BayeSMM: Robust Deep Combined Computing Tackling Heavy-tailed Distribution in Medical Images

Yuanye Liu[1†], Ruoxuan Zhen[1†], Shangqi Gao[2], Xinzhe Luo[3], Xin Gao[1], Qingchao Chen[4✉], and Xiahai Zhuang[1✉]

[1] Fudan University, Shanghai, China
[2] University of Cambridge, Cambridge, United Kingdom
[3] Imperial College London, London, United Kingdom
[4] Peking University, Beijing, China

Abstract. Abnormal structures in multi-modality medical images often lead to heterogeneous heavy-tailed distributions. However, traditional models, especially those relying on Gaussian distributions, struggle to effectively capture these outliers. To address this, we propose BayeSMM, a novel framework that leverages Student's t distribution mixture models (SMM) to simultaneously perform registration and segmentation for misaligned multi-modality medical images. Specifically, we construct a Bayesian Student's t mixture model incorporating the heavy-tailed nature of the Student's t distribution and develop variational inference to optimize the model. Guided by variational inference, we design a novel deep learning architecture that performs registration and segmentation jointly. We demonstrate the effectiveness of BayeSMM with experiments on the MS-CMR dataset, where the results show superior performance compared to existing combined computing methods, and yield enhanced robustness under the simulated heavy-tailed setting. The code is available at https://github.com/HenryLau7/BayeSMM.

Keywords: Bayesian statistical modeling · Medical image combined computing · Robustness · Student's t distribution

1 Introduction

Jointly performing registration and segmentation, also termed as, combined computing [22], could benefit each other, and have gained popularity in multi-modality scenarios [2, 4, 13, 16, 22]. Early methods addressed this problem via Maximum a Posteriori (MAP) estimation within probabilistic frameworks [20, 21]. Bayesian frameworks [1, 6] established principled generative models that jointly perform segmentation and registration. By using the conditional independence of multi-modality image intensity distribution given the tissue label, [4, 13, 22] have the advantage of computational efficiency, in a Bayesian

† These two authors contributed equally.
✉ Corresponding authors: Qingchao Chen, Xiahai Zhuang (zxh@fudan.edu.cn)

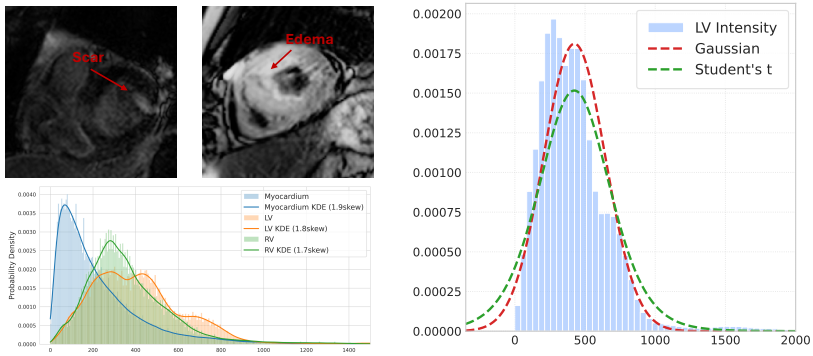


Fig. 1. (a) **Top:** In cardiac MR images, scar and edema may appear, influencing the distribution of anatomical structures. **Bottom:** The distribution histogram illustrates the heavy-tailed nature of an LGE modality image. (b) The Student's t distribution, with its heavier tail, is more robust to outliers, such as those caused by abnormal structures, exemplified by the LV tissue in the LGE modality.

framework. Deep learning-based methods substantially improved the performance [11, 19]. However, they only focus on localized pathological structures, such as tumors, limiting their applicability in holistic segmentation scenarios.

