Malaysian Journal of Mathematical Sciences 14(1): 77–93 (2020)

MALAYSIAN JOURNAL OF MATHEMATICAL SCIENCES

Journal homepage: http://einspem.upm.edu.my/journal

MRI-Based Brain Tumor Segmentation Using Gaussian and Hybrid Gaussian Mixture Model-Spatially Variant Finite Mixture Model with Expectation-Maximization Algorithm

Pravitasari, A. A. *[1,2], Qonita, S. F.[1], Nur Iriawan[1], Irhamah[1], Fithriasari, K.[1], Purnami, S. W.[1], and Ferriastuti, W.[3]

[1]*Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*
[2]*Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Bandung, Indonesia*
[3]*Department of Radiology, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia*

E-mail: nur_i@statistika.its.ac.id
* *Corresponding author*

Received: 10 March 2019
Accepted: 29 December 2019

ABSTRACT

A brain tumor is found in the nervous system and several studies have been conducted to assist medical personnel in dealing with the problem, one of which is through detection by using image-based medical segmentation of Magnetic Resonance Imaging (MRI). This is mostly used to separate the Region of Interest (ROI) or segments considered important in the medical point of view from others (Non-ROI), especially noise. The method commonly used is the model-based clustering with a Gaussian Mixture Model (GMM). However, this is limited by the consideration of the image between the pixels independent, thereby making the segmentation results lack noise robustness. In order to minimize this, the

Markov Random Field (MRF) model, which fully considers the spatial dependencies between pixels, was used. The combination was named Spatially Variant Finite Mixture Model (SVFMM), with the initial parameter generated from the GMM, making the proposed model a hybrid GMM-SVFMM. In the inference process, the maximum likelihood estimation method was employed to estimate the proposed model parameters using the Expectation-Maximization (EM) algorithm. The results from the correct classification ratio (*CCR*) showed that MRI-based brain image segmentation coupled with hybrid GMM-SVFMM was able to provide more accurate results in separating ROI from noise compared to GMM, with *CCR* of 0.9876 for hybrid GMM-SVFMM and 0.8735 for GMM.

Keywords: Expectation-Maximization, GMM, Hybrid GMM-SVFMM, Markov Random Fields, Image Segmentation.

1. Introduction

A brain tumor, besides the spinal and periphery, is another example of a tumor in the nervous system. It can be either primary or as a metastatic of other organs (Wahjoepramono, 2006). Several works of research have been conducted to assist medical personnel deal with the disease, one of which is detection through segmentation of medical image based on MRI. This is, however, an advanced medical imaging technique that provides information about the anatomy of the soft tissue in the human brain. MRI depiction techniques are relatively complex because the resulting image depends on the number of parameters involved. Such that the right selection provides quality and detailed picture of the human body with contrast differences for an accurate evaluation of anatomy and pathology of the tissues. Furthermore, the amount of data available on the MRI is too big to be analyzed or interpreted manually and this forms one of the biggest obstacles to its efficiency. Therefore, there is a need for automatic or semi-automatic computational analysis techniques for the data.

One of the important steps in image analysis is the medical image segmentation. This is usually conducted in an MRI to separate ROI or segments considered medically important from others or the Non-ROI. Moreover, Finite Mixture Model (FMM) is one of the Model-Based Clustering techniques widely applied in image segmentation with each mixture component being a probabilistic distribution (Pravitasari et al., 2019). Furthermore, the FMM is called a GMM when the Gaussian distribution is selected as its component function. This has the possibility to obtain a good segmentation result for an image without noise, but usually, it gives an unsatisfactory result when there is noise. The main reason is that it assumes the relationship between pixels in an image is statistically independent. However, in a previous study conducted by Sanjay and Hebert (1988), MRF, due to its pixel spatial dependencies, was used to improve the robustness of GMM against noise. The model is relatively simple but effective to cover prior knowledge in the segmentation process. By modeling the image as MRF, each pixel is assumed to be statistically dependent on others and can be easily expressed through the definition of potential functions corresponding to the Gibbs distribution.

Another research conducted by Iriawan et al. (2018) implemented the EM-GMM method and showed that it is more robust against Salt and Pepper Noise, while Fuzzy Clustering Method (FCM) is more robust against Gaussian Noise. Therefore, in order to overcome the robustness of noise, this study applied MRF in accordance with Sanjay and Hebert (1988) and named Spatially Variant Finite Mixture Model (SVFMM). Based on the explanation above, this

study conducted segmentation by using GMM and the SVFMM, known as hybrid GMM-SVFMM with initialization parameters obtained from the GMM as one way to speed up convergence. This was necessary due to the consideration of the results to be capable of approaching the true solution in the algorithm (Chapra and Canale, 2015). Furthermore, the EM algorithm was employed to optimize these two methods and the results were compared using the Correct Classification Ratio (CCR). The expected benefit of this research was to maximize quality and reduce noise in image segmentation, especially in medical images or MRI.

