

# Bayesian Learning of Probabilistic Dipole Inversion for Quantitative Susceptibility Mapping

**Jinwei Zhang**<sup>1,2</sup>

JZ853@CORNELL.EDU

<sup>1</sup> *Department of Biomedical Engineering, Cornell University, Ithaca, NY, USA*

<sup>2</sup> *Department of Radiology, Weill Medical College of Cornell University, New York, NY, USA*

**Hang Zhang**<sup>2,3</sup>

HZ459@CORNELL.EDU

**Mert Sabuncu**<sup>1,2,3</sup>

MSABUNCU@CORNELL.EDU

<sup>3</sup> *Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA*

**Pascal Spincemaille**<sup>2</sup>

PAS2018@MED.CORNELL.EDU

**Thanh Nguyen**<sup>2</sup>

TDN2001@MED.CORNELL.EDU

**Yi Wang**<sup>1,2</sup>

YIWANG@MED.CORNELL.EDU

## Abstract

A learning-based posterior distribution estimation method, Probabilistic Dipole Inversion (PDI), is proposed to solve quantitative susceptibility mapping (QSM) inverse problem in MRI with uncertainty estimation. A deep convolutional neural network (CNN) is used to represent the multivariate Gaussian distribution as the approximated posterior distribution of susceptibility given the input measured field. In PDI, such CNN is firstly trained on healthy subjects' data with labels by maximizing the posterior Gaussian distribution loss function as used in Bayesian deep learning. When tested on new dataset without any label, PDI updates the pre-trained CNN's weights in an unsupervised fashion by minimizing the *KullbackLeibler* divergence between the approximated posterior distribution represented by CNN and the true posterior distribution given the likelihood distribution from known physical model and pre-defined prior distribution. Based on our experiments, PDI provides additional uncertainty estimation compared to the conventional MAP approach, meanwhile addressing the potential discrepancy issue of CNN when test data deviates from training dataset.

**Keywords:** Bayesian deep learning, variational inference, convolutional neural network, quantitative susceptibility mapping

## 1. Introduction

Consider the following biomedical imaging model:

$$y = Ax + n \tag{1}$$

where  $A$  the forward imaging system model,  $x$  the underlining biomedical image variable,  $n$  the system noise, and  $y$  the measured data/signal variable. Because of the intrinsic ill-posedness of forward imaging operator  $A$ , prior term is needed in the following Maximum a posteriori (MAP) estimation problem (Kaipio and Somersalo, 2006):

$$\hat{x} = \arg \max_x p(x|y) \propto p(y|x)p(x) \quad (2)$$

where  $p(x)$  is the prior term to regularize the inverse problem. Assuming zero mean Gaussian noise with covariance matrix  $\Sigma$ , Eq. 2 is equivalent to the following minimum  $-\log p(x|y)$  problem:

$$\hat{x} = \arg \min_x \|Ax - y\|_{\Sigma^{-1/2}}^2 + R(x) \quad (3)$$

where  $R(x) = -\log p(x)$ . Convex optimization solvers have been widely used to solve Eq. 3 with both accuracy and efficiency, such as quasi-newton method (Dennis and Moré, 1977), alternating direction method of multipliers (ADMM) (Boyd et al., 2011) and primal dual method (PD) (Chambolle and Pock, 2011).

In recent years, posterior distribution estimation in imaging inverse problems has been a new topic in medical imaging field (Repetti et al., 2019; Chappell et al., 2009; Tezcan et al., 2018), in which random variable’s variance is provided from posterior distribution to measure the uncertainty of the solution. However, posterior distribution estimation requires complicated or even intractable integral from Bayes formula, therefore sampling or approximation method is used to reduce the computational cost and intractability of the problem. Markov chain Monte Carlo (MCMC) (Andrieu et al., 2003) and variational inference (VI) (Bishop, 2006) are two common frameworks in Bayesian estimation problem. In MCMC, efficient sampling methods are used to get random samples from posterior distribution. After proper sampling procedure, random samples can represent an empirical distribution which resembles the true distribution. However, in imaging inverse problem, the computational cost of approximating integrals for Bayesian estimation is often several magnitude higher than the optimization method of MAP estimation, suffering from curse of dimensionality (Pereyra, 2017).