An unresolved challenge in multi-modality medical imaging is the heavy-tailed distributions arising from abnormal structures [8]. In cardiac MR images, as illustrated in Fig. 1(a), scars in LGE (Late Gadolinium Enhanced) and edema in T2-weighted MR are rare yet significant pathological features resulting in extreme values in intensity distributions, while they are not consistent in different modalities. This could influence learning processes and model fitting [7, 15]. Previous parametric models, which typically rely on Gaussian distributions, fail to capture these outliers.

Motivated by the heavy-tailed nature of Student's t distribution [18], which is more robust to model medical images with abnormal structures, as shown in Fig. 1(b), we propose BayeSMM, a novel approach that models misaligned multi-modality medical images with a Student's t mixture model, which could model multi-modality heterogeneous abnormal outliers, to perform registration and segmentation in an end-to-end unified framework. In BayeSMM, we employ variational inference [3] and propose a deep computation framework to estimate the variables. This combined approach allows us to effectively learn from misaligned multi-modality data and enhance its robustness.

Our key contributions are as follows: (1) We are the first to introduce a method for modeling multi-modality medical images using a Student's t mixture model, which is robustness with abnormal outliers. (2) We employ variational inference and propose a novel deep learning framework that simultaneously performs both image registration and segmentation. (3) Our experimental results demonstrate the superiority of BayeSMM in registration and segmentation, and

we explore the heavy-tailed nature of Student's t distribution through a visualization analysis for our defined outlier-detection variable.

2 Methodology

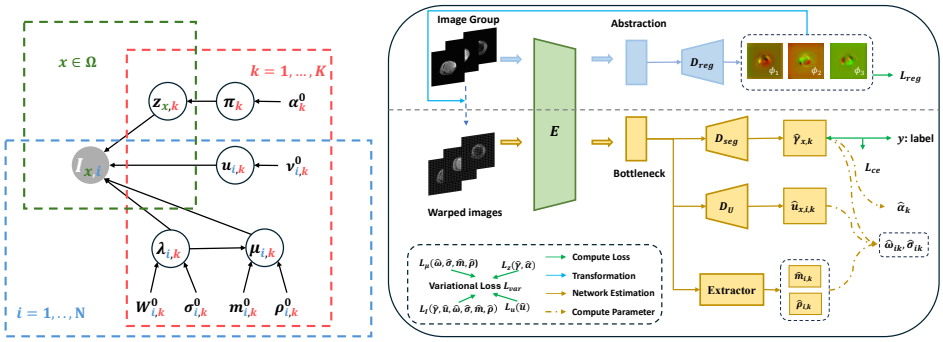
This work introduces a robust Bayesian Student's t Mixture Model (BayeSMM) for the simultaneous registration and segmentation of multi-modality medical images, with a particular focus on addressing the heavy-tail problem often encountered in medical image analysis. The proposed framework consists of two main components: (1) statistical modeling for the multi-modality medical images using the Student's t distribution, as illustrated in Fig. 2.1(a), and (2) a deep learning-based registration and segmentation framework for estimating posteriors of the variables introduced in statistical modeling, as depicted in Fig. 2.1(b). At step (1), we model the multi-modality images via a multivariate BayeSMM, incorporating a class variable $\boldsymbol{z}$, its prior probability $\boldsymbol{\pi}$, a defined outlier-detection variable $\boldsymbol{u}$, the mean $\boldsymbol{\mu}$ and inverse-variance $\boldsymbol{\lambda}$ variables. These variables are governed by conjugate priors, determining the observed images $\boldsymbol{I}$. We define $\boldsymbol{u}$ as the outlier-detection variable, which could adaptively adjust the variance to enhance the robustness. Given observed multi-modality images $\boldsymbol{I}$, we propose an alternating approach consisting of learning-based inference and analytical computation, to estimate posteriors at step (2). This alternating approach separates dependencies between variables, enabling effective estimation by leveraging the representational power of neural networks.

