$). In all cases, the R^2 values

obtained between the manual mask and each mask generation method were statistically significant ($p<0.05$).

Figure 6 shows mask differences near hemorrhagic lesions for two representative patients. For Patient 1, no noticeable differences were observed in the masks near hemorrhages within the brain. In contrast, Patient 2 presented hemorrhagic lesions near the brain boundary regions. In this case, the FSL-based

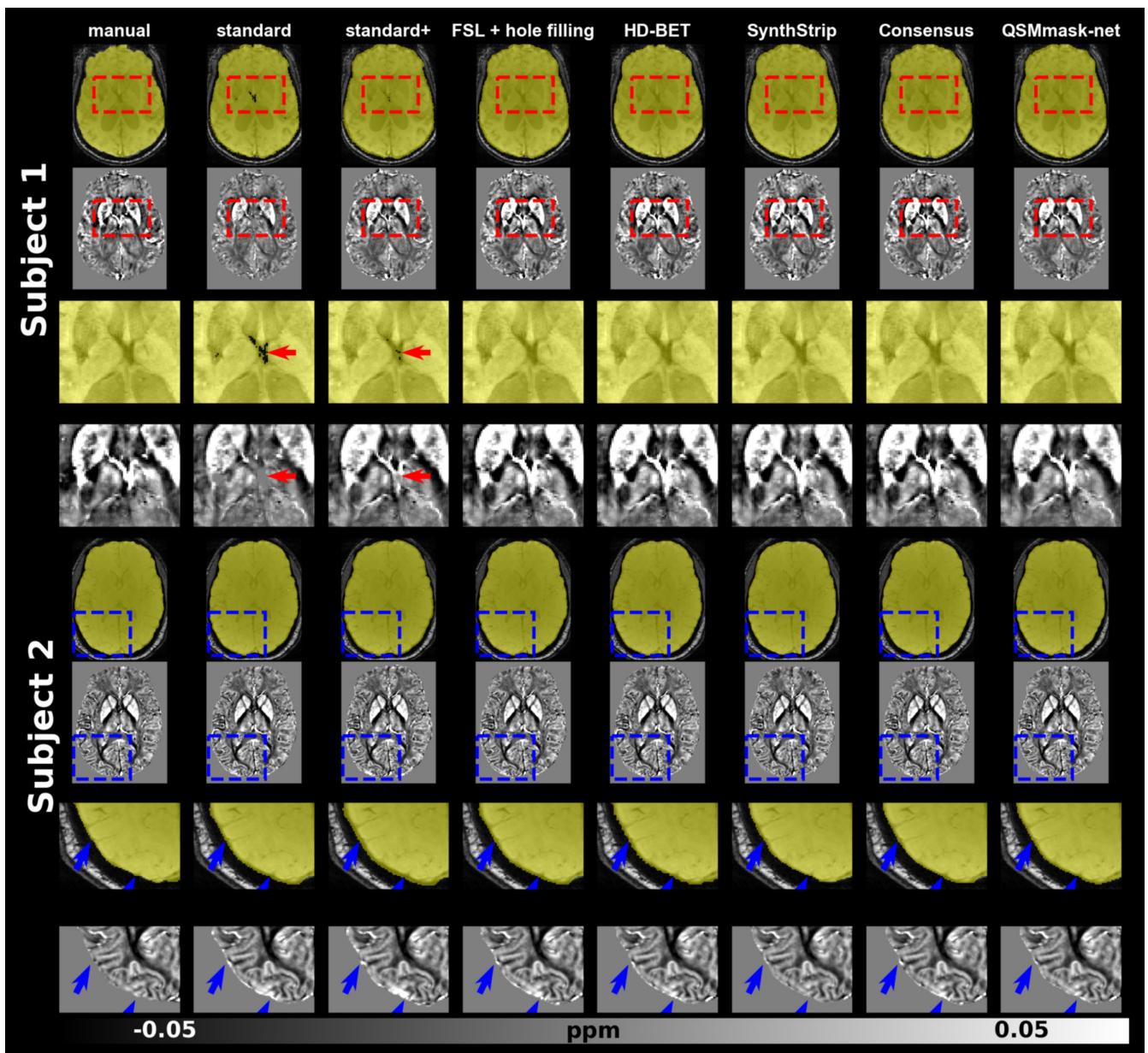


FIGURE 3 | QSM reconstruction results with manual, standard, standard+, FSL + hole filling, HD-BET, SynthStrip, Consensus, and QSMmask-net masks. Subject 1 represents a case of varying susceptibility due to unwanted holes in the center of the brain (red arrows), whereas Subject 2 shows varying susceptibility due to over-dilated voxels (blue arrows). The first and fifth rows display the GRE magnitude images with each overlaid mask, with bounding boxes highlighting the prominent differences. The second and sixth rows show the QSM maps reconstructed with each mask. The third, fourth, seventh, and eighth rows show zoomed-in views of the masks and QSM maps.

methods (standard, standard+, and FSL + hole filling masks) showed incomplete segmentation around lesions located at the brain boundaries. Conversely, the HD-BET, SynthStrip, and QSMmask-net masks successfully included the hemorrhagic lesions.

