Comparative Analysis of Probabilistic Methods for Brain Tumor Segmentation Images

1 Abstract

Accurate brain tumor segmentation in MRI images is essential for diagnosis and treatment planning. This study compares three probabilistic methods: Gaussian Mixture Models (GMM), Hidden Markov Models (HMM), and a hybrid Hidden Monte Carlo (HMC) with the Chan-Vese model. GMM achieved a Dice Similarity coefficient (DSC) of 0.7550, effectively clustering pixel intensities but lacking spatial structure modeling. HMM demonstrated improved spatial modeling with adaptive state creation, achieving a DSC of 0.6389, but faced challenges as the number of clusters rapidly increased when processing complex images. Integrating Markov Random Fields with variational energy minimization, the hybrid model approaches a DSC above 0.8 and offers robust segmentation and better handling of intensity inhomogeneities but comes with a higher computational cost. These results provide a comparative analysis, highlighting how different probabilistic models impact the accuracy and robustness of brain tumor image segmentation.

2 Introduction

Image segmentation is critical in medical imaging, particularly for brain tumor detection and treatment planning, as it provides essential details about tumor size, shape, and location. However, the complexity of brain images and tumor variability make robust segmentation methods crucial.

Our original project aimed to evaluate GMM and bidirectional Hidden Markov Models (bi-HMM) for brain tumor segmentation. While bi-HMM theoretically captures horizontal and vertical pixel dependencies, its implementation faced challenges, including computational inefficiency, instability, and scalability issues with high-resolution images. These limitations led us to pivot to more practical alternatives: Hidden Markov Models (HMM) and Hidden Markov Chain (HMC) models. Both approaches balance computational efficiency and segmentation accuracy, leveraging probabilistic modeling principles to address the challenges posed by complex medical images.

In addition to these methods, we will also include GMM as a baseline model. These three methods— GMM, HMM, and HMC-based model will be evaluated for their ability to segment brain tumors in real-world medical imaging data. The dataset, sourced from Hugging Face's Brain Tumor Object Detection repository, provides annotated brain images with tumor regions, which serves as a robust testbed for assessing segmentation performance.

The primary goal of this research is to identify the most effective method for brain tumor segmentation by comparing the accuracy of GMM, HMM, and the HMC-based model. Segmentation performance will be measured using widely accepted metrics such as the Dice Similarity Coefficient (DSC). Through this comprehensive evaluation, we aim to provide insights into the strengths and limitations of each method, contributing to the development of improved segmentation techniques for medical imaging.

3 Previous Work

Brain tumor segmentation from MRI images has evolved from traditional methods like thresholding (Gordillo et al., 2013) and clustering (e.g., k-means, Fuzzy-Possibilistic Clustering (Bal et al., 2018)) to machine learning and deep learning techniques. While early approaches were efficient,

they needed help with intensity inhomogeneity and the complex morphology of brain tumors, limiting their robustness in medical imaging tasks (Wadhwa et al., 2019). Deep learning methods, including Fully Convolutional Neural Networks (FCNN) (Shelhamer et al., 2017) and 3D Convolutional Neural Networks (3D-CNN) (Kamnitsas et al., 2017), have achieved state-of-the-art results by learning hierarchical features for precise tumor boundary delineation.

While deep learning models demonstrate high accuracy, probabilistic approaches like GMM, HMM, and the HMC-base hybrid model remain attractive alternatives. These methods do not rely on labeled data, making them particularly useful in medical scenarios with limited annotations. Moreover, their lower computational demands make them practical for deployment in resource-constrained settings. To summarise, this study compares the performance of GMM, HMM, and the HMC-based hybrid model for brain tumor segmentation to bridge the gap between traditional and modern methods, offering adaptable solutions to medical image analysis challenges.

4 Methods

4.1 GMM

In this project, we first employed the Expectation-Maximization (EM) algorithm as a baseline to estimate the parameters of a Gaussian Mixture Model (GMM) for clustering pixel intensities in an image of the dataset into distinct classes. The GMM assumes that the pixel intensities are drawn from a mixture of Gaussian distributions, each characterized by a set of parameters: weight w_k, mean μ_k, and variance σ_k^2, where k denotes the class index. The algorithm iteratively updates these parameters to maximize the likelihood of the observed data under the model.