2.1 Robust Bayesian Student's t Mixture Model

Mixture Model. Let $\boldsymbol{I} = \{\boldsymbol{I}_i\}_{i=1}^{N}$ denotes the set of N observed images, all acquired from the same subject but in different imaging modalities. Following the framework proposed by [22], we assume the existence of a common latent space, $\Omega \subset \mathbb{R}^d$, from which all image modalities are generated, where d represents the spatial dimensionality of the images. The anatomical structure of the image is modeled using a multinomial variable $\boldsymbol{z} = \{\boldsymbol{z}_x\}_{x\in\Omega}$, where each $\boldsymbol{z}_x = [z_{x,1}, \ldots, z_{x,K}]^\top$ is a one-hot vector, with $z_{x,k} = 1$ indicating that the pixel x belongs to the k-th tissue class, with K classes in total. For a given modality i, we model the anatomical structure as a finite mixture model [14], where the likelihood of an image pixel $I_{x,i}$ conditioned on the tissue label $\boldsymbol{z}_x$ is given by

$$p(I_{x,i} \mid \boldsymbol{z}, \boldsymbol{\theta}) = \prod_{k=1}^{K} p(I_{x,i} \mid \boldsymbol{\theta}_{i,k})^{z_{x,k}}, \tag{1}$$

where $\boldsymbol{\theta}$ represents the parameters of the mixtures, with a corresponding multinomial prior for the tissue label distribution $p(\boldsymbol{z}_x \mid \boldsymbol{\pi}) = \prod_{k=1}^{K} \pi_k^{z_{x,k}}$, where $\boldsymbol{\pi}$ represents the label proportions, with $\sum_k \pi_k = 1$. When the tissue label of a pixel is known, the intensities from different modalities become conditionally independent [4], and the likelihood of the image pixel simplifies to $p(\boldsymbol{I}_x \mid z_{x,k} = 1, \boldsymbol{\theta}) = \prod_i p(I_{x,i} \mid z_{x,k} = 1, \boldsymbol{\theta}_i)$.



(a)PGM (b)Deep learning framework

Fig. 2. Overview of the BayeSMM framework. (a) The probabilistic graphical model (PGM) of BayeSMM, illustrating statistical modeling of multi-modality images. (b) The deep learning framework for BayeSMM, consisting of two stages: the registration stage and the segmentation stage, sharing an encoder. The warped images after registration go through segmentation network and other branches for estimating posteriors.

Student's t Mixture Model. To address the heavy-tail issue that commonly arises in medical image analysis, we introduce the Student's t distribution, with an additional hyperparameter $\boldsymbol{\nu}^0$, *i.e.*, degree of freedom, enhancing its flexibility, to model the intensity distributions of each tissue class. It could be derived by compounding a Gaussian distribution with an outlier-detection variable $\boldsymbol{u}$, which is itself distributed according to a gamma distribution. Integrating over $\boldsymbol{u}$ yields the Student's t distribution, expressed as follows,

$$St(I_{x,i} \mid z_{x,k}=1, \theta_{k,i}) = \int \mathcal{N}\left(I_{x,i} \mid \mu_{k,i}, \frac{1}{u_{x,i,k}\lambda_{k,i}}\right) \mathcal{G}(u_{x,i,k} \mid \frac{\nu^0_{k,i}}{2}, \frac{\nu^0_{k,i}}{2})\, \mathrm{d}u_{x,i,k}, \quad (2)$$

where $\mathcal{N}$ represents a normal distribution and $\mathcal{G}$ a Gamma distribution. The prior parameter $\boldsymbol{\nu}^0$ is manually chosen, and its posterior is implicitly updated through approximate inference.