4 | Discussion

In this study, we propose an efficient mask generation method for QSM using deep learning. To create a mask generation method requiring minimal labor input while producing results comparable to those seen with the manual mask method, we developed a deep neural network called QSMmask-net.

Furthermore, we aimed to provide a simplified framework for QSM processing that is easily accessible, even to individuals lacking knowledge of medical image processing. Employing deep learning-based QSM processing tools at each step, including QSMmask-net for mask generation and QSMnet+ for calculating dipole inversion, enables users to effortlessly generate QSM maps that are comparable in quality to those achieved with manually segmented masks and COSMOS, even with a single-head orientation.

We evaluated the QSMmask-net using simulation data, the healthy control evaluation dataset, and data from patients with hemorrhagic lesions to confirm its effectiveness. We found that the QSMmask-net was able to generate masks of similar quality

TABLE 4 | Mean (standard deviation) susceptibility values of ROIs in four healthy controls.

ROI		Mask						
		Manual	Standard	Standard+	FSL + hole filling	HD-BET	SynthStrip	Consensus
Caudate	0.0372 (0.021)	0.0316 (0.024) <i>p</i> < 0.05	0.0332 (0.023) <i>p</i> < 0.05	0.0396 (0.028) <i>p</i> = 0.2308	0.0398 (0.028) <i>p</i> = 0.2087	0.035 (0.027) <i>p</i> = 0.5717	0.0396 (0.028) <i>p</i> = 0.2367	0.0366 (0.020) <i>p</i> = 0.5597
Globus pallidus	0.1342 (0.046)	0.1014 (0.048) <i>p</i> < 0.05	0.1173 (0.049) <i>p</i> < 0.05	0.1373 (0.059) <i>p</i> = 0.4190	0.1396 (0.059) <i>p</i> = 0.2135	0.1272 (0.062) <i>p</i> = 0.2352	0.1372 (0.059) <i>p</i> = 0.4436	0.1313 (0.042) <i>p</i> = 0.5253
Putamen	0.0459 (0.031)	0.0378 (0.031) <i>p</i> < 0.05	0.0400 (0.030) <i>p</i> < 0.05	0.0479 (0.036) <i>p</i> = 0.3672	0.0484 (0.035) <i>p</i> = 0.2561	0.0443 (0.036) <i>p</i> = 0.3977	0.0479 (0.036) <i>p</i> = 0.3720	0.0443 (0.030) <i>p</i> = 0.1929
Red nucleus	0.1062 (0.046)	0.0886 (0.052) <i>p</i> < 0.05	0.0861 (0.049) <i>p</i> < 0.05	0.1060 (0.060) <i>p</i> = 0.9581	0.1122 (0.058) <i>p</i> = 0.1634	0.0880 (0.061) <i>p</i> = 0.0627	0.1061 (0.060) <i>p</i> = 0.9634	0.1009 (0.044) <i>p</i> = 0.1158
Substantia nigra	0.1147 (0.040)	0.0910 (0.047) <i>p</i> < 0.05	0.0972 (0.048) <i>p</i> < 0.05	0.1187 (0.060) <i>p</i> = 0.4098	0.1236 (0.059) <i>p</i> = 0.1162	0.1063 (0.061) <i>p</i> = 0.1041	0.1177 (0.060) <i>p</i> = 0.5383	0.1219 (0.041) <i>p</i> = 0.2521
Thalamus	0.0120 (0.018)	0.0102 (0.017) <i>p</i> < 0.05	0.0103 (0.017) <i>p</i> < 0.05	0.0121 (0.019) <i>p</i> = 0.5197	0.0121 (0.019) <i>p</i> = 0.3436	0.0116 (0.020) <i>p</i> < 0.05	0.0121 (0.019) <i>p</i> = 0.3995	0.0125 (0.017) <i>p</i> = 0.3240

Note: *p* values shown are for comparison with the manual mask. Bold represents the lowest difference between the mean ROI susceptibility of the manual mask and those of other mask generation methods.
Abbreviation: ROI, region of interest.

to manual masks in a much shorter time. Specifically, whereas the manual mask method required approximately 1.5 h to generate a mask for one subject, the QSMmask-net was able to produce a mask of comparable quality in just 30 s.