To determine the optimal number of classes K, we first evaluated the Akaike Information Criterion (AIC) and Minimum Description Length (MDL) followed by Liang et al. (1992). These criteria balance model fit and complexity, with AIC defined as $\text{AIC} = -2\ln(L) + 2p$, and MDL as $\text{MDL} = -\ln(L) + p\ln(N)$, where L is the likelihood, p the number of parameters, and N the number of data points. However, subsequent experiments revealed that both methods favored overly complex models needed to be more suitable for our image data.

To address this issue, we used the Silhouette Score, which evaluates clustering quality based on how similar points in the same cluster are and how distinct clusters are. For a data point i, the Silhouette Score $s(i)$ is defined as: $s(i) = \frac{b(i)-a(i)}{\max(a(i),b(i))}$, where $a(i)$ is the average distance between i and other points in the same cluster, and $b(i)$ is the minimum average distance between i and points in other clusters. By maximizing the Silhouette Score across different values of K, we identified the best clustering structure for the image.

Using the most optimal K from the above three methods, the EM algorithm effectively segmented the image. The EM algorithm operates in the Expectation step (E-step) and the Maximization step (M-step). In the E-step, we calculate the responsibility z_{kij}, which represents the probability of a given data point belonging to a specific class. The responsibility is computed using the Gaussian probability density function, normalized across all classes for each data point. Mathematically, this step involves:

$$z_{kij} = \frac{w_k \cdot p(Y_{ij} \mid \mu_k, \sigma_k^2)}{\sum_{l=1}^{K} w_l \cdot p(Y_{ij} \mid \mu_l, \sigma_l^2)},$$

where $p(Y_{ij} \mid \mu_k, \sigma_k^2)$ is the Gaussian density function evaluated at Y_{ij}.

In the M-step, the parameters w_k, μ_k, and σ_k^2 are updated based on the responsibilities computed in the E-step. The weight w_k is updated as the proportion of the total responsibility for class k to

the total number of data points. The mean μ_k is updated as a weighted average of the data, and the variance σ_k^2 is updated based on the weighted squared deviations of the data from the mean:

$$w_k = \frac{\sum_{i=1}^{N} z_{kij}}{N}, \quad \mu_k = \frac{\sum_{i=1}^{N} z_{kij} \cdot Y_{ij}}{\sum_{i=1}^{N} z_{kij}}, \quad \sigma_k^2 = \frac{\sum_{i=1}^{N} z_{kij} \cdot (Y_{ij} - \mu_k)^2}{\sum_{i=1}^{N} z_{kij}}.$$

The pseudocode outlines our EM algorithm for GMM (See **Algorithm 1**). It alternates between the E-step, where responsibilities (γ_{ik}) are calculated based on the current parameters, and the M-step, where weights, means, and variances are updated to maximize the likelihood of the data. The process repeats until the log-likelihood converges or a maximum number of iterations is reached, assigning each data point to the class with the highest responsibility.

Algorithm 1 EM Algorithm for GMM Parameter Estimation

Require: Data $\{x_i\}_{i=1}^{N}$, components K, max iterations max_iter, tolerance tol
Ensure: Parameters $\{w_k, \mu_k, \sigma_k^2\}$, labels $\{\text{label}_i\}$
Initialize w_k, μ_k, σ_k^2
for $n = 1$ to max_iter **do**
    E-Step: $\gamma_{ik} \leftarrow \frac{w_k \cdot p_k(x_i|\mu_k,\sigma_k^2)}{\sum_j w_j \cdot p_j(x_i|\mu_j,\sigma_j^2)}$
    M-Step: $w_k \leftarrow \frac{\sum_i \gamma_{ik}}{N}$, $\mu_k \leftarrow \frac{\sum_i \gamma_{ik} x_i}{\sum_i \gamma_{ik}}$, $\sigma_k^2 \leftarrow \frac{\sum_i \gamma_{ik}(x_i - \mu_k)^2}{\sum_i \gamma_{ik}}$
    Compute log_likelihood $\leftarrow \sum_i \log\left(\sum_k w_k \cdot p_k(x_i \mid \mu_k, \sigma_k^2)\right)$
    if |log_likelihood − log_likelihood_old| < tol **then**
        break
    end if
end for
Assign $\text{label}_i \leftarrow \arg\max_k \gamma_{ik}$
return $\{w_k, \mu_k, \sigma_k^2, \text{label}_i\}$