Hierarchical Bayesian Modeling. Given the hierarchical nature of the Student's t distribution derived from Eq.(2), we model the intensity data within a PGM, as shown in Fig. 2.1(a). The likelihood of the observed images $\boldsymbol{I}$, given the tissue class variable $\boldsymbol{z}$, outlier-detection variable $\boldsymbol{u}$, mean $\boldsymbol{\mu}$, and inverse variance $\boldsymbol{\lambda}$, is given by,

$$p(\boldsymbol{I} \mid \boldsymbol{z}, \boldsymbol{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \prod_{x\in\Omega}\prod_{i}\prod_{k} \mathcal{N}\left(I_{x,i} \mid \mu_{i,k}, (u_{x,i,k}\lambda_{i,k})^{-1}\right)^{z_{x,k}}, \quad (3)$$

where $\{\mu_{i,k}, \boldsymbol{u}_{i,k}, \lambda_{i,k}\}$ are the parameters $\boldsymbol{\theta}_{i,k}$ in Eq.(1), representing the parameters of Student's t distribution.

As for the prior selection, we place a Normal-Gamma prior on the mean $\mu_{i,k}$ and inverse variance $\lambda_{i,k}$, and a Dirichlet prior is placed on the label proportions $\boldsymbol{\pi}$. Conjugate priors are chosen to simplify posterior estimation. This hierarchical probabilistic model enables the integration of intensity values from multiple

modalities, accounting for the complex distributions of medical image intensities while effectively addressing the heavy-tail issue inherent in such data.

2.2 Variational Inference

In this section, we present a variational inference approach to estimate the variables of SMM, given the observed image set $\boldsymbol{I}$. We perform maximum a posteriori (MAP) estimation for the set of variables $\boldsymbol{\Psi} = \{\boldsymbol{z}, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{u}\}$.

Since the exact inference of the posterior distribution $p(\boldsymbol{\Psi} \mid \boldsymbol{I})$ is intractable, we employ a variational Bayesian (VB) method [5] to approximate the posterior. To ensure tractable inference with efficient updates under conjugate priors, we adopt the mean-field assumption [3], which assumes that the variables in $\boldsymbol{\Psi}$ are mutually independent. Thus, we factorize the variational distribution as

$$q(\boldsymbol{\Psi}) = q(\boldsymbol{z})q(\boldsymbol{\pi})q(\boldsymbol{\mu})q(\boldsymbol{\lambda})q(\boldsymbol{u}). \tag{4}$$

Since we assign the conjugate priors, their variational posteriors would have the same forms as given priors. And they have a closed-form solution derived by,

$$\ln q^*\left(\boldsymbol{\Psi}_j\right) = \mathbb{E}_{\boldsymbol{\Psi}_{-j}}[\ln p(\boldsymbol{I}, \boldsymbol{\Psi})] + \text{const.} \tag{5}$$

where $\boldsymbol{\Psi}_{-j}$ represents all variables except $\boldsymbol{\Psi}_{\boldsymbol{j}}$. Variables $\boldsymbol{\pi}$ and $\boldsymbol{\lambda}$ are computed explicitly by Eq. (5).

To optimize other variational variables, *i.e.*, $\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{u}$, we minimize the Kullback-Leibler (KL) divergence between the variational distribution $q(\boldsymbol{\Psi})$ and the true posterior $p(\boldsymbol{\Psi} \mid \boldsymbol{I})$, which results in the following optimization problem,

$$\min_{q(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{u})} \mathcal{L}_{\text{var}} = KL(q(\boldsymbol{\Psi}) \mid p(\boldsymbol{\Psi})) - \mathbb{E}[\ln p(\boldsymbol{I} \mid \boldsymbol{\Psi})]. \tag{6}$$

This optimization leads to the following variational loss, which corresponds to the individual components of the lower bound, $L_{var} = L_I + L_u + L_\mu + L_z$. Each of these components is derived through the variational formulation.

2.3 Deep Computing and Training Strategy

This section describes the network architecture used to achieve variational inference and the training strategy for medical image combined computing. As shown in Fig. 2.1(b), it consists of two stages, *i.e.*, registration and segmetnation. We adopt an Encoder-Decoder architecture, with the SwinUNETR [9] as the backbone, which consists of multiple levels of convolutional blocks to generate multi-level feature maps, with residual connections between the encoder and decoder to facilitate information flow. The encoder is shared across registration and segmentation to extract features.