Our findings revealed a notable variation in QSM maps based on the mask used. By using simulation data with tissue-typical susceptibility values referenced from the literature [44, 45], we were able to objectively compare the mean susceptibility values obtained with different mask generation methods. Our analysis demonstrated a significant difference in the ROI susceptibility values depending on the mask employed. Notably, the susceptibility values obtained with the QSMmask-net exhibited the closest resemblance to the values obtained with the manual mask in the ROI analysis of simulation data.

We also demonstrated the impact of different masks on QSM results using the healthy control evaluation dataset. We found that the type of mask used can lead to inaccuracies in susceptibility values caused by holes in the mask and the introduction of noise in the QSM due to overdilation. Holes in the mask result in intensity inhomogeneity within the same tissue, leading to a miscalculation of susceptibility values. Overdilation of the mask results in increased noise in the QSM due to the background field. In this study, the standard mask failed to adequately encompass the cortical regions adjacent to adipose tissue in the skull and the superior sagittal sinus, leading to holes within images of the brain. Pixel dilation is frequently used to enhance the standard mask; however, this process may inadvertently incorporate background pixels, leading to the creation of artifacts. This was seen with the standard+ mask in the current study. In our qualitative comparison using healthy control data, the FSL + hole filling, HD-BET, SynthStrip, Consensus, and QSMmask-net masks effectively mitigated the limitations found with standard and standard+ masks. The manual mask yields comparatively accurate QSM results, but at the expense of a substantial amount of time required for mask generation. Consequently, its use is restricted in situations requiring extensive analysis of substantial datasets. Our proposed mask generation approach could provide masks with comparable performance to that of the manual masks, but with notably reduced labor. Thus, the QSMmask-net could be useful for assessing large quantities of clinical QSM data.

In terms of time efficiency, the FSL + hole filling, HD-BET, SynthStrip, and Consensus masks are methods comparably effective to QSMmask-net. Additionally, the analysis of healthy controls in this study demonstrated no significant differences between the susceptibility values reconstructed with each mask and those with the manual mask across all compared ROIs. Furthermore, the Consensus and the QSMmask-net masks produced susceptibility values comparable to those of manual masks. However, notable differences were observed in the QSM results when comparing the manual mask with the standard and standard+ masks.

In patients with hemorrhagic lesions, there was a strong correlation between the mean susceptibility values obtained with the QSMmask-net and the values obtained with the manual mask. Notably, the QSMmask-net mask also exhibited the least variation in mean susceptibility values within the lesion across the