2. Methodology

2.1 Gaussian Mixture Model

Model-Based Clustering is a method founded on the probability model of data. It is assumed that if there is only a limited (finite) number of clusters in the model, then it could be assumed fixed. Each cluster would represent a group with certain probability distributions as the best fitting to their characteristics of data. As stated previously, the FMM is considered GMM when the Normal or Gaussian distribution is selected as a component and structured as a weighted linear combination (McLachlan and Peel, 2000). It can be approximated by the formula stated below (Astuti et al. (2017)):

$$f\left(y|\boldsymbol{\pi},\boldsymbol{\mu},\boldsymbol{\sigma}\right)=\sum_{j=1}^{K}\pi_j f\left(y|\mu_j,\sigma_j^2\right), \tag{1}$$

where $f\left(y|\boldsymbol{\pi},\boldsymbol{\mu},\boldsymbol{\sigma}\right)$ is the density distribution function of the mixture model and $f\left(y|\mu_j,\sigma_j^2\right)$ is the Normal density function of the j^{th} component, defined as:

$$f\left(y|\mu_j,\sigma_j^2\right)=\frac{1}{\sqrt{2\pi}\sigma_j}\exp\left(-\frac{1}{2}\left(\frac{y-\mu_j}{\sigma_j}\right)^2\right). \tag{2}$$

The vector of mixture proportions is denoted by $\{\pi_1,\pi_2,...,\pi_j,...,\pi_K\}$, that satisfies $0\leq\pi_j\leq 1$ and $\sum_{j=1}^{K}\pi_j=1$, K is the number of $Normal$ distribution components in the mixture model.

In this research, the greyscale intensity of the MRI image was defined as the response, y variable, and $y_1,y_2,...,y_N$ was assumed to follow the mixture Gaussian distribution. The data y could be in the domain of $0\leq y\leq 255$, as the greyscale intensity range from black to white, with mean and variance $0<\mu_j<255$, and $\sigma_j^2>0$ respectively.

Thus, to facilitate this research, the benefits of the GMM function were used. The Normal density function has a simple form with only two parameters. The log-likelihood of the GMM is as follows:

$$L(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\pi}|\mathbf{y}) = \sum_{i=1}^{N} \log \left(\sum_{j=1}^{K} \pi_j \frac{1}{\sqrt{2\pi}\sigma_j} \exp \left(-\frac{1}{2} \left(\frac{y_i - \mu_j}{\sigma_j} \right)^2 \right) \right) \tag{3}$$

This, in general, has no direct closed-form solution to be maximized. Therefore, the latent analysis technique that could be used is a relocation algorithm such as Expectation-Maximization (EM).

2.2 Hybrid GMM-SVFMM

In order to improve the robustness to noise, a mixture model with MRF has been widely used to model the joint distribution for pixel labeling (Nguyen and Wu (2013), Ji et al. (2016)). Furthermore, the Hammersley-Clifford theorem states that the random variable is the MRF with respect to the neighborhood system if and only if it is corresponding to a Gibbs distribution with the nearest neighbor Gibbs potential $V_{C_i}(\mathbf{w})$, such as:

$$p(\mathbf{w}) = \frac{1}{Z} \exp \left(-\beta \sum_{i=1}^{N} V_{C_i}(\mathbf{w}) \right), \tag{4}$$

where $\mathbf{w}$ is a set of the probability vector, Z is normalizing constant, $V_{C_i}(\mathbf{w})$ is the clique potential function of pixel label vector $\{\pi_i\}$ to its neighborhood pixel, C is set of cliques (which can be written as $C = C_1 \cup C_2 \cup ... \cup C_N$), and β is the prior parameter that controls the degree to which smoothness is imposed.

The SVFMM modified the classical mixture model for pixel labeling by assuming K components, each of which has a parameter density vector $\boldsymbol{\theta}_j = \left(\mu_j, \sigma_j^2\right)$, follows a Gaussian distribution (Sanjay and Hebert, 1988). This approach classified the probability of the i^{th} pixel into the j^{th} cluster (class label) and defined it by $p(y_i = j) = \pi_{ij}$. Therefore, the density function in equation (1) was changed to:

$$f(y_i|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\sigma}) = \sum_{j=1}^{K} \pi_{ij} f\left(y_i|\mu_j, \sigma_j^2\right) \tag{5}$$

with log-likelihood function as follows:

$$L(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\pi}|\mathbf{y}) = \sum_{i=1}^{N} \log \left(\sum_{j=1}^{K} \pi_{ij} f\left(y_i|\mu_j, \sigma_j^2\right) \right) \tag{6}$$