Registration Stage. The registration stage begins by fusing the extracted features from each modality through an abstraction layer that computes the first and second moments of the feature maps [10]. The fused feature maps are then

passed through the registration decoder, D_{reg}, which predicts the spatial transformation in the form of dense displacement fields $\boldsymbol{\phi} = \{\phi_i\}$ for each modality.
Segmentation Stage. Following the registration step, the estimated transformations $\boldsymbol{\phi}$ are applied to the images to register them to the common space. The warped images are then fed into a segmentation network to estimate the variational variables (or their posterior parameters). Specifically, for class variable $\boldsymbol{z} \sim \text{Mult}(\hat{\boldsymbol{\gamma}})$, where posterior parameters of the Multinomial distribution $\hat{\gamma}_{x,k}$ represents the predicted class probability, and for the outlier-detection variable $\hat{\boldsymbol{u}}$, we employ decoders D_{seg} and D_U to estimate, respectively. For the posterior parameters $\hat{\boldsymbol{m}}, \hat{\boldsymbol{\rho}}$ of the mean values $\boldsymbol{\mu}$, we use an extractor block consisting of convolutional and linear layers to estimate them.
Training Strategy. We train the BayeSMM by balancing the cross-entropy loss L_{CE}, the variational loss L_{var} and the regularization term L_{reg} computed by bending energy [17] as $L = L_{CE} + \lambda_1 L_{var} + \lambda_2 L_{reg}$, where λ_1, λ_2 are balancing weights that determines the contribution of each loss. Cross-entropy loss is expressed as,

$$L_{CE} = \sum_{i \in \mathcal{I}} \frac{1}{|\Omega|} \sum_{x \in \Omega} -y_{x,i} \log \hat{\gamma}_{x,i}, \tag{7}$$

where $\mathcal{I}$ is the index set of modalities with labels, $\boldsymbol{y}$ represents the ground truth. This means that BayeSMM could address the weakly supervised scenario, where the labels of certain modalities are provided. This combined loss function ensures that the network learns both the segmentation task and the variational posterior inference effectively, facilitating the joint optimization of registration and segmentation tasks within the same framework.

3 Experiments

3.1 Experiment Setups

Dataset. We used a cardiac dataset to evaluate BayeSMM. The MS-CMR dataset provides multi-modality cardiac MR images for 45 patients [22]. Each patient was scanned with three sequences: LGE, bSSFP, and T2-weighted MR, containing 3-5 short-axis slices. We resampled the images to $1 \times 1mm^2$ resolution, and center-cropped to 256×256 pixels, background suppression and z-score normalization was employed. For all experiments, a three-fold cross-validation strategy was employed.
Baseline methods. We compared BayeSMM with several baselines. DGR uses a generic framework where no segmentation masks are provided, reducing the problem to deep groupwise registration with neural network estimation. MvMM [22] uses iterative optimization for combined computing. $\mathcal{X}$-metric [13] estimates using a deep neural network. For fairness of comparison, all methods employed the same settings.
Implementation details. We implemented BayeSMM using a SwinUNETR backbone for encoder-decoders. The model was trained for 50 epochs using the Adam optimizer [12] with an initial learning rate of 3×10^{-4}, and a batch size

Table 1. Comparison of joint registration and segmentation performance on multi-modality cardiac MRI. Results show DSC for registration (Reg DSC) and segmentation (Seg DSC) across three sequences under two scenarios: without affine preprocessing (top) and with affine preprocessing (bottom). BayeSMM variants are evaluated with synthetic noise levels (0%/50%/80%), demonstrating robustness compared to ablated version (w/o) and baselines. Best results in each column are bolded.