This approach allows for an unsupervised image segmentation into distinct regions based on intensity levels. The output of the segmentation process is a labeled image in which each pixel is assigned a cluster corresponding to its region. We used the Dice Similarity Coefficient (DSC) to evaluate segmentation performance, which measures the overlap between the segmented region and the ground truth bounding box. The cluster corresponding to the highest mean pixel intensity was assumed to represent the target region, as higher intensity often correlates with the area of interest. The Dice coefficient was computed as:

$$\text{Dice} = \frac{2 \cdot |\text{Predicted Region} \cap \text{Ground Truth}|}{|\text{Predicted Region}| + |\text{Ground Truth}|},$$

Where the predicted region was derived from the cluster labels and the ground truth was represented as a binary mask generated from the bounding box. This approach effectively evaluated the segmentation's accuracy.

4.2 HMM

HMMs (Joumad et al., 2023) effectively model spatial relationships and label pixels based on underlying structures, overcoming limitations of traditional methods in noisy environments or with subtle boundaries. We implement an adaptive HMM framework inspired by (Song et al., 2020), which dynamically introduces new states during segmentation to capture image complexity. This

approach uses row-by-row initialization combined with an adaptive Viterbi decoding algorithm for efficient and accurate segmentation.

Our experiments utilize images with annotated bounding boxes indicating tumor regions. Preprocessing involves enhancing contrast and reducing noise using Contrast Limited Adaptive Histogram Equalization, highlighting subtle intensity variations, and preparing images for segmentation.

egmentation begins with row-by-row initialization. The first row defines the initial states of the HMM. Each pixel is assigned to a state based on its intensity. If no state represents the intensity well, a new state is created. Emission probabilities are modeled using Gaussian distributions, with means and variances estimated from the data. Transition probabilities are set uniformly to avoid bias. For subsequent rows, the adaptive Viterbi algorithm calculates the most probable state sequence, using emission and transition probabilities for spatial consistency. The model dynamically adds states if no existing state sufficiently represents a pixel, guided by the parameter α. The pseudocode for the segmentation and evaluation process is presented in **Algorithm 2**.

The segmentation produces a labeled image where each pixel corresponds to a region. Segmentation performance is evaluated using the Dice Similarity Coefficient (DSC). Since the dataset provides bounding boxes instead of pixel-level ground truth, binary ground-truth masks are generated from bounding boxes. To refine the segmentation, clusters in the labeled image are merged iteratively, starting with the brightest regions. The Dice coefficient is calculated at each step, and merging continues until the coefficient declines. This step ensures that the segmented region aligns closely with the tumor indicated by the bounding box.

Algorithm 2 Adaptive HMM Segmentation and Evaluation

```
function PREPROCESS_IMAGE(image)
    Convert image to grayscale and enhance contrast using CLAHE
    return Preprocessed image
end function
function ADAPTIVE_HMM_SEGMENTATION(image, α, max_iter)
    Initialize states for the first row
    for each row in the image do
        for each pixel do
            Assign pixel to the most likely state or create a new state if necessary
        end for
        Update states using the adaptive Viterbi algorithm
    end for
    return Segmentation map
end function
function EVALUATE_SEGMENTATION(segmented_map, ground_truth_bbox)
    Convert bounding box to binary mask
    Iteratively merge clusters and compute Dice coefficient
    return Best Dice coefficient and refined segmentation map
end function
```

4.3 HMC

Algorithm 3 HMC and Chan-Vese Hybrid Framework

Step 1: Preprocessing and Superpixel Segmentation Convert I to grayscale, normalize pixel values, and reduce noise. Then, Segment I into K superpixels $\{S_i\}_{i=1}^{K}$ using SLIC

Step 2: MRF and HMC for Parameter Estimation

Construct graph $G = (V, E)$ with nodes V as superpixels S_i and pairwise potentials:

$$V(a_i, a_j) = \beta \cdot \exp\left(-\frac{\|f_i - f_j\|^2}{2\sigma^2}\right) \cdot \delta(a_i \neq a_j)$$

Apply posterior $P(\theta|\mathbf{a}) \propto P(\mathbf{a}|\theta) \cdot P(\theta)$ and energy function:

$$U(\theta) = -\log P(\mathbf{a}|\theta) - \log P(\theta)$$

Use HMC to sample $P(\theta|\mathbf{a})$ and estimate $\theta = \{\beta, \sigma^2\}$.