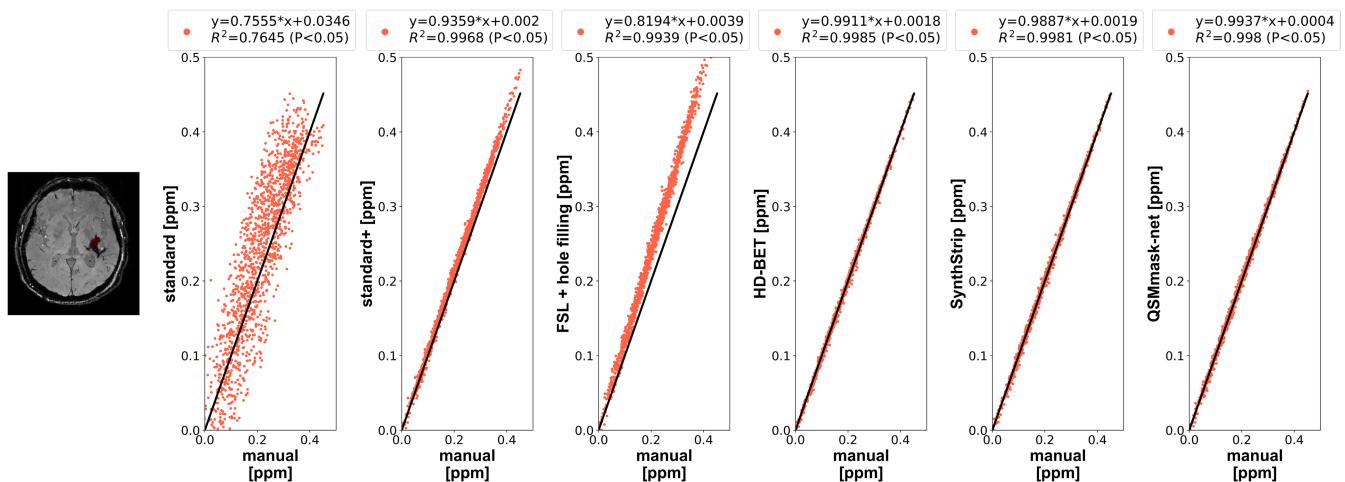


FIGURE 4 | Comparison of the linearity of hemorrhagic lesion susceptibility values reconstructed using the manual, standard, standard+, FSL + hole filling, HD-BET, SynthStrip, and QSMmask-net masks. In a scatter plot of susceptibility values in the lesions from one representative patient, the black solid line is the unity-slope line for lesion susceptibility values reconstructed with the manual mask, whereas the orange dots are lesion susceptibility values reconstructed with the standard, standard+, FSL + hole filling, HD-BET, SynthStrip, and QSMmask-net masks. The susceptibility values reconstructed using the standard mask exhibit more scatter compared to those obtained with the standard+ and the QSMmask-net masks. The standard+, FSL + hole filling, HD-BET, SynthStrip, and QSMmask-net masks show improved linearity than the standard mask. Among these methods, the QSMmask-net shows a slope closest to 1, with R^2 values of 0.9980.

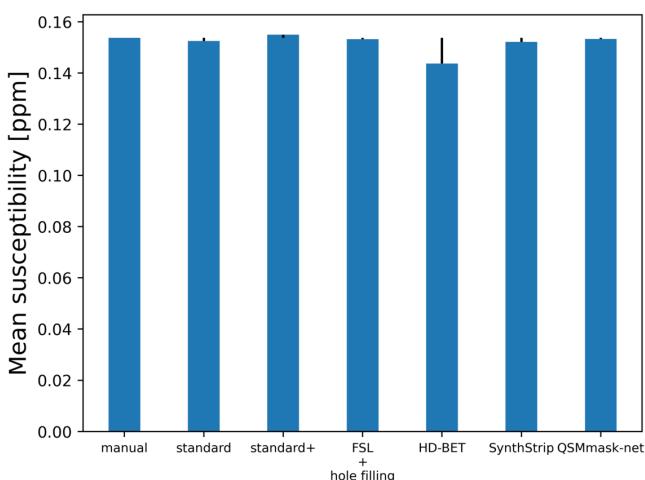


FIGURE 5 | Mean susceptibility values of hemorrhagic lesions in eight participants, obtained using the manual, standard, standard+, FSL + hole filling, HD-BET, SynthStrip, and QSMmask-net masks. For the hemorrhagic lesion case, the QSMmask-net mask and the FSL + hole filling mask show the smallest differences from the manual mask compared to other masks.

eight participants. In contrast, FSL-based masks occasionally excluded hemorrhagic lesions located near brain boundaries. In such cases, the accuracy of FSL-based masks could be improved by fine-tuning parameters (e.g., threshold) or by combining masks generated from different input scans, such as MPRAGE T_1w or multi-echo scans.

However, such adjustments necessitate a laborious process of optimization that must be repeated for each dataset. The QSMmask-net offers a convenient and ready-to-use solution that automatically generates optimal masks from a single GRE magnitude image without the need for any personal intervention.

HD-BET and SynthStrip are state-of-the-art, deep learning-based brain extraction methods. These methods were trained on various MRI contrasts, such as T_1w , T_2w , and FLAIR, ensuring robustness across different input scans. In our study, the robustness of HD-BET and SynthStrip on GRE scans was evaluated by comparing the results obtained from healthy subjects. However, our patient dataset consisted of single-echo GRE images acquired with a relatively long echo time (TE), whereas previous studies utilized either masks averaged across multiple echoes [46, 47] or combined masks derived from MPRAGE $T1w$ and GRE images [48]. This difference in dataset characteristics might explain the susceptibility differences observed in hemorrhagic lesions. Additionally, we found larger variations in susceptibility values within lesions when using the HD-BET mask and manual mask compared to other masks, likely influenced by differences in image acquisition parameters.

The principal objective of our study was to improve usability for inexperienced users, necessitating a comparison with a widely utilized tool such as FSL. Our proposed approach rectifies the deficiencies of FSL in QSM processing by reducing user intervention and enabling the generation of masks that attain quality akin to manually segmented masks in a singular step. Although QSMmask-net requires minimal user intervention, reorienting the data to the standard MNI space is necessary when the image orientation is not axial.

The participants in this study whose data were used for training and evaluation of the QSMmask-net were scanned on a Siemens 3T scanner. To enhance the generalizability of the QSMmask-net, more diverse datasets from scanners produced by different vendors and with different field strengths are required. However, there is a lack of such datasets for GRE magnitude images. Consequently, future research should explore training methodologies for scanner-invariant segmentation, in addition to generating QSM-optimized segmentation masks. In

TABLE 5 | Linear regression analysis results for the mean susceptibility values of hemorrhagic lesions in eight participants.