Method	Noise	Reg DSC	Seg DSC		
			LGE	bSSFP	T2
Before Affine Pre-process					
Init	0	34.34 ± 0.71	-	-	-
DGR	0	$\mathbf{67.93 \pm 0.34}$	-	-	-
MvMM [22]	0	34.14 ± 0.62	26.83 ± 0.51	36.48 ± 2.03	34.07 ± 3.47
$\mathcal{X}$-metric [13]	0	55.65 ± 3.93	66.60 ± 4.25	73.00 ± 1.07	71.55 ± 1.43
BayeSMM(w/o)	0	65.31 ± 0.87	73.98 ± 2.36	72.94 ± 2.21	74.26 ± 1.04
BayeSMM	0	66.09 ± 0.57	$\mathbf{76.39 \pm 1.35}$	$\mathbf{74.53 \pm 1.65}$	$\mathbf{76.12 \pm 1.14}$
After Affine Pre-process					
Init	0	76.08 ± 0.63	-	-	-
DGR	0	79.53 ± 0.84	-	-	-
MvMM [22]	0	75.00 ± 0.59	62.32 ± 1.18	60.83 ± 1.09	59.18 ± 2.65
$\mathcal{X}$-metric [13]	0	$\mathbf{81.19 \pm 0.95}$	85.50 ± 1.24	86.40 ± 0.98	85.38 ± 0.86
BayeSMM(w/o)	80%	76.97 ± 0.62	83.96 ± 0.87	83.51 ± 0.87	80.52 ± 0.90
BayeSMM(w/o)	50%	78.00 ± 0.90	84.66 ± 1.03	84.75 ± 0.70	82.15 ± 0.91
BayeSMM(w/o)	0	81.02 ± 0.89	85.60 ± 0.82	86.56 ± 0.62	85.78 ± 1.26
BayeSMM	80%	76.99 ± 0.63	84.44 ± 0.57	83.81 ± 0.35	81.05 ± 1.36
BayeSMM	50%	78.06 ± 0.93	84.81 ± 1.07	84.26 ± 0.55	82.81 ± 0.30
BayeSMM	0	80.93 ± 0.85	$\mathbf{85.62 \pm 0.78}$	$\mathbf{86.66 \pm 0.32}$	$\mathbf{86.27 \pm 1.00}$

of 1. The coefficient of variational loss and regularization were set as $\lambda_1 = 1$ and $\lambda_2 = 1$. Network parameters for the two tasks were optimized alternately, with each task being optimized for one iteration step at a time. Before joint training, we pretrained the model for 50 epochs with fixed outlier-detection variables ($\boldsymbol{u}$) to stabilize feature learning. All experiments used PyTorch on an NVIDIA RTX 3090 GPU. For fairness of comparison, all methods employed the same settings.

3.2 Results

Evaluation of Combined Computing. To evaluate BayeSMM's capability for joint registration and segmentation, we conducted comprehensive experiments under two clinically relevant scenarios: (1) severe misalignment without pre-processing and (2) affine-preprocessed inputs. As shown in Table 1, BayeSMM exhibits superior performance in both settings. Compared to the $\mathcal{X}$-metric, we observed improvements in segmentation performance across both scenarios, particularly in the absence of affine pre-registration, achieving a improve of 18.7%. This suggests that our two-stage iterative framework can effectively handle such