If all constants are ignored and the MRF distribution is applied to adapt the spatial correlation between label values, the log-likelihood function for SVFMM is given by:

$$\begin{aligned} L(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\pi}|\mathbf{y}) &= \sum_{i=1}^{N} \log \left\{ \sum_{j=1}^{K} \pi_{ij} f\left(y_i|\mu_j, \sigma_j^2\right) \right\} + \log p\left(\boldsymbol{\pi}\right) \\ &= \sum_{i=1}^{N} \log \left\{ \sum_{j=1}^{K} \pi_{ij} f\left(y_i|\mu_j, \sigma_j^2\right) \right\} - \beta \sum_{i=1}^{N} \sum_{j=1}^{K} \sum_{m \in C} \left(\pi_{ij} - \pi_{mj}\right)^2. \end{aligned} \tag{7}$$

Based on previous research, the model parameters can be estimated through maximum likelihood (ML) coupled with the EM algorithm (Nguyen, 2011). Therefore, the initialization parameters used for SVFMM were obtained from the results of the parameter estimation with GMM. This application of the EM algorithm on the SVFMM method coupled with the initialization result from GMM is a new innovative approach and was called a Hybrid GMM-SVFMM method.

2.3 Expectation-Maximization Algorithm

This algorithm is a well-known iterative numerical approach for estimating a model which has non-close-form of MLE. It has a simple and tractable step of iterations and a convergent achievement in the stable iterations (Nguyen, 2011). It was, therefore, used in this paper to estimate GMM parameters implemented in **Algorithm 1** and Hybrid GMM-SVFMM parameters implemented in **Algorithm 2**.

Algorithm 1. The EM Algorithm for GMM

1. Initialization parameters$\left\{\mu_j^{(0)}, \sigma_j^{2(0)}, \pi_j^{(0)}\right\}$, mean $\mu_j^{(0)}$; variance $\sigma_j^{2(0)}$; proportion $\pi_j^{(0)}$, and set $t = 0$.
2. E-Step: Calculate z_{ij}

$$z_{ij}^{(t+1)} = \frac{\pi_j^{(t)} f\left(y_i|\mu_j^{(t)}, \sigma_j^{(t)}\right)}{\sum_{l=1}^{K} \pi_l^{(t)} f\left(y_i|\mu_l^{(t)}, \sigma_l^{(t)}\right)}. \tag{8}$$

3. M-Step: Re-estimate parameters$\{\mu_j, \sigma_j, \pi_j\}$

a. Update mean μ_j

$$\mu_j^{(t+1)} = \frac{\sum_{i=1}^{N} z_{ij}^{(t+1)} y_i}{\sum_{i=1}^{N} z_{ij}^{(t+1)}}. \tag{9}$$

b. Update variance σ_j^2

$$\left(\sigma_j^2\right)^{(t+1)} = \frac{\sum_{i=1}^{N} z_{ij}^{(t+1)} \left(y_i - \mu_j^{(t+1)}\right)^2}{\sum_{i=1}^{N} z_{ij}^{(t+1)}}. \tag{10}$$

c. Update proportion π_j

$$\pi_j^{(t+1)} = \frac{1}{N} \sum_{i=1}^{N} z_{ij}^{(t+1)}. \tag{11}$$

4. Evaluate log-likelihood function in equation (3) and check for convergence, both from the log-likelihood function and the parameter values. If the convergent criterion error value <0.0001 has not been fulfilled, set $t = t + 1$ and return to step 2.

Algorithm 2. The EM Algorithm Hybrid GMM-SVFMM

1. Initialization parameters $\left\{\mu_j^{(0)}, \Sigma_j^{(0)}, \pi_{ij}^{(0)}\right\}$, mean $\mu_j^{(0)}$; covariant $\Sigma_j^{(0)}$; and proportion $\pi_{ij}^{(0)}$, given by the results of the estimated parameter of GMM using **Algorithm 1**, and set $t = 0$.
2. E-Step: Calculate z_{ij}

$$z_{ij}^{(t+1)} = \frac{\pi_{ij}^{(t)} f\left(y_i | \mu_j^{(t)}, \Sigma_j^{(t)}\right)}{\sum_{l=1}^{K} \pi_{il}^{(t)} f\left(y_i | \mu_l^{(t)}, \Sigma_l^{(t)}\right)}. \tag{12}$$

3. M-Step: Re-estimate parameters $\left\{\mu_j, \sigma_j^2, \pi_j\right\}$

   Update mean (μ_j) and varians (σ_j^2) using equation (9) and (10) along with proportion (π_{ij}) using equation (13).

$$\pi_{ij}^{(t+1)} = \frac{\sum_{m \in C} \pi_{mj}^{(t)} + \left| \left(\left(\sum_{m \in C} \pi_{mj}^{(t)} \right)^2 + \frac{\mathcal{N}_i}{\beta} z_{ij}^{(t+1)} \right)^{\frac{1}{2}} \right|}{2\mathcal{N}_i}, \tag{13}$$

where $\mathcal{N}_i$ is the neighborhood for the i^{th} pixel.

4. Evaluate log-likelihood function in equation (7) and check for convergence, both from the log-likelihood function and the parameter values. If the convergent criterion error value <0.0001 has not been fulfilled, set $t = t + 1$ and return to step 2.