Step 3: Chan-Vese Model

Initialize contour C from segmentation map L. Optimize the Chan-Vese energy function:

$$F(C) = \lambda \cdot \text{Length}(C) + \mu \left(\int_{\text{inside}(C)} |I(x,y) - c_1|^2 dx + \int_{\text{outside}(C)} |I(x,y) - c_2|^2 dx \right)$$

Iterate min-cut refinement until convergence.

HMC is an efficient sampling algorithm that leverages the geometry of the target distribution to explore high-dimensional parameter spaces, making it suitable for image segmentation tasks. In this study, we apply combined models to brain tumor segmentation by effectively sampling the posterior distribution of segmentation parameters, enabling robust and probabilistic delineation of tumor regions.

We propose a technique that combines the Chan-Vese model(Huang et al., 2021), enhanced Markov chains for improved segmentation accuracy in brain tumor imaging, and HMC to explore parameter spaces, enabling robust convergence and uncertainty quantification, as outlined in **Algorithm 3**.

First, we perform preprocessing and initialization using Contrast Limited Adaptive Histogram Equalization (CLAHE) for contrast enhancement and Simple Linear Iterative Clustering (SLIC) for superpixel segmentation. CLAHE enhances contrast in regions with subtle intensity variations like brain tumors, making them more distinguishable. SLIC reduces computational complexity by grouping pixels into spatially adjacent superpixels based on color and intensity similarity, simplifying subsequent probabilistic modeling and segmentation while preserving boundary details.

Second, we apply spatial probabilistic modeling with Markov Random Fields (MRF) to capture spatial relationships between superpixels. This action ensures that the segmentation accounts for dependencies between neighboring regions. Initial labels are assigned based on pixel intensity and spatial features using thresholds, clustering algorithms, or pre-trained models. Parameters of the MRF are estimated using algorithms like MLE, EM, or sampling methods such as Gibbs sampling.

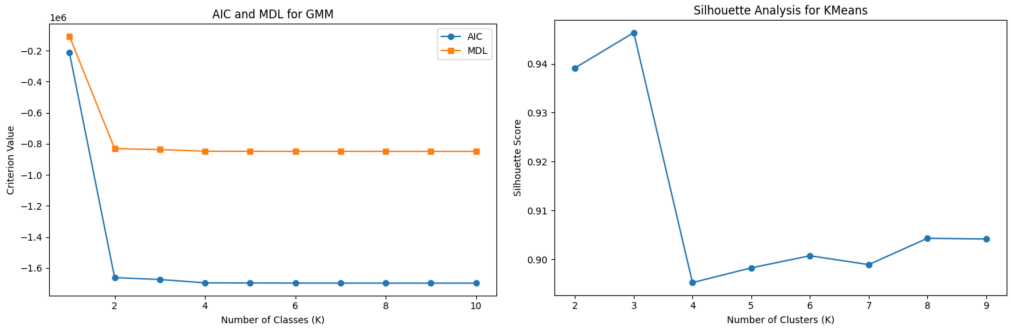
Third, we integrate with the Chan-Vese model by constructing an undirected graph with K + 2 nodes and establishing a weight map for each superpixel node. We update node values and edge weights and apply the min-cut algorithm to obtain a new segmentation. This action combines intensity-based segmentation with spatial context, yielding more accurate and context-aware

results.

Finally, we optimize and refine the segmentation by iteratively minimizing the combined models. We use iterative loops to ensure the segmentation contour evolves towards the optimal boundary between tumor and non-tumor regions.

5 Results

5.1 GMM



(a) AIC and MDL trends for determining the optimal number of clusters K.

(b) Silhouette Score analysis for identifying the optimal K.

Figure 1: Comparison of criteria for optimal clustering: (a) AIC and MDL trends suggest $K = 2$, while (b) Silhouette Score identifies $K = 4$ as the best clustering solution.

The determination of the optimal number of clusters K is illustrated in the figures fig. 1. From the AIC and MDL curves, the optimal K value is 2, suggesting that these criteria primarily distinguish between the background and the brain region. While suitable for simple segmentation tasks, this result fails to capture finer details such as the tumor. In contrast, the Silhouette Score provides a more reasonable result, identifying K=4 as the optimal number of clusters. This value better represents the structural complexity of the image.