An alternative approach is to use VI, in which an approximation distribution is proposed with specific function form and unknown parameters, and then optimization algorithm is implemented (for example, expectation-maximization (EM) algorithm (Blei et al., 2017)) to learn these parameters by minimizing the divergence between true posterior and approximate posterior. After learning/fitting, the approximate posterior represents the true posterior. However, approximation quality is determined by the flexibility of approximate function form and trainable parameters. More complicated approximate function has better representation ability, however, the computational cost becomes higher. Therefore, the trade-off between number of trainable parameters of approximate function and learning/fitting efficiency needs careful consideration for VI.

Over the past years, thanks to the advances of deep learning, using deep neural network as the approximate function has become a new trend in VI, especially for generative models (Rezende et al., 2014; Kingma and Welling, 2013), in which low dimensional latent space variables are modeled and encoder and decoder networks are built to approximate the latent variable distribution conditioned on observed data and reversely observed data distribution conditioned on latent variable. Due to the approximation and generalization

power of deep neural network with millions of trainable weights, neural network can approximate any function/distribution with high accuracy. In addition, advanced stochastic optimization algorithms such as ADAM (Kingma and Ba, 2014) have been proposed for efficient backpropagation in network weights’ updating. Another topic related to the posterior distribution estimation with deep learning is discussed in Bayesian deep learning framework (Kendall and Gal, 2017), in which data uncertainties are captured by maximizing the posterior distribution with the labels as the samples assuming they follow multivariate Gaussian.

In this paper, we come up with a framework by combining Bayesian deep learning to model data uncertainties and VI with deep learning to approximate true posterior distribution, and apply it to one important imaging inverse problem in MRI: *quantitative susceptibility mapping* (QSM) (de Rochefort et al., 2010; Wang and Liu, 2015), which has the advantages of mapping iron decomposition (Wang et al., 2017) and calcification (Chen et al., 2014). Assuming multivariate Gaussian represented by a CNN as the posterior distribution of susceptibility given the input local field, golden standard susceptibility maps COSMOS (Calculation Of Susceptibility through Multiple Orientation Sampling (Liu et al., 2009)) are used to train such CNN with a maximal posterior loss function. With physical model-based likelihood term and delicately designed prior term, the pre-trained CNN can be enhanced when tested on patient dataset by minimizing the Kullback-Leibler (KL) divergence between true posterior distribution and approximation distribution represented by such CNN. Our experimental results show the proposed method gives mean and variance estimation of the solution automatically, and yields optimal results compared to two types of benchmark methods: deep learning QSM ((Yoon et al., 2018; Zhang et al., 2020)) and *maximum a posteriori* (MAP) QSM with convex optimization (Liu et al., 2012; Kee et al., 2017; Milovic et al., 2018), both of which do not provide uncertainty estimation.

## 2. Method

### 2.1. Modeling

In Magnetic Resonance Imaging (MRI), the forward model of generating relative local field  $b$  from tissue susceptibility  $\chi$  is:

$$b = d * \chi + n \tag{4}$$

where  $*$  denotes convolution operation,  $d$  denotes *dipole kernel*, which is ill-posed inherent in the structure of dipole convolution operator.  $b$  is derived from multi-echo gradient echo MR signal with noise  $n$  estimated as well. The inverse problem of estimating  $\chi$  from measured  $b$  is called *Quantitative Susceptibility Mapping* (QSM). From convolution theory, the forward convolution process in Eq. 4 is equivalent to the following Fourier space multiplication process:

$$b = F^H D F \chi + n \tag{5}$$

where  $F$  is Fourier matrix,  $D$  is dipole kernel in Fourier space. Eq. 5 is computationally friendly since Fast Fourier Transform (FFT) can be used efficiently. We will use forward model in Eq. 5 for computation in this paper.

One successful approach to solving QSM from single orientation field  $b$  is MEDI (Morphology enabled dipole inversion) (Liu et al., 2011b, 2012), where weighted total variation regularization was imposed onto the area except tissue in brain and the following MAP estimation is deployed:

$$\hat{\chi} = \arg \min_{\chi} \|W(F^H DF\chi - b)\|_2^2 + \lambda \|M\nabla\chi\|_1 \quad (6)$$

where  $W$  is derived from observation noise covariance matrix and  $M$  is gradient’s weight to penalize only region’s outside brain tissues. Computational methods for solving Eq. 6 is reviewed in (Kee et al., 2017).