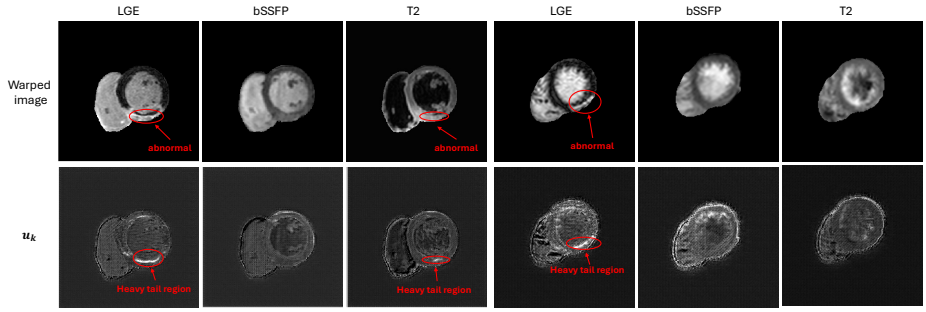


Fig. 3. Illustration of two cases. **Top row:** Warped multi-modality images. **Bottom row:** Visualization of the outlier-detection variable $\boldsymbol{u}_k$, where k corresponds to the true class label. The red arrows and circles highlight the abnormal regions (e.g., scars and edema) and their corresponding heavy-tailed regions in $\boldsymbol{u}_k$ (lighter areas with high u values). The consistency between these regions demonstrates that BayeSMM is capable of identifying abnormal areas, which contributes to its robustness.

terrible misalignments through alternating registration and segmentation refinement. On the other hand, with affine preprocessing, BayeSMM performs slightly worse than the $\mathcal{X}$-metric in registration. This may be due to the fact that multi-modality segmentations in $\mathcal{X}$-metric benefits from simultaneous processing across modalities, whereas BayeSMM provides a single segmentation map defined in the common space, which may influence the subsequent registration performance supervised by likelihood.

Ablation Study: Robustness Evaluation. To validate our outlier-detection mechanism, we conducted controlled experiments introducing synthetic abnormalities by replacing 50% and 80% of pixels with extreme intensities (below 10-th or above 90-th percentile of tissue histograms). We compared BayeSMM against its ablated variant, BayeSMM(w/o), where the outlier variable $\boldsymbol{u}$ is fixed to $\mathbf{1}$, reducing the distribution to a standard Gaussian distribution (referred to Eq. 2). As shown in Table 1, while both setting's performance degrades with increasing noise, BayeSMM demonstrates superior robustness. At 50% noise intensity, it maintains a Seg DSC of 82.81 for T2 images, compared to 82.15 for the ablated model. For 80% noise, with BayeSMM achieving lower degradation compared with BayeSMM(w/o). This results demonstrate our probabilistic outlier modeling effectively recognize outliers through robust Student's t distribution.

Visualization analysis. To better understand the impact of the outlier-detection variable $\boldsymbol{u}$, we visualize its estimated values to explore how it influences BayeSMM's handling of outliers (abnormal regions). To make it obvious, we adjusted the brightness. As illustrated in Fig. 3, the red circle highlights the abnormal regions in the image, which correspond to the heavy-tailed region in the $\boldsymbol{u}_k$ map. These regions overlap consistently, indicating that BayeSMM could automatically identify abnormal areas, such as scars or edema, and be more tolerant of

them. This feature enhances the robustness of BayeSMM, enabling it to handle outliers effectively.