Cluster	Weight	Mean	Variance
1	0.061663	0.223183	0.016112
2	0.779238	0.000026	0.000001
3	0.006370	0.749161	0.013064
4	0.152730	0.328559	0.001501

Table 1: Cluster statistics including Weight, Mean, and Variance.

With K=4 determined from the Silhouette Score, we applied the EM algorithm to fit a Gaussian Mixture Model (GMM) and visualize the clusters. The parameters for each cluster, including weights, means, and variances, are shown in Table 1. Among the clusters, the region with the highest mean intensity (0.749161), or cluster 4 in fig. 2, was identified as the tumor region. The segmentation results demonstrate that the GMM successfully detects the tumor area and is further validated by a Dice Similarity Coefficient (DSC) of 0.7550, indicating a substantial overlap between the detected tumor and the ground truth. These results confirm that the GMM, guided by the Silhouette Score, is highly effective for tumor detection and segmentation.

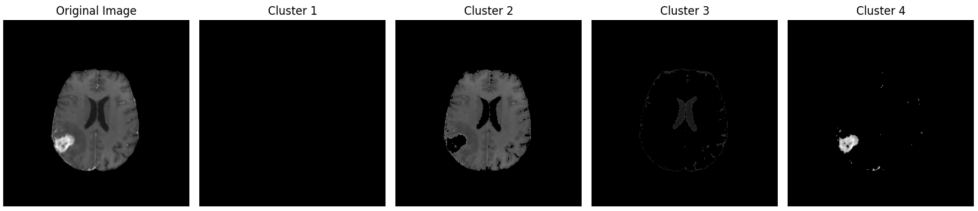
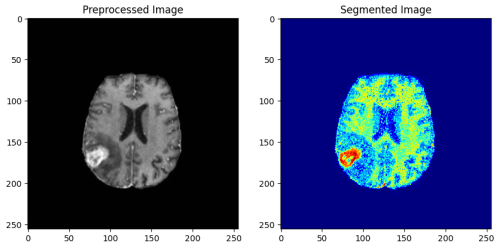


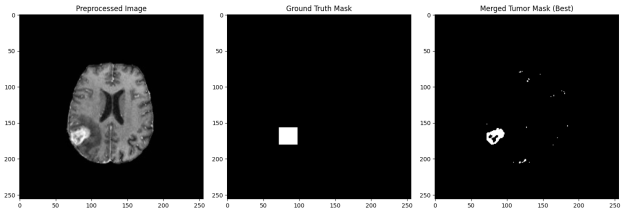
Figure 2: GMM-based segmentation results. The original image is shown alongside four clusters identified by the GMM with $K = 4$. Cluster 4, characterized by the highest mean intensity, corresponds to the tumor region.

5.2 HMM

The results of the Adaptive HMM segmentation are illustrated in fig. 3a and fig. 3b. fig. 3a shows the preprocessed brain image (left) and the resulting segmentation map (right), where each pixel is assigned a cluster label based on its intensity and spatial relationships. The segmentation highlights multiple regions of interest, including the suspected tumor area. fig. 3b compares the detected tumor cluster (right) with the binary ground truth mask generated from the bounding box (center) alongside the preprocessed brain image (left). The DSC, calculated based on the overlap between the tumor cluster and the ground truth mask, achieves a value of 0.6389, demonstrating a moderate match. While the method effectively identifies the tumor region, slight discrepancies arise due to intensity overlaps and structural complexity, which warrant further optimization.



(a) Adaptive HMM Segmentation: (Left) Preprocessed Image after CLAHE; (Right) Labeled Segmentation.

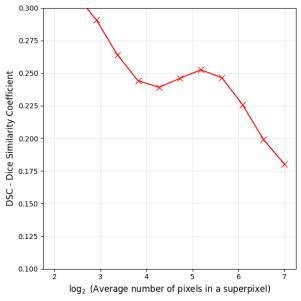


(b) Dice Similarity Coefficient Evaluation: (Left) Preprocessed Image, (Center) Ground Truth Mask, (Right) Refined Tumor Mask.