Starting from forward model in Eq. 5, we develop the fully probabilistic model of QSM and use approximate Bayesian inference to solve this problem. We assume conditional distribution of field  $b$  given susceptibility  $\chi$  as a Gaussian distribution:

$$p(b|\chi) = \mathcal{N}(b|F^H DF\chi, \Sigma_{b|\chi}) \quad (7)$$

where we assume  $n \sim \mathcal{N}(0, \Sigma_{b|\chi})$  with  $\Sigma_{b|\chi}$  diagonal in Eq. 5. The prior distribution from Eq. 6 reads:

$$p(\chi) \propto e^{-\lambda \|M\nabla\chi\|_1}. \quad (8)$$

Other types of prior distributions can also be applied. Because of the intractability of estimating the posterior distribution  $p(\chi|b) = p(b|\chi)p(\chi) / \int_{\chi} p(b|\chi)p(\chi)d\chi$  in most cases, approximate posterior distribution  $q(\chi|b) = \mathcal{N}(\mu_{\chi|b}, \Sigma_{\chi|b})$  with diagonal covariance matrix is assumed to approximate the true posterior distribution  $p(\chi|b)$ . In this work, we use a dual-decoder network architecture (Figure 1) extended from 3D U-Net (Ronneberger et al., 2015; Çiçek et al., 2016) to represent the approximate posterior  $q_{\psi}(\chi|b)$ , with each decoder’s output representing mean  $\mu_{\chi|b}$  and variance  $\Sigma_{\chi|b}$  map, respectively.

## 2.2. Supervised Bayesian Training

For training dataset with COSMOS as golden standard labels, we can treat these labels as samples from the true posterior distribution, and train the approximate distribution  $q_{\psi}(\chi|b)$  in a supervised fashion with the following MAP loss function:

$$-\log q_{\psi}(\chi_i|b_i) = \frac{1}{2}(\chi_i - \mu_{\chi|b_i})^T \Sigma_{\chi|b_i}^{-1} (\chi_i - \mu_{\chi|b_i}) + \frac{1}{2} \ln |\Sigma_{\chi|b_i}|, \quad (9)$$

where  $\{b_i, \chi_i\}$  denote the  $i$ -th input and label data pair in the training dataset. Note that because of multiple orientations’ scanning for COSMOS, this dataset is quite limited and usually only on healthy subjects. We denote this supervised Bayesian learning approach as Probabilistic Dipole Inversion (PDI).

## 2.3. Unsupervised Variational Inference

After training using COSMOS data with loss function Eq. 9 and obtaining optimal parameters  $\psi^*$ , given a test local field  $b'$ , we can simply estimate  $p(\chi|b')$  as  $q_{\psi^*}(\chi|b')$ . However, for