Variable	Mask						
	Standard	Standard+	FS + hole filling	HD-BET	SynthStrip	Consensus	QSMmask-net
Slope	0.8604	0.8656	0.8639	1.076	1.007	N/A	0.9814
(Standard error)	(0.029)	(0.023)	(0.062)	(0.080)	(0.036)		(0.007)
Intercept	0.032	0.0184	0.0214	-0.0010	0.0010	N/A	0.0031
(Standard error)	(0.005)	(0.004)	(0.011)	(0.013)	(0.006)		(0.001)
R ²	0.9840	0.9899	0.9697	0.9681	0.9925	N/A	0.9992
(p value)	(p < 0.05)	(p < 0.05)	(p < 0.05)	(p < 0.05)	(p < 0.05)		(p < 0.05)

Abbreviation: N/A, not applicable.

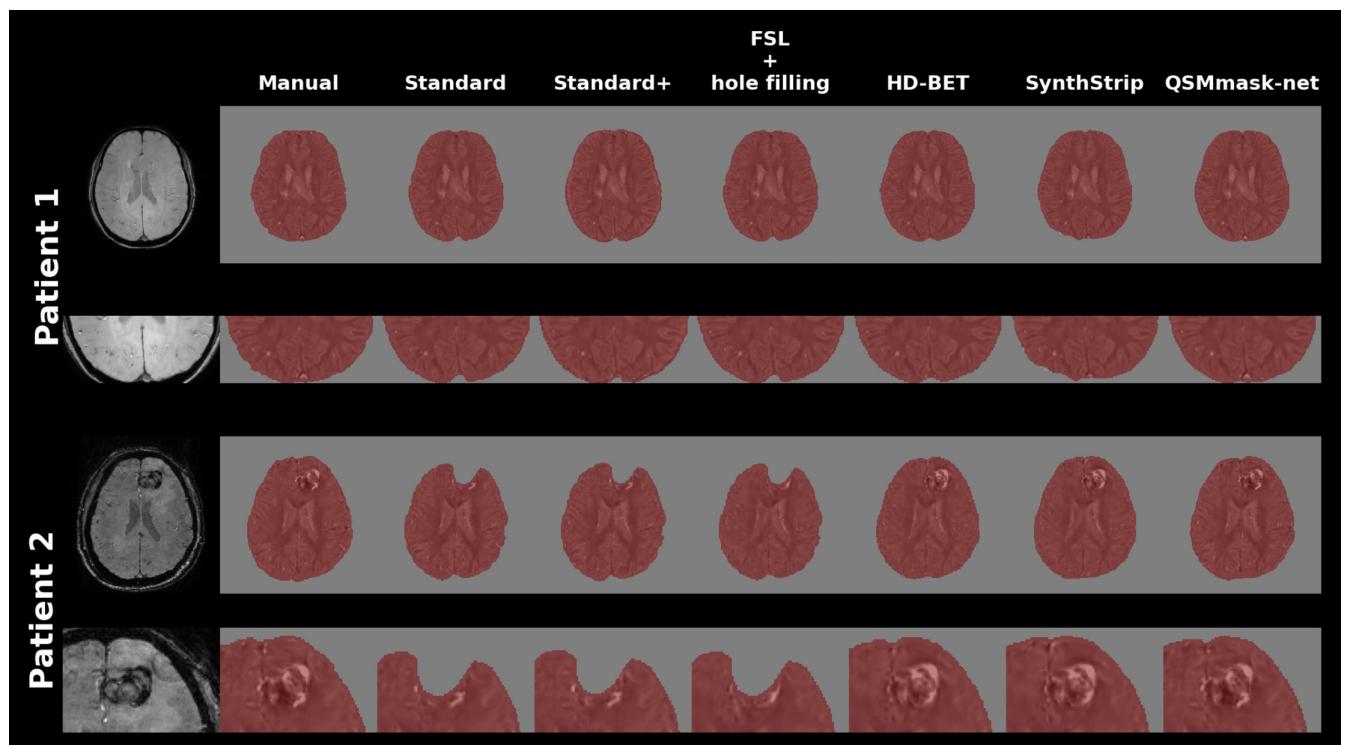


FIGURE 6 | Comparison of masking in hemorrhage patients. The magnitude images of two hemorrhage patients and zoomed-in views are shown in the leftmost column. In each mask column, QSM results with overlaid masks are shown. For Patient 1, no noticeable differences in the masks are observed near the microhemorrhage lesions across the different mask generation methods. In contrast, for Patient 2, FSL-based methods exhibit incomplete segmentation near the hemorrhage lesion.

this study, we also used a single-echo GRE magnitude image for training purposes. Future research should explore the resilience of averaged magnitude images from multi-echo scan data.