4 Conclusion

In this work, we proposed a robust statistical framework for medical image registration and segmentation, termed as BayeSMM. Specifically, we incorporate the heavy-tailed nature of the Student's t distribution and develop variational inference to optimize the model. Guided by variational inference, we design a novel deep learning architecture that performs registration and segmentation jointly.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (62372115, 62201014, and 62111530195), the Shanghai Municipal Education Commission–Artificial Intelligence Initiative to Promote Research Paradigm Reform and Empower Disciplinary Advancement Plan (24KXZNA13), the Clinical Medicine Plus X–Young Scholars Project of Peking University (PKU2024LCXQ028), and the Fundamental Research Funds for the Central Universities.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ashburner, J., Friston, K.J.: Unified segmentation. NeuroImage **26**(3), 839–851 (2005)
2. Bhatia, K.K., Aljabar, P., Boardman, J.P., Srinivasan, L., Murgasova, M., Counsell, S.J., Rutherford, M.A., Hajnal, J.V., Edwards, A.D., Rueckert, D.: Groupwise combined segmentation and registration for atlas construction. In: Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI). pp. 532–540. Springer (2007)
3. Bishop, C.M., Nasrabadi, N.M.: Pattern recognition and machine learning, vol. 4. Springer (2006)
4. Blaiotta, C., Freund, P., Cardoso, M.J., Ashburner, J.: Generative diffeomorphic modelling of large mri data sets for probabilistic template construction. NeuroImage **166**, 117–134 (2018)
5. Blei, D.M., Kucukelbir, A., McAuliffe, J.D.: Variational inference: A review for statisticians. Journal of the American statistical Association **112**(518), 859–877 (2017)
6. Elmahdy, M.S., Beljaards, L., Yousefi, S., Sokooti, H., Verbeek, F., Van Der Heide, U.A., Staring, M.: Joint registration and segmentation via multi-task learning for adaptive radiotherapy of prostate cancer. IEEE Access **9**, 95551–95568 (2021)
7. Guan, H., Liu, M.: Domain adaptation for medical image analysis: a survey. IEEE Transactions on Biomedical Engineering **69**(3), 1173–1185 (2021)
8. Hachama, M., Richard, F., Desolneux, A.: A mammogram registration technique dealing with outliers. In: IEEE International Symposium on Biomedical Imaging. pp. 458–461 (2006)
9. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images (2022)

10. Havaei, M., Guizard, N., Chapados, N., Bengio, Y.: Hemis: Hetero-modal image segmentation. In: Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI). pp. 469–477. Springer (2016)
11. Joshi, A., Hong, Y.: Metaregnet: Metamorphic image registration using flow-driven residual networks. In: Computational Mathematics Modeling in Cancer Analysis. pp. 160–170 (2023)
12. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: Proceedings of the International Conference on Learning Representations (ICLR) (2015)
13. Luo, X., Zhuang, X.: $\mathcal{X}$-metric: An n-dimensional information-theoretic framework for groupwise registration and deep combined computing. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) **45**(7), 9206–9224 (2023)
14. McLachlan, G.J., Peel, D.: Finite mixture models. John Wiley & Sons (2000)
15. Mårtensson, G., Ferreira, D., Granberg, T., Cavallin, L., Oppedal, K., Padovani, A., Rektorova, I., Bonanni, L., Pardini, M., Kramberger, M.G., Taylor, J.P., Hort, J., Snædal, J., Kulisevsky, J., Blanc, F., Antonini, A., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Lovestone, S., Simmons, A., Aarsland, D., Westman, E.: The reliability of a deep learning model in clinical out-of-distribution mri data: A multicohort study. Medical Image Analysis **66**, 101714 (2020)
16. Orchard, J., Mann, R.: Registering a multisensor ensemble of images. IEEE Transactions on Image Processing **19**(5), 1236–1247 (2009)
17. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: application to breast mr images. IEEE Transactions on Medical Imaging (TMI) **18**(8), 712–721 (1999)
18. Svensén, M., Bishop, C.M.: Robust bayesian mixture modelling. Neurocomputing **64**, 235–252 (2005)
19. Wang, J., Xing, J., Druzgal, J., Wells, W.M., Zhang, M.: Metamorph: Learning metamorphic image transformation with appearance changes. In: Proceedings of the Information Processing in Medical Imaging (IPMI). pp. 576–587 (2023)
20. Wyatt, P.P., Noble, J.: Map mrf joint segmentation and registration of medical images. Medical Image Analysis **7**(4), 539–552 (2003)
21. Xiaohua, C., Brady, M., Rueckert, D.: Simultaneous segmentation and registration for medical image. In: Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI). pp. 663–670 (2004)
22. Zhuang, X.: Multivariate mixture model for myocardial segmentation combining multi-source images. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) **41**(12), 2933–2946 (2019)