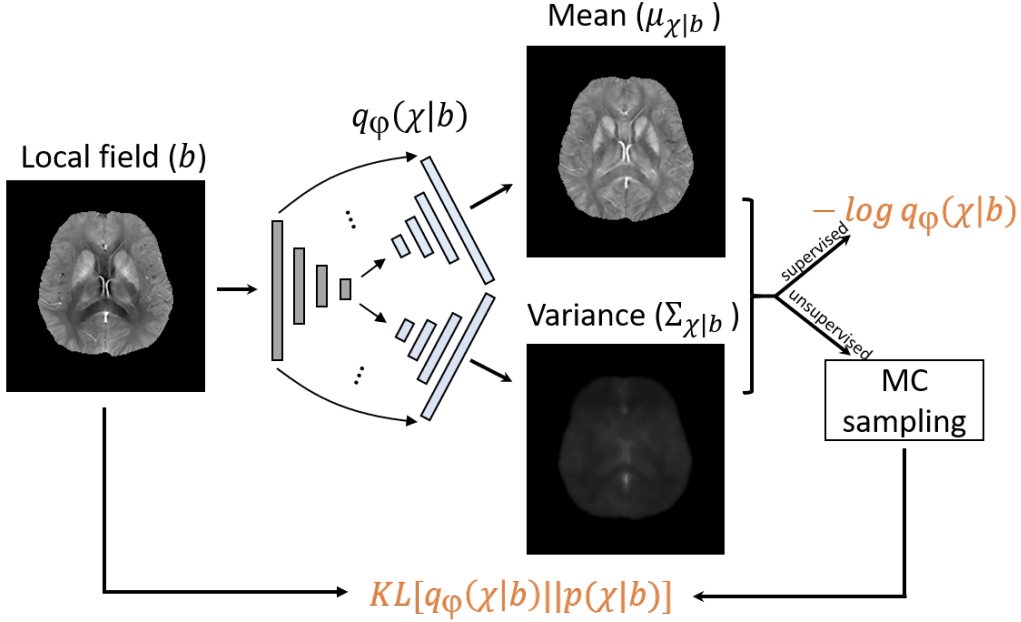


Figure 1: Network architecture of the proposed method. Dual decoders' outputs represent mean and variance maps. COSMOS dataset was used to do supervised Bayesian training via MAP in Eq. 9. Unsupervised VI with MC sampling in Eq. 11 and 12 was applied on other test dataset.

new test dataset which has input field  $b'$  deviating from COSMOS training dataset (such as having new pathologies), inferior outputs could be produced. In this case,  $q_{\psi^*}(\chi|b')$  can be enhanced by deploying variational inference on a subset of this new test dataset as another training set. Specifically, the pre-trained approximation  $q_{\psi}(\chi|b')$  with weights  $\psi$  initialized as  $\psi^*$  can be fine-tuned by minimizing the KL divergence between  $p(\chi|b')$  and  $q_{\psi}(\chi|b')$ :

$$\begin{aligned}
 & \text{KL}[q_{\psi}(\chi|b')||p(\chi|b')] \\
 &= \mathbb{E}_q[\log q_{\psi}(\chi|b') - \log p(\chi|b')] \\
 &= \mathbb{E}_q[\log q_{\psi}(\chi|b') - \log p(\chi, b')] + \log p(b') \\
 &= \text{KL}[q_{\psi}(\chi|b')||p(\chi)] - \mathbb{E}_q[\log p(b'|\chi)]
 \end{aligned} \tag{10}$$

where the first term in the last equation above imposes the posterior to be similar to the prior, and the second term encourages data consistency in QSM forward model. Applying the prior term defined in Eq. 8 and likelihood term in Eq. 7, KL divergence in Eq. 10 becomes:

$$\begin{aligned}
 & \text{KL}[q_{\psi}(\chi|b')||p(\chi|b')] \\
 &= -\frac{1}{2}\ln|\Sigma_{\chi|b'}| + \frac{1}{2K} \sum_{k=1}^K \lambda \|M \nabla_{\chi_k}\|_1 + \frac{1}{2K} \sum_{k=1}^K (\chi_k * d - b')^T \Sigma_{b'|\chi}^{-1} (\chi_k * d - b')
 \end{aligned} \tag{11}$$

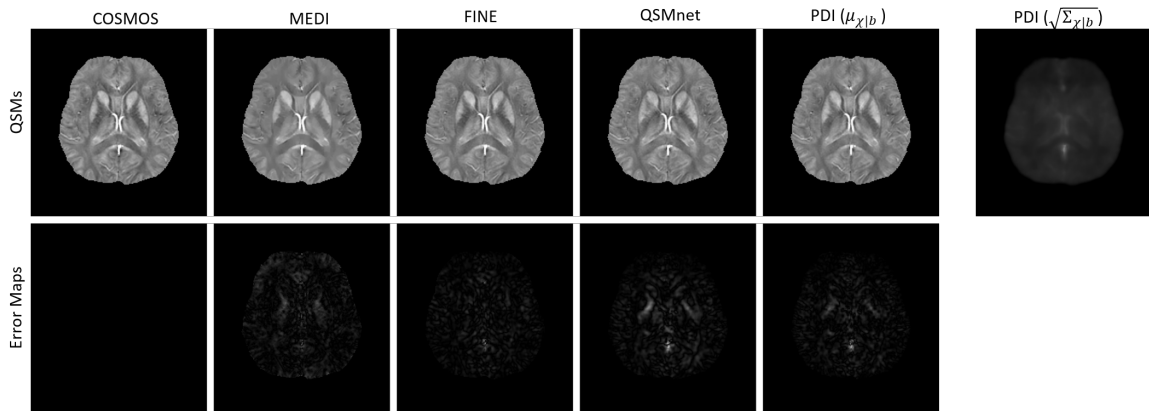


Figure 2: Reconstructions (first row) and error maps (second row) of one COSMOS test subject in one orientation, with COSMOS as the golden standard. FINE gives the best reconstruction at the expense of significantly increased computational time. The other three methods have comparable results. The standard deviation map (last column) provided by PDI resembled its error map, with high uncertainties/errors locating at sagittal sinus and globus pallidus.