In our evaluation using GRE datasets with varying spatial resolutions, QSMmask-net demonstrated good generalizability on standard-resolution scans. However, its robustness was relatively reduced on high-resolution data compared to other brain masking methods. Although these results are not shown in detail, they suggest that the performance of QSMmask-net may be resolution dependent. The QSMmask-net could be further improved by developing a resolution-free segmentation method and training on MR images of other areas of the body. First, the QSMmask-net was trained in this study using scans with a voxel

size of $1 \times 1 \times 1 \text{ mm}^3$. Future work should focus on developing segmentation approaches that are robust to variations in image resolution. Ensuring resolution robustness is essential for the generalizability of deep neural networks used in QSM reconstruction pipelines. Notably, performance degradation has been reported in deep learning-based dipole inversion methods when applied to data with different resolutions [49]. A recent study [50] proposed generating synthetic data using diverse combinations of contrast, spatial resolution, anatomical morphology, and noise levels. This strategy may be further optimized in future work to improve brain mask generation for QSM reconstruction across a wide range of imaging conditions. Second, although the QSMmask-net was trained using brain magnitude images in this study, it can be further trained on images of other areas of

the body. Specifically, segmentation foundation models such as MedSAM [51] could be fine-tuned to generate optimal masks for the QSM reconstruction of various body parts.

5 | Conclusion

In conclusion, this study represents the first step toward an automated and optimal workflow for QSM reconstruction. Through simulation data, *in vivo* healthy control comparison, and analysis of lesion susceptibility values in participants with hemorrhagic lesions, we have demonstrated how unwanted holes or overdilated voxels in masks can affect calculations of susceptibility values. We have shown that the use of the QSMmask-net can greatly reduce the labor required to generate optimal masks for QSM reconstruction while also enhancing the accuracy of susceptibility values. In future work, we plan to examine the effect of the QSMmask-net on reproducibility, which will be important when studying susceptibility changes in longitudinal studies of neurological disease.