2.4 Measure of Evaluation

The CCR is an evaluation criterion used to test whether the ROI obtained is in accordance with its ground truth. The MRI data used as an input in this study have been taken in medical action. Therefore, the brain tumor target was already physically known as the ground truth. The CCR value can be defined as in equation (14).

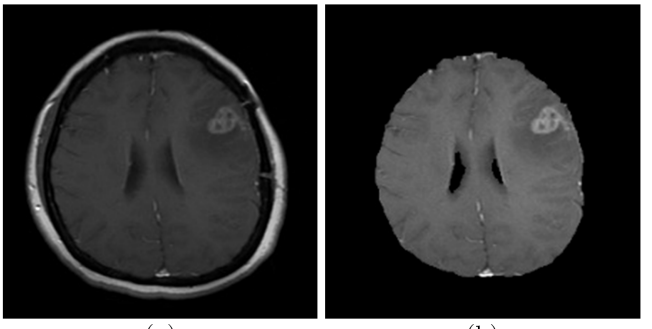
$$CCR = \sum_{j=1}^{2} \frac{|GT_j \cap Seg_j|}{|GT_j|}, \tag{14}$$

where GT_j is the ground truth for non-ROI area ($j = 1$) and ROI area ($j = 2$), meanwhile Seg_j describes a pixel clustered to the non-ROI area ($j = 1$) or ROI area ($j = 2$), and $GT = \sum_{j=1}^{2} GT_j$. Evaluation of the image segmentation can be seen from the greater the CCR value produced, the better the segmentation model (Nikou et al., 2007).

3. Results and Discussions

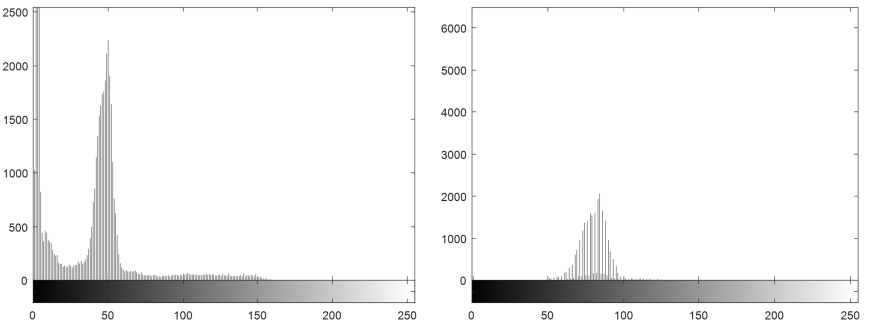
3.1 Image Selection and Preprocessing Data

The first step taken in MRI-based image segmentation was the selection of MRI image data provided by the Regional Public Hospital Dr. Soetomo from Surabaya, Indonesia. This was followed by preprocessing to improve image quality. Furthermore, MRI images are usually selected based on certain sequences and slices recommended from a medical point of view and for the purpose of this study, sequence ax T1 memp+C was used. This is an MRI image produced by axial pieces of the patient's brain with contrast media. However, the image of the comparison between the initial/input and the preprocessing result is as shown in Figure 1.



(a) (b)

Figure 1: Preprocessing results (a) Original image (b) Image after preprocessing



(a) (b)

Figure 2: Image histogram (a) Before preprocessing (b) After preprocessing

Figure 1 (b) shows the preprocessed image to include the composition of greyscale intensity representing the brain area only with the skull image seen in Figure 1(a) was removed. Furthermore, data exploration was conducted by constructing greyscale histograms based on the image matrix to estimate the density. As shown in Figure 2, the input image histogram before (Figure 2 (a)) and after (Figure 2 (b)) preprocessing, demonstrate the multimodal pattern as a mixture distribution.

Figure 2(a) shows the initial histogram has 3 peaks and from the point of the highest, ROI could be assumed absent. Therefore, it was removed from the observation through preprocessing as shown in Figure 2(b). The figure shows a multimodal pattern which can be represented as a mixture distribution. With

respect to this information, the MRI image segmentation was conducted based on model-based clustering through the use of GMM method and the proposed Hybrid GMM-SVFMM coupled with the EM algorithm.

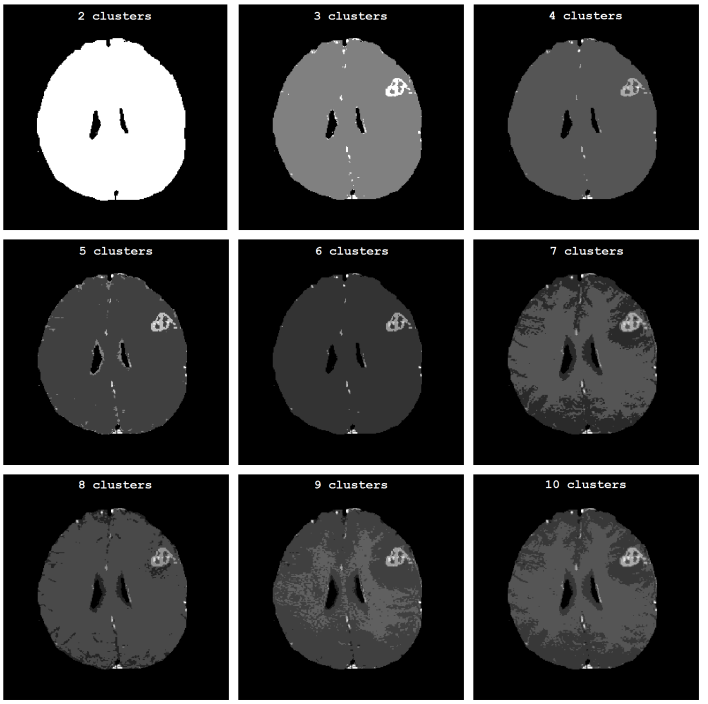


Figure 3: MRI-based Segmentation Results using GMM for 2 to 10 Number of Clusters (from top left to bottom right)

3.2 Trial, Selection, and Cluster Validation

There is no standard guide to determining the optimal number of clusters in the MRI-based image segmentation using GMM and Hybrid GMM-SVFMM. Therefore, several numbers of clusters were tested to get the best through the use of CCR coupled with a subjective assessment from medical experts. The results of GMM for 2 to 10 number of clusters are as shown in Figure 3, while the ones conducted with the Hybrid GMM-SVFMM method is shown in Figure 4.

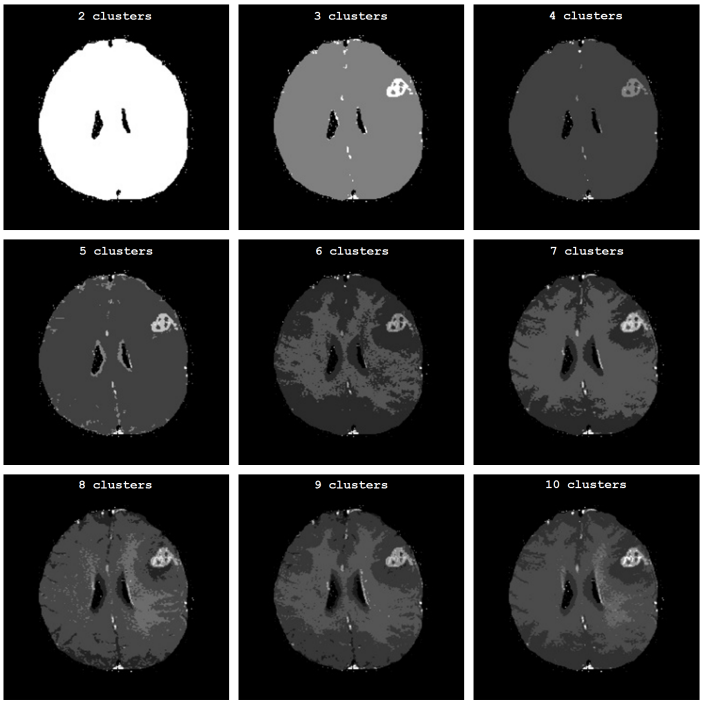


Figure 4: MRI-based Segmentation Results using Hybrid GMM-SVFMM for 2 to 10 Number of Clusters (from top left to bottom right)

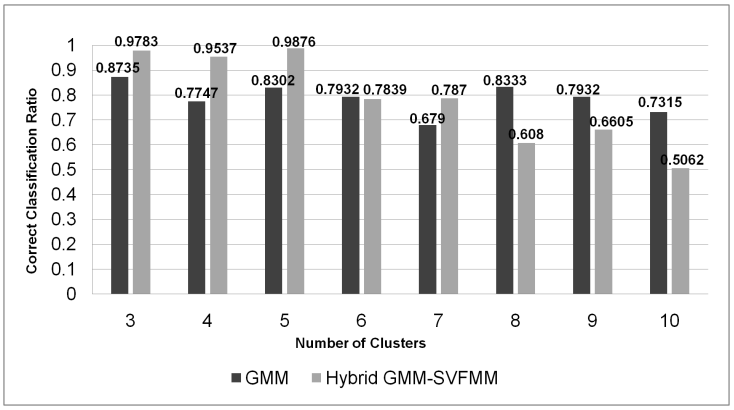
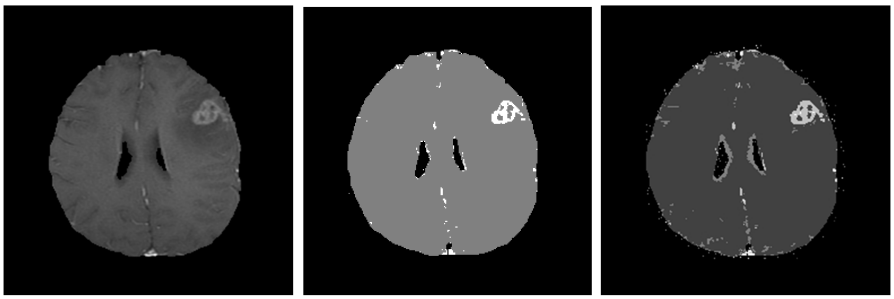


Figure 5: The CCR for GMM and Hybrid GMM-SVFMM

Figure 3 and Figure 4 reveal brain tumor ROI can be caught when segmentation is conducted for 3 to 10 number of clusters. The comparison of the CCR value for GMM and Hybrid GMM-SVFMM is as shown in Figure 5. It was found that the greatest value was 0.8735 for 3 clusters with GMM and 0.9876 for 5 clusters with the Hybrid GMM-SVFMM. These have been validated by medical experts.

3.3 Model Comparison

Image segmentation with both GMM and hybrid GMM-SVFMM were conducted by assuming MRI image data consists of Normal distribution components. However, the difference is that Hybrid GMM-SVFMM considered the spatial correlation of pixels. The results of the image segmentation using both methods with an optimum number of clusters are shown in Figure 6.



(a) (b) (c)

Figure 6: Image Segmentation (a) Preprocessing Image (b) GMM with 3 Clusters (c) Hybrid GMM-SVFMM with 5 Clusters

The Figure 6(a) demonstrates the input image for segmentation of MRI of brain tumors after preprocessing while Figure 6(b) shows the 3 clusters as the optimal GMM segmentation with the indication of the brain tumor area with a light gray color. Whereas, Figure 6 (c) presents the 5 clusters as the optimal segmentation using Hybrid GMM-SVFMM with the area indicated by a lighter gray color. Furthermore, the estimated parameters of both approaches were shown in Table 1.

Table 1: Parameters Estimation of GMM and Hybrid GMM-SVFMM

Method	**Cluster:**	$\boldsymbol{\pi}$	$\boldsymbol{\mu}$	$\boldsymbol{\sigma}$
GMM	1	0.5325	0.6684	0.5591
	2	0.4487	79.9858	68.2674
	3	0.0188	94.4135	769.0575
Hybrid GMM-SVFMM	1	$\pi_{11}, \pi_{21}, \ldots, \pi_{N1}$	0	0
	2	$\pi_{12}, \pi_{22}, \ldots, \pi_{N2}$	80.7825	54.4523
	3	$\pi_{13}, \pi_{23}, \ldots, \pi_{N3}$	62.523	231.603
	4	$\pi_{14}, \pi_{24}, \ldots, \pi_{N4}$	111.3844	66.0721
	5	$\pi_{15}, \pi_{25}, \ldots, \pi_{N5}$	111.321	3976.1

The ROI area of the brain tumor was defined by the third cluster either through the use of GMM or hybrid GMM-SVFMM. MRI-based Image segmentation using GMM produced an ROI area with the proportion of 0.0188, intensity mean of 94.4135 and variance of 769.0575, while the Hybrid GMM-SVFMM produced an area with intensity mean and variance of 62.5230 and 231.6030 respectively. The segmentation density function of both GMM and the Hybrid GMM-SVFMM can be composed as shown in equation (17) and (18). The class label parameter π_{ij} for the Hybrid GMM-SVFMM was not written one by one due to the big number of estimated parameters i.e.$5 \times 65,536 = 327,680$ parameters. The $\mathbf{y}_{i\in j}$ denotes the i^{th} data that belongs to the j^{th} cluster.

$$
\begin{aligned}
f\left(\mathbf{y}|\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\pi}\right) &= \sum_{j=1}^{3} \pi_j f\left(\mathbf{y}_{i\in j}|\mu_j,\sigma_j\right) \\
&= \pi_1 N\left(\mathbf{y}_{i\in j=1}|\mu_1,\sigma_1\right) + \pi_2 N\left(\mathbf{y}_{i\in j=2}|\mu_2,\sigma_2\right) + \\
&\quad \pi_3 N\left(\mathbf{y}_{i\in j=3}|\mu_3,\sigma_3\right) \\
&= 0.5325 \tfrac{1}{\sqrt{2\pi 0.5591}} \exp\left(-\tfrac{1}{2}\tfrac{\left(\mathbf{y}_{i\in j=1}-0.6684\right)^2}{0.5591}\right) + \\
&\quad 0.4487 \tfrac{1}{\sqrt{2\pi 68.2674}} \exp\left(-\tfrac{1}{2}\tfrac{\left(\mathbf{y}_{i\in j=2}-79.9858\right)^2}{68.2674}\right) + \\
&\quad 0.0188 \tfrac{1}{\sqrt{2\pi 769.0575}} \exp\left(-\tfrac{1}{2}\tfrac{\left(\mathbf{y}_{i\in j=3}-94.4135\right)^2}{769.0575}\right).
\end{aligned} \tag{15}
$$

$$
\begin{aligned}
f\left(\mathbf{y}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\pi}\right) &= \sum_{j=1}^{5}\boldsymbol{\pi}_{ij}f\left(\mathbf{y}_{i\in j}|\mu_j,\Sigma_j\right)\\
&= \boldsymbol{\pi}_{i1}N\left(\mathbf{y}_{i\in j=1}|\mu_1,\sigma_1\right)+\boldsymbol{\pi}_{i2}N\left(\mathbf{y}_{i\in j=2}|\mu_2,\sigma_2\right)+\\
&\quad \boldsymbol{\pi}_{i3}N\left(\mathbf{y}_{i\in j=3}|\mu_3,\sigma_3\right)+\boldsymbol{\pi}_{i4}N\left(\mathbf{y}_{i\in j=4}|\mu_4,\sigma_4\right)+\\
&\quad \boldsymbol{\pi}_{i5}N\left(\mathbf{y}_{i\in j=5}|\mu_5,\sigma_5\right)\\
&= \boldsymbol{\pi}_{i1}\frac{1}{\sqrt{2\pi 0}}\exp\left(-\frac{1}{2}\frac{\left(\mathbf{y}_{i\in j=1}-0\right)^2}{0}\right)+\\
&\quad \boldsymbol{\pi}_{i2}\frac{1}{\sqrt{2\pi 54.4523}}\exp\left(-\frac{1}{2}\frac{\left(\mathbf{y}_{i\in j=2}-80.7825\right)^2}{54.4523}\right)+\\
&\quad \boldsymbol{\pi}_{i3}\frac{1}{\sqrt{2\pi 231.6030}}\exp\left(-\frac{1}{2}\frac{\left(\mathbf{y}_{i\in j=3}-62.523\right)^2}{231.6030}\right)+\\
&\quad \boldsymbol{\pi}_{i4}\frac{1}{\sqrt{2\pi 66.0721}}\exp\left(-\frac{1}{2}\frac{\left(\mathbf{y}_{i\in j=4}-111.3844\right)^2}{66.0721}\right)+\\
&\quad \boldsymbol{\pi}_{i5}\frac{1}{\sqrt{2\pi 3976.1}}\exp\left(-\frac{1}{2}\frac{\left(\mathbf{y}_{i\in j=5}-111.3210\right)^2}{3976.1}\right).
\end{aligned}
\tag{16}
$$

The clear differences between the ROI and Non-ROI areas in the 3 clusters of GMM and 5 clusters of Hybrid GMM-SVFMM are presented in Figure 7. Figure 7(a) shows the area suspected to be the ROI of a brain tumor while Figure 7(b) shows the Non-ROI. The results of GMM segmentation indicate the method is not robust to noise as shown by the presence of several small dots far from the area of the original tumor. Meanwhile, the Hybrid GMM-SVFMM segmentation results show the suspected ROI in Figure 7(c) and the area not affected or Non-ROI in Figure 7(d). These show this method is better compared to the results of GMM segmentation. The hybrid method has successfully demonstrated its expertise in estimating ROI by showing its findings to be more solid and free from noise.

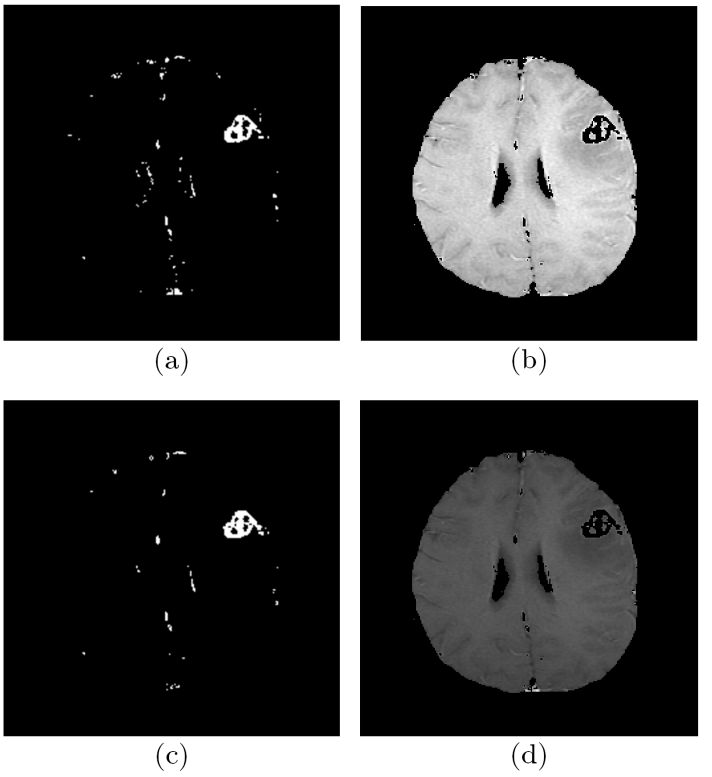


Figure 7: Segmentation Results; (a) GMM with 3 Clusters ROI Area, (b) GMM with 3 clusters Non-ROI Area, (c) Hybrid GMM-SVFMM with 5 Clusters ROI Area, and (d) Hybrid GMM-SVFMM with 5 Clusters Non-ROI Area

The results of both methods show that there are differences in finding the ROI area of brain tumors. However, their evaluation and characteristics using the CCR are as shown in Table 2. From the information provided in Table 2, it can be concluded that by using the same hardware, Intel (R) Core (TM) i5-4200U CPU, 1.60GHz CPU with 6.00GB RAM and using EM algorithm, the Hybrid GMM-SVFMM method has the ability to provide more accurate segmentation of brain tumor ROI area compared to the GMM method. This is indicated by its higher CCR value of 0.9876 which is bigger than 0.8735 for GMM. Nevertheless, it is quite complex in terms of computing because it required a longer running time of 45.399 seconds compared to 8.333 seconds for the other method. In addition, its high CCR has to be paid with a complex model possessing a greater number of parameters of 327,680 compared to only 9 in the other method.

Table 2: The Comparison of GMM and the Hybrid GMM-SVFMM with EM Algorithm

Method	Number of Cluster	Number of Parameters	Consuming Computer Time Run (second)	CCR
GMM	3	9	8.334	0.8735
Hybrid GMM-SVFMM	5	327,680	45.399	0.9876

4. Conclusions and Future Research

Based on the implementation, trials, results, and discussion conducted, some conclusions were obtained from this study. It was observed that after preprocessing, the data of MRI sequence ax T1 memp+C made the quality of the MRI image better. Furthermore, validation of the best number of clusters based on CCR and subjective assessment of medical experts found 3 clusters for the GMM and 5 clusters for the hybrid GMM-SVFMM method. Evaluation of the segmentation results through the CCR value also showed that the Hybrid GMM-SVFMM is better than the GMM. This was indicated by the CCR value of 0.9876 for the Hybrid GMM-SVFMM and a smaller 0.8735 for the GMM.

For future research, other Integrated Development Environment (IDE) implementations should be employed for faster and efficient computing time. Furthermore, other distribution should be employed for the greyscale that could represent the original image pattern in order to provide a better segmentation.

Acknowledgment

The Authors are grateful to the Directorate for Research and Community Service (DRPM), Ministry of Research, Technology and Higher Education Indonesia for supporting this research under PSN research grant no. 965/PKS/ITS/2018.

References

Astuti, A. B., Iriawan, N., Irhamah, Kuswanto, H., and Sasiarini, L. (2017). Bayesian Mixture Modeling for Blood Sugar Levels of Diabetes Mellitus Pa-

tients (Case Study in RSUD Saiful Anwar Malang Indonesia). *Journal of Physics: Conference Series*, 893:012036.

Chapra, S. C. and Canale, R. P. (2015). *Numerical Methods for Engineers*. McGraw-Hill Higher Education, Boston, 7th edition.

Iriawan, N., Pravitasari, A. A., Fithriasari, K., Irhamah, Purnami, S. W., and Ferriastuti, W. (2018). Comparative Study of Brain Tumor Segmentation using Different Segmentation Techniques in Handling Noise. In *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM)*, pages 289–293.

Ji, Z., Huang, Y., Sun, Q., and Guo, C. (2016). A Spatially Constrained Generative Asymmetric Gaussian Mixture Model for Image Segmentation. *Journal of Visual Communication and Image Representation*, **40**(B):600–626.

McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Model, Wiley Series in Probability and Statistics*. John Wiley and Sons Inc, New York.

Nguyen, T. (2011). Gaussian Mixture Model Based Spatial Information Concept for Image Segmentation. *Electronic Theses and Dissertation 438, https://scholar.uwindsor.ca/etd/438*.

Nguyen, T. M. and Wu, Q. M. J. (2013). Fast and Robust Spatially Constrained Gaussian Mixture Model for Image Segmentation. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(4):621–635.

Nikou, C., Galatsanos, N. P., and Likas, A. C. (2007). A Class-Adaptive Spatially Variant Mixture Model for Image Segmentation. *IEEE Transactions on Image Processing*, 16(4):1121–1130.

Pravitasari, A. A., Iriawan, N., Safa, M., Fithriasari, K., Irhamah, Purnami, S. W., and Ferriastuti, W. (2019). MRI-based Brain Tumor Segmentation using Modified Stable Student's t from Burr Mixture Model with Bayesian Approach. *Malaysian Journal of Mathematical Science*, 13(3):297–310.

Sanjay, G. S. and Hebert, T. J. (1988). Bayesian Pixel Classification using Spatially Variant Finite Mixtures and The Generalized EM Algorithm. *IEEE Trans. Image Process*, **7**(7):1014–1028.

Wahjoepramono (2006). *Tumor Otak*. Subur Jaringan Cetak Terpadu, Jakarta.