where  $-\mathbb{E}_q[\ln p(\chi)]$  in  $\text{KL}[q_\psi(\chi|b')||p(\chi)]$  and  $-\mathbb{E}_q[\log p(b'|\chi)]$  are calculated through Monte Carlo (MC) sampling with  $\chi_k$  sampled from  $q_\psi(\chi|b')$  because of the intractability of both expectations. We denote the fine-tuned approximate distribution with Eq. 11 as PDI-VII.

Another possible prior term for  $\chi$  is simply constant prior  $p(\chi) \propto c$ , which means no prior information is given regarding the distribution of  $\chi$ . With such non-informative prior, the corresponding loss function is simply:

$$\text{KL}[q_\psi(\chi|b')||p(\chi|b')] = -\frac{1}{2}\ln|\Sigma_{\chi|b'}| + \frac{1}{2K} \sum_{k=1}^K (\chi_k * d - b')^T \Sigma_{b'|\chi}^{-1} (\chi_k * d - b') \quad (12)$$

We denote the fine-tuned approximate distribution with Eq. 12 as PDI-VI2.

### 3. Experiments

MRI was performed on 7 healthy subjects with 5 brain orientations using a 3T GE scanner equipped with a multi-echo 3D gradient echo (GRE) sequence. Acquisition matrix was  $256 \times 256 \times 48$  and voxel size was  $1 \times 1 \times 3 \text{ mm}^3$ . The input local tissue field data  $b$  was generated by sequentially deploying non-linear fitting across multi-echo phase data (Kressler et al., 2009), graph-cut based phase unwrapping (Dong et al., 2014) and background field removal (Liu et al., 2011a). COSMOS reconstruction (Liu et al., 2011b) was calculated from 5 orientations' GRE imaging and was used as the gold standard label in the experiment. A second dataset was obtained by performing single orientation GRE MRI on 8 patients with intracerebral hemorrhage (ICH), which were acquired using the same scanner and imaging parameters as above.

Network architecture is shown in Figure 1. Dual decoders' outputs were used to represent mean and variance maps in the posterior susceptibility distribution given input local field. The 3D convolutional kernel size was  $3 \times 3 \times 3$ . The numbers of filters from the highest