Acknowledgments

The authors also appreciate Megan Griffiths for editing the manuscript. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (RS-2023-NR076864) and the Hankuk University of Foreign Studies Research Fund.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. L. De Rochefort, T. Liu, B. Kressler, et al., “Quantitative Susceptibility Map Reconstruction From MR Phase Data Using Bayesian Regularization: Validation and Application to Brain Imaging,” *Magnetic Resonance in Medicine* 63, no. 1 (2010): 194–206.
2. Q. C. O. Committee, B. Bilgic, M. Costagli, et al., “Recommended Implementation of Quantitative Susceptibility Mapping for Clinical Research in the Brain: A Consensus of the ISMRM Electro-Magnetic Tissue Properties Study Group,” *Magnetic Resonance in Medicine* 91, no. 5 (2024): 1834–1862.
3. J. H. Duyn and J. Schenck, “Contributions to Magnetic Susceptibility of Brain Tissue,” *NMR in Biomedicine* 30, no. 4 (2017): e3546.
4. J. Acosta-Cabronero, G. B. Williams, A. Cardenas-Blanco, R. J. Arnold, V. Lupson, and P. J. Nestor, “In Vivo Quantitative Susceptibility Mapping (QSM) in Alzheimer’s Disease,” *PLoS ONE* 8, no. 11 (2013): e81093.
5. L. Li and J. S. Leigh, “Quantifying Arbitrary Magnetic Susceptibility Distributions With MR,” *Magnetic Resonance in Medicine* 51, no. 5 (2004): 1077–1082.
6. E. M. Haacke, N. Y. Cheng, M. J. House, et al., “Imaging Iron Stores in the Brain Using Magnetic Resonance Imaging,” *Magnetic Resonance Imaging* 23, no. 1 (2005): 1–25.
7. W. Chung, J. Jang, and Y. Nam, “Quantitative Susceptibility Mapping of Oxygen Metabolism: A Feasibility Study Utilizing a Large-Scale Clinical Dataset,” *Investigative Magnetic Resonance Imaging* 27, no. 4 (2023): 221–225.
8. J. Yang, S. Lee, Y. Moon, J. Lee, and W.-J. Moon, “Relationship Between Cortical Iron and Diabetes Mellitus in Older Adults With Cognitive Complaints: A Quantitative Susceptibility Map Study,” *Investigative Magnetic Resonance Imaging* 27, no. 2 (2023): 84–92.
9. S. Ayton, A. Fazlollahi, P. Bourgeat, et al., “Cerebral Quantitative Susceptibility Mapping Predicts Amyloid- β -Related Cognitive Decline,” *Brain* 140, no. 8 (2017): 2112–2119.
10. C. Langkammer, T. Liu, M. Khalil, et al., “Quantitative Susceptibility Mapping in Multiple Sclerosis,” *Radiology* 267, no. 2 (2013): 551–559.
11. Y. Zhang, S. A. Gauthier, A. Gupta, et al., “Longitudinal Change in Magnetic Susceptibility of New Enhanced Multiple Sclerosis (MS) Lesions Measured on Serial Quantitative Susceptibility Mapping (QSM),” *Journal of Magnetic Resonance Imaging* 44, no. 2 (2016): 426–432.
12. A. K. Lotfipour, S. Wharton, S. T. Schwarz, et al., “High Resolution Magnetic Susceptibility Mapping of the Substantia Nigra in Parkinson’s Disease,” *Journal of Magnetic Resonance Imaging* 35, no. 1 (2012): 48–55.
13. H. Sun, A. C. Klahr, M. Kate, et al., “Quantitative Susceptibility Mapping for Following Intracranial Hemorrhage,” *Radiology* 288, no. 3 (2018): 830–839.
14. W. Chen, W. Zhu, I. Kovankaya, et al., “Intracranial Calcifications and Hemorrhages: Characterization With Quantitative Susceptibility Mapping,” *Radiology* 270, no. 2 (2014): 496–505.
15. Y. Zhang, S. A. Gauthier, A. Gupta, et al., “Magnetic Susceptibility From Quantitative Susceptibility Mapping Can Differentiate New Enhancing From Nonenhancing Multiple Sclerosis Lesions Without Gadolinium Injection,” *American Journal of Neuroradiology* 37, no. 10 (2016): 1794–1799.
16. S. Zhang, T. D. Nguyen, Y. Zhao, S. A. Gauthier, Y. Wang, and A. Gupta, “Diagnostic Accuracy of Semiautomatic Lesion Detection Plus Quantitative Susceptibility Mapping in the Identification of New and Enhancing Multiple Sclerosis Lesions,” *NeuroImage: Clinical* 18 (2018): 143–148.
17. B. Wu, W. Li, A. Guidon, and C. Liu, “Whole Brain Susceptibility Mapping Using Compressed Sensing,” *Magnetic Resonance in Medicine* 67, no. 1 (2012): 137–147.
18. Y. Wen, D. Zhou, T. Liu, P. Spincemaille, and Y. Wang, “An Iterative Spherical Mean Value Method for Background Field Removal in MRI,” *Magnetic Resonance in Medicine* 72, no. 4 (2014): 1065–1071.
19. W. Li, B. Wu, and C. Liu, “Quantitative Susceptibility Mapping of Human Brain Reflects Spatial Variation in Tissue Composition,” *NeuroImage* 55, no. 4 (2011): 1645–1656.
20. H. Sun and A. H. Wilman, “Background Field Removal Using Spherical Mean Value Filtering and Tikhonov Regularization,” *Magnetic Resonance in Medicine* 71, no. 3 (2014): 1151–1157.
21. F. Schweser, S. D. Robinson, L. De Rochefort, W. Li, and K. Bredies, “An Illustrated Comparison of Processing Methods for Phase MRI and QSM: Removal of Background Field Contributions From Sources Outside the Region of Interest,” *NMR in Biomedicine* 30, no. 4 (2017): e3604.
22. X. Zhu, Y. Gao, F. Liu, S. Crozier, and H. Sun, “BFRnet: A Deep Learning-Based MR Background Field Removal Method for QSM of the Brain Containing Significant Pathological Susceptibility Sources,” *Zeitschrift für Medizinische Physik* 33, no. 4 (2022): 578–590.
23. A. W. Stewart, S. D. Robinson, K. O’Brien, et al., “QSMxT: Robust Masking and Artifact Reduction for Quantitative Susceptibility Mapping,” *Magnetic Resonance in Medicine* 87, no. 3 (2022): 1289–1300.
24. H. Alloui, M. Sadgal, and A. Elfazziki, “Deep MRI Segmentation: A Convolutional Method Applied to Alzheimer Disease Detection,” *International Journal of Advanced Computer Science and Applications* 10, no. 11 (2019): 365–371.