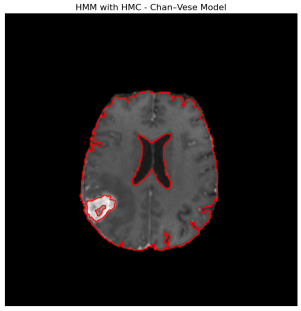
Figure 3: Segmentation and Evaluation Results of Adaptive HMM.

5.3 HMC

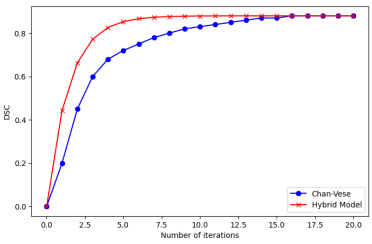
The results of HMC contain fig. 4a, fig. 4b, and fig. 4c. For fig. 4a, as the superpixel size increases, the DSC has a trend to decline, indicating a loss of segmentation accuracy due to reduced granularity. Additionally, the fig. 4b is achieved by combining the HMC framework for probabilistic parameter estimation with the Chan-Vese model for contour refinement. The red contours indicate the refined tumor boundaries detected by the method, showcasing the integration of spatial coherence and intensity homogeneity in identifying tumor and non-tumor regions. Moreover, the fig. 4c compares the convergence of the DSC for the traditional Chan-Vese model (blue) and the Hybrid HMC-Chan-Vese model (red), which both achieved a DSC above 0.8 over 20 iterations. However, the hybrid model approaches a higher DSC quickly, and the cost of computational demands is also higher.



(a) Impact of Superpixel Size on Dice Similarity Coefficient (DSC).



(b) Segmentation of Brain Tumor Using HMC and Chan-Vese Model.



(c) DSC for Chan-Vese and Hybrid HMC-Chan-Vese Models

Figure 4: Results for HMC

6 Conclusion

This study explored the performance of GMM, HMM, and a hybrid HMC-Chan-Vese model for MRI brain tumor segmentation. Each method demonstrated unique strengths: GMM excelled in efficiency and intensity-based clustering, HMM provided adaptive spatial modeling but faced challenges with cluster proliferation in complex images, and the hybrid model achieved the most accurate segmentation at the cost of higher computational demands. These findings underscore the importance of balancing accuracy and efficiency when designing segmentation algorithms. Future work could focus on enhancing the scalability of spatial models like HMM and integrating boundary refinement techniques to improve segmentation in diverse medical imaging datasets.

References

Bal, A., M. Banerjee, P. Sharma, and M. Maitra (2018). Brain tumor segmentation on mr image using k-means and fuzzy-possibilistic clustering. *2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, 1–8.

Gordillo, N., E. Montseny, and P. Sobrevilla (2013). State of the art survey on mri brain tumor segmentation. *Magnetic Resonance Imaging 31*(8), 1426–1438.

Huang, Q., Y. Zhou, L. Tao, W. Yu, Y. Zhang, L. Huo, and Z. He (2021). A chan–vese model based on the markov chain for unsupervised medical image segmentation. *Tsinghua Science and Technology 26*(6), 833–844.

Joumad, A., A. E. Moutaouakkil, A. Nasroallah, O. Boutkhoum, F. Rustam, and I. Ashraf (2023). Unsupervised statistical image segmentation using bi-dimensional hidden markov chains model with application to mammography images. *Journal of King Saud University - Computer and Information Sciences 35*(9), 101715.

Kamnitsas, K., C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker (2017). Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical Image Analysis 36*, 61 – 78. Cited by: 2614; All Open Access, Hybrid Gold Open Access.

Liang, Z., R. Jaszczak, and R. Coleman (1992). Parameter estimation of finite mixtures using the em algorithm and information criteria with application to medical image processing. *IEEE Transactions on Nuclear Science 39*(4), 1126–1133.

Shelhamer, E., J. Long, and T. Darrell (2017). Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence 39*(4), 640–651.

Song, Y., B. Adobah, J. Qu, and C. Liu (2020). Segmentation of ordinary images and medical images with an adaptive hidden markov model and viterbi algorithm. *Current Signal Transduction Therapy 15*(2), 109–123.

Wadhwa, A., A. Bhardwaj, and V. Singh Verma (2019). A review on brain tumor segmentation of mri images. *Magnetic Resonance Imaging 61*, 247–259.