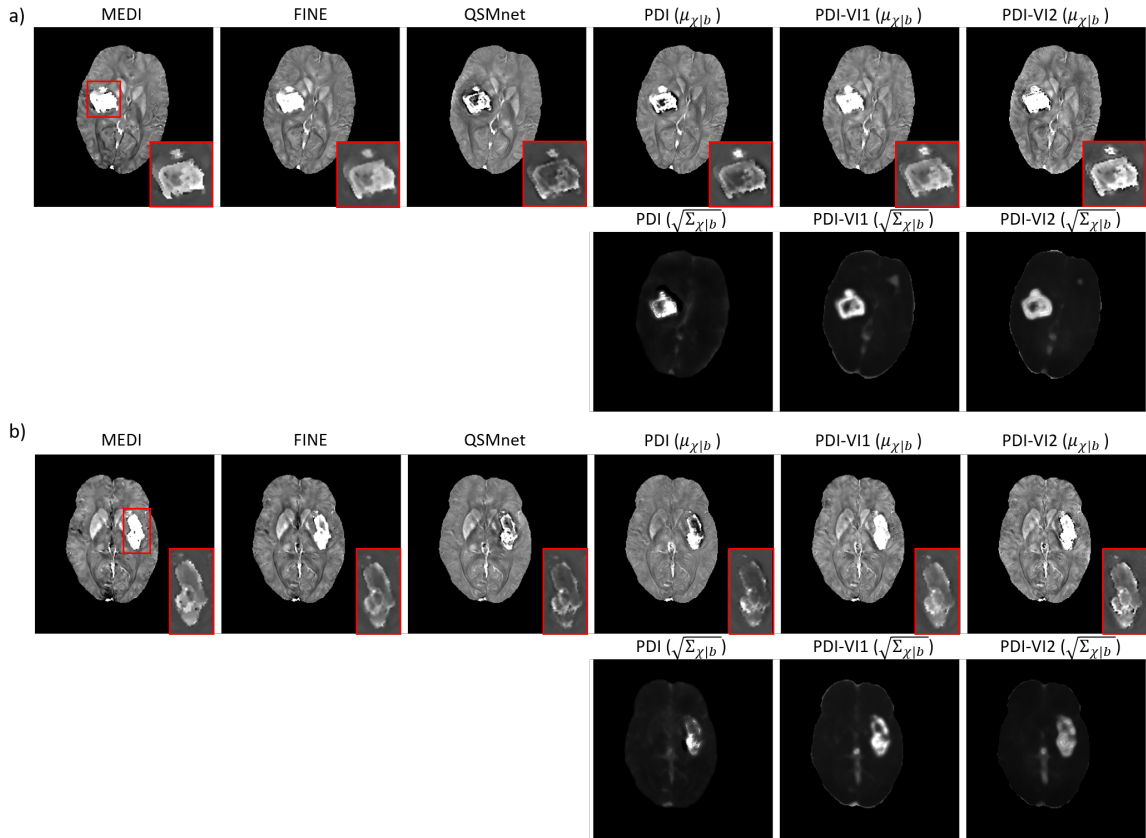


Figure 3: Reconstructions (first row in (a) and (b)) and standard deviation maps (second row in (a) and (b)) of two ICH patients. Compared to MEDI and FINE, underestimation issue inside hemorrhage happened on QSMnet and PDI. This issue was reduced in PDI-VI1 and PDI-VI2 by fine-tuning the pre-trained network using unsupervised variational inference. High variance inside the hemorrhage was consistent with high measured noise in the same region.

feature level to the lowest were 32, 64, 128, 256 and 512, respectively. Batch normalization (Ioffe and Szegedy, 2015), max pooling for downsampling and deconvolution operation for upsampling were used. For COSMOS dataset, 4/1 subjects (20/5 brain volumes) were used as training/validation dataset, with augmentation by in-plane rotation of  $\pm 15^\circ$ . Each brain volume data in the training and validation dataset was divided into 3D patches with patch size  $64 \times 64 \times 32$  and extraction step  $21 \times 21 \times 11$ . The remaining 2 subjects (10 brain volumes in total) were used for testing. For ICH patients dataset, 5/1 subjects were used as training/validation dataset for PDI-VI1 and PDI-VI, and the remaining 2 subjects were used for testing.

Loss function in Eq. 9 was applied for supervised Bayesian training on COSMOS dataset with ADAM optimizer (Kingma and Ba, 2014) (learning rate:  $10^{-3}$ , Number of epochs: 60), yielding trained network  $q_{\psi^*}(\chi|b)$ . The outputs of  $q_{\psi^*}(\chi|b)$  were denoted as PDI. Initialized with the pre-trained PDI using COSMOS training data, unsupervised variational inference with loss function Eq. 11 and 12 was also applied on ICH dataset using ADAM optimizer (learning rate:  $10^{-3}$ , Number of iterations: 100). MC sampling size  $K$  was chosen as 5 due



	pSNR	RMSE	SSIM	HFEN	GPU time (s)
MEDI (Liu et al., 2012)	46.39	41.16	0.9569	31.30	17.54
FINE (Zhang et al., 2020)	48.12	33.66	0.9789	31.97	65.42
QSMnet (Yoon et al., 2018)	46.35	41.29	0.9705	43.31	0.60
PDI (Eq. 9)	47.77	35.08	0.9772	35.17	0.61

Table 1: Mean quantitative metrics of 10 test COSMOS brains reconstructed by different methods. FINE gives the best reconstruction at the expense of significantly increased computational time. The other three methods have comparable results.

to limited GPU memory and reparameterization trick (Kingma and Welling, 2013) was used for MC sampling in order to do backpropagation. The outputs were denoted as PDI-VI1 and PDI-VI2, respectively. The whole brain volume was fed into the network during testing, including unsupervised variational inference step. We implemented the proposed method using PyTorch (Python 3.6) on an RTX 2080Ti GPU.

For COSMOS test dataset, we compared PDI with MAP estimation MEDI (Liu et al., 2012) and two deep learning reconstructions QSMnet (Yoon et al., 2018) and FINE (Zhang et al., 2020). Reconstruction maps of one orientation from one test subject are shown in Figure 2 ([-0.15ppm, 0.15ppm]). Quantitative metrics of each reconstruction method averaged among 10 test brains are shown in Table 1. FINE gave the best overall quantitative results; However, it overfitted to every test case by minimizing the fidelity loss, which had the major drawback of significantly increased computational time. PDI gave slightly better results than MEDI and QSMnet, meanwhile achieved fast inference time on GPU comparable to QSMnet. In figure 2, error map of PDI’s mean output  $\mu_{\chi|b}$  was coincident with PDI’s standard deviation output  $\sqrt{\Sigma_{\chi|b}}$ , with high uncertainty/error happening at sagittal sinus and globus pallidus.

For ICH test dataset, PDI-VI1 and PDI-VI2 were also performed and compared. Two representative ICH patients’ QSMs are shown in Figure 3 ([-0.6ppm, 1.5ppm] for zoomed-in hemorrhage). Compared to MEDI and FINE which had hyperintensity inside the hemorrhage, both QSMnet and PDI suffered from underestimation issue inside this region, which might result from the fact that such pathology was not encountered during training since long scan COSMOS was not practical for the patients. After PDI-VI1 and PDI-VI2, such underestimation issue was reduced and variance maps’ structures inside the hemorrhage were also better depicted. High uncertainties inside hemorrhage as shown in Figure 3 were consistent with high local field noise level which was approximately proportional to the underlining susceptibility values.

#### 4. Conclusion

We developed a Bayesian dipole inversion framework for quantitative susceptibility mapping by combining variational inference and Bayesian deep learning. Our method generated high fidelity susceptibility maps meanwhile provided uncertainty quantifications. When applied to other datasets not encountered during training, the proposed method was able to correct the undesirable outputs in an unsupervised fashion based on variational inference principle.



## Acknowledgments

This research was supported in part by National Institute of Health (R01NS090464, R01NS095562, R01NS105144, R01DK116126, R01CA181566, S10OD021782, 1R21AG050122, R01LM012719 and R01AG053949), National Science Foundation (1748377 and 1707312) and National Multiple Sclerosis Society (RR-1602-07671).

## References

- Christophe Andrieu, Nando De Freitas, Arnaud Doucet, and Michael I Jordan. An introduction to mcmc for machine learning. *Machine learning*, 50(1-2):5–43, 2003.
- Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.
- David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017.
- Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.
- Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40(1):120–145, 2011.
- Michael A Chappell, Adrian R Groves, Brandon Whitcher, and Mark W Woolrich. Variational bayesian inference for a nonlinear forward model. *IEEE Transactions on Signal Processing*, 57(1):223–236, 2009.
- Weiwei Chen, Wenzhen Zhu, Ilhami Kovanlikaya, Arzu Kovanlikaya, Tian Liu, Shuai Wang, Carlo Salustri, and Yi Wang. Intracranial calcifications and hemorrhages: characterization with quantitative susceptibility mapping. *Radiology*, 270(2):496–505, 2014.
- Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.
- Ludovic de Rochefort, Tian Liu, Bryan Kressler, Jing Liu, Pascal Spincemaille, Vincent Lebon, Jianlin Wu, and Yi Wang. Quantitative susceptibility map reconstruction from mr phase data using bayesian regularization: validation and application to brain imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 63(1):194–206, 2010.
- John E Dennis, Jr and Jorge J Moré. Quasi-newton methods, motivation and theory. *SIAM review*, 19(1):46–89, 1977.

- Jianwu Dong, Tian Liu, Feng Chen, Dong Zhou, Alexey Dimov, Ashish Raj, Qiang Cheng, Pascal Spincemaille, and Yi Wang. Simultaneous phase unwrapping and removal of chemical shift (spurs) using graph cuts: application in quantitative susceptibility mapping. *IEEE transactions on medical imaging*, 34(2):531–540, 2014.
- Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.
- Jari Kaipio and Erkki Somersalo. *Statistical and computational inverse problems*, volume 160. Springer Science & Business Media, 2006.
- Youngwook Kee, Zhe Liu, Liangdong Zhou, Alexey Dimov, Junghun Cho, Ludovic De Rochefort, Jin Keun Seo, and Yi Wang. Quantitative susceptibility mapping (qsm) algorithms: mathematical rationale and computational implementations. *IEEE Transactions on Biomedical Engineering*, 64(11):2531–2545, 2017.
- Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? In *Advances in neural information processing systems*, pages 5574–5584, 2017.
- Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- Bryan Kressler, Ludovic De Rochefort, Tian Liu, Pascal Spincemaille, Quan Jiang, and Yi Wang. Nonlinear regularization for per voxel estimation of magnetic susceptibility distributions from mri field maps. *IEEE transactions on medical imaging*, 29(2):273–281, 2009.
- Jing Liu, Tian Liu, Ludovic de Rochefort, James Ledoux, Ildar Khalidov, Weiwei Chen, A John Tsiouris, Cynthia Wisnieff, Pascal Spincemaille, Martin R Prince, et al. Morphology enabled dipole inversion for quantitative susceptibility mapping using structural consistency between the magnitude image and the susceptibility map. *Neuroimage*, 59(3):2560–2568, 2012.
- Tian Liu, Pascal Spincemaille, Ludovic De Rochefort, Bryan Kressler, and Yi Wang. Calculation of susceptibility through multiple orientation sampling (cosmos): a method for conditioning the inverse problem from measured magnetic field map to susceptibility source image in mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 61(1):196–204, 2009.
- Tian Liu, Ildar Khalidov, Ludovic de Rochefort, Pascal Spincemaille, Jing Liu, A John Tsiouris, and Yi Wang. A novel background field removal method for mri using projection onto dipole fields. *NMR in Biomedicine*, 24(9):1129–1136, 2011a.
- Tian Liu, Jing Liu, Ludovic De Rochefort, Pascal Spincemaille, Ildar Khalidov, James Robert Ledoux, and Yi Wang. Morphology enabled dipole inversion (medi) from

- a single-angle acquisition: comparison with cosmos in human brain imaging. *Magnetic resonance in medicine*, 66(3):777–783, 2011b.
- Carlos Milovic, Berkin Bilgic, Bo Zhao, Julio Acosta-Cabronero, and Cristian Tejos. Fast nonlinear susceptibility inversion with variational regularization. *Magnetic resonance in medicine*, 80(2):814–821, 2018.
- Marcelo Pereyra. Maximum-a-posteriori estimation with bayesian confidence regions. *SIAM Journal on Imaging Sciences*, 10(1):285–302, 2017.
- Audrey Repetti, Marcelo Pereyra, and Yves Wiaux. Scalable bayesian uncertainty quantification in imaging inverse problems via convex optimization. *SIAM Journal on Imaging Sciences*, 12(1):87–118, 2019.
- Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. *arXiv preprint arXiv:1401.4082*, 2014.
- Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- Kerem C Tezcan, Christian F Baumgartner, Roger Luechinger, Klaas P Pruessmann, and Ender Konukoglu. Mr image reconstruction using deep density priors. *IEEE transactions on medical imaging*, 2018.
- Yi Wang and Tian Liu. Quantitative susceptibility mapping (qsm): decoding mri data for a tissue magnetic biomarker. *Magnetic resonance in medicine*, 73(1):82–101, 2015.
- Yi Wang, Pascal Spincemaille, Zhe Liu, Alexey Dimov, Kofi Deh, Jianqi Li, Yan Zhang, Yihao Yao, Kelly M Gillen, Alan H Wilman, et al. Clinical quantitative susceptibility mapping (qsm): biometal imaging and its emerging roles in patient care. *Journal of Magnetic Resonance Imaging*, 46(4):951–971, 2017.
- Jaeyeon Yoon, Enhao Gong, Itthi Chatnuntawech, Berkin Bilgic, Jingu Lee, Woojin Jung, Jingyu Ko, Hosan Jung, Kawin Setsompop, Greg Zaharchuk, et al. Quantitative susceptibility mapping using deep neural network: Qsmnet. *Neuroimage*, 179:199–206, 2018.
- Jinwei Zhang, Zhe Liu, Shun Zhang, Hang Zhang, Pascal Spincemaille, Thanh D Nguyen, Mert R Sabuncu, and Yi Wang. Fidelity imposed network edit (fine) for solving ill-posed image reconstruction. *NeuroImage*, page 116579, 2020.