25. S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Transactions on Medical Imaging* 35, no. 5 (2016): 1240–1251.
26. R. Balakrishnan, M. C. V. Hernández, and A. J. Farrall, "Automatic Segmentation of White Matter Hyperintensities From Brain Magnetic Resonance Images in the Era of Deep Learning and Big Data—A Systematic Review," *Computerized Medical Imaging and Graphics* 88 (2021): 101867.
27. F. Isensee, M. Schell, I. Pflueger, et al., "Automated Brain Extraction of Multisequence MRI Using Artificial Neural Networks," *Human Brain Mapping* 40, no. 17 (2019): 4952–4964.
28. A. Hoopes, J. S. Mora, A. V. Dalca, B. Fischl, and M. Hoffmann, "SynthStrip: Skull-Stripping for Any Brain Image," *NeuroImage* 260 (2022): 119474.
29. S. M. Smith, "Fast Robust Automated Brain Extraction," *Human Brain Mapping* 17, no. 3 (2002): 143–155.
30. P. J. LaMontagne, T. L. Benzinger, J. C. Morris, et al., "OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease," medrxiv, (2019), 2019.12.13.19014902.
31. R. Beare, B. Lowekamp, and Z. Yaniv, "Image Segmentation, Registration and Characterization in R With SimpleITK," *Journal of Statistical Software* 86 (2018): 1–35.
32. Z. Yaniv, B. C. Lowekamp, H. J. Johnson, and R. Beare, "SimpleITK Image-Analysis Notebooks: A Collaborative Environment for Education and Reproducible Research," *Journal of Digital Imaging* 31, no. 3 (2018): 290–303.
33. B. C. Lowekamp, D. T. Chen, L. Ibáñez, and D. Blezek, "The Design of SimpleITK," *Frontiers in Neuroinformatics* 7 (2013): 45.
34. N. Yaghmaie, W. T. Syeda, C. Wu, et al., "QSMART: Quantitative Susceptibility Mapping Artifact Reduction Technique," *NeuroImage* 231 (2021): 117701.
35. P. S. Özbay, A. Deistung, X. Feng, D. Nanz, J. R. Reichenbach, and F. Schweser, "A Comprehensive Numerical Analysis of Background Phase Correction With V-SHARP," *NMR in Biomedicine* 30, no. 4 (2017): e3550.
36. O. Ronneberger, P. Fischer, and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation* (Springer, 2015): 234–241.
37. D. P. Kingma, "Adam: A Method for Stochastic Optimization," arXiv preprint arXiv:14126980, (2014).
38. M. J. Cardoso, W. Li, R. Brown, et al., "Monai: An Open-Source Framework for Deep Learning in Healthcare," arXiv preprint arXiv:221102701, (2022).
39. J. P. Marques, J. Meineke, C. Milovic, et al., "QSM Reconstruction Challenge 2.0: A Realistic In Silico Head Phantom for MRI Data Simulation and Evaluation of Susceptibility Mapping Procedures," *Magnetic Resonance in Medicine* 86, no. 1 (2021): 526–542.
40. M. A. Schofield and Y. Zhu, "Fast Phase Unwrapping Algorithm for Interferometric Applications," *Optics Letters* 28, no. 14 (2003): 1194–1196.
41. W. Li, N. Wang, F. Yu, et al., "A Method for Estimating and Removing Streaking Artifacts in Quantitative Susceptibility Mapping," *NeuroImage* 108 (2015): 111–122.
42. W. Jung, J. Yoon, S. Ji, et al., "Exploring Linearity of Deep Neural Network Trained QSM: QSMnet+," *NeuroImage* 211 (2020): 116619.
43. P. A. Yushkevich, J. Piven, H. C. Hazlett, et al., "User-Guided 3D Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability," *NeuroImage* 31, no. 3 (2006): 1116–1128.
44. S. Wharton and R. Bowtell, "Effects of White Matter Microstructure on Phase and Susceptibility Maps," *Magnetic Resonance in Medicine* 73, no. 3 (2015): 1258–1269.
45. A. Deistung, A. Schäfer, F. Schweser, U. Biedermann, R. Turner, and J. R. Reichenbach, "Toward In Vivo Histology: A Comparison of Quantitative Susceptibility Mapping (QSM) With Magnitude-, Phase-, and R2*-Imaging at Ultra-High Magnetic Field Strength," *NeuroImage* 65 (2013): 299–314.
46. J. Yao, M. A. Morrison, A. Jakary, et al., "Altered Iron and Microstructure in Huntington's Disease Subcortical Nuclei: Insight From 7T MRI," *Journal of Magnetic Resonance Imaging* 60, no. 4 (2024): 1484–1499.
47. S. Straub, X. Zhou, S. Tao, E. M. Westerhold, J. Jin, and E. H. Middlebrooks, "Feasibility of Submillimeter Functional Quantitative Susceptibility Mapping Using 3D Echo Planar Imaging at 7 T," *NMR in Biomedicine* 38, no. 1 (2025): e5263.
48. M. N. Voelker, O. Kraff, S. Goerke, et al., "The Traveling Heads 2.0: Multicenter Reproducibility of Quantitative Imaging Methods at 7 Tesla," *NeuroImage* 232 (2021): 117910.
49. W. Jung, S. Bollmann, and J. Lee, "Overview of Quantitative Susceptibility Mapping Using Deep Learning: Current Status, Challenges and Opportunities," *NMR in Biomedicine* 35, no. 4 (2022): e4292.
50. B. Billot, D. N. Greve, O. Puonti, et al., "SynthSeg: Segmentation of Brain MRI Scans of Any Contrast and Resolution Without Retraining," *Medical Image Analysis* 86 (2023): 102789.
51. J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment Anything in Medical Images," *Nature Communications* 15, no. 1 (2024